Louisiana State University
LSU Digital Commons

LSU Doctoral Dissertations | Graduate School

6-15-2018

Focality in Prospective Memory: A Thorough Examination of Cue Specificity and Task-Appropriateness

Bethany Alice Lyon
Louisiana State University and Agricultural and Mechanical College, bethanyalyon@gmail.com

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_dissertations

Part of the Cognitive Psychology Commons

Recommended Citation
Lyon, Bethany Alice, "Focality in Prospective Memory: A Thorough Examination of Cue Specificity and Task-Appropriateness" (2018). *LSU Doctoral Dissertations*. 4623.
https://digitalcommons.lsu.edu/gradschool_dissertations/4623

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Doctoral Dissertations by an authorized graduate school editor of LSU Digital Commons. For more information, please contactgradetd@lsu.edu.

FOCALITY IN PROSPECTIVE MEMORY:
A THOROUGH EXAMINATION OF CUE SPECIFICITY
AND TASK-APPROPRIATENESS

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Psychology

by
Bethany Alice Lyon
B.A., Augustana College, 2012
M.A., Louisiana State University, 2014
August 2018

ACKNOWLEDGMENTS

My completion of this dissertation, and all of the training prior to this point would not have been possible without the support and encouragement from many important individuals. Each of the people mentioned below, as well as countless others have helped me succeed and reach this point in my life, and for that I owe each and every one of them a sincere thank you. It is difficult to express how great an impact each person has had on my success and overall wellbeing throughout my academic training building up to the completion of this dissertation, but I feel certain that I would not be here without their support, encouragement, and advice throughout the years. It is incredible the difference that a support system can make in one individual's success, and I feel like my support system is one of the best I've ever witnessed. All of these individuals have encouraged me through the discouraging portions of my training. They have celebrated with me during my successes. They have helped me persevere through the daily grinds of academia, and I cannot thank them enough!

To Dr. Jason Hicks- my continued progress in the program would not have been possible without your support, time, and mentorship. Thank you for taking me on as a student. Your patience and thoughtfulness as an advisor have directed me towards a path for a successful and happy future. I am humbled that you took the time to sort through numerous research ideas and iterations with me; and allowed and encouraged all of the spontaneous meetings when I felt slightly panicked about my progress. Thank you for all of the times you sat down with me and talked through snags and questions throughout the entirety of this dissertation. I cannot express how much your guidance, knowledge, and support have meant to me these past six years, I am incredibly grateful for all of the efforts you have put forth to see me through to success, professionally and academically. I feel proud to have you as an advisor!

To Dr. Katie Cherry- I would not be where I am today without the training, support, and guidance you offered me throughout the years we worked together at LSU. I learned so much from working with you and I'm very glad I had the opportunity to work on the diverse research projects you had in the lab. You have played a large part in shaping me into the researcher and educator I am today, thank you. I wish you all the success in your future projects!

To Dr. Emily Elliott- Thank you for your guidance as an educator! I appreciate the time and energy you put into your work, and all of the effort you put in to help and encourage me throughout my training at LSU. You were always a source of support, whether it be during a committee meeting, a brown bag presentation, or class evaluations, or how you structured your feedback in a positive manner and worked to build me up. Thank you!

To Dr. Janet McDonald- I greatly appreciate the time and energy you have invested in my successes, especially during the past few years! I know I would not be where I am now, both academically and professionally, without your involvement. You were always available to listen when issues came up, especially in navigating the stresses of graduate training. I'm grateful for your guidance in how to handle some of those tougher situations that came up. Thank you so much for all of your feedback and insight on this dissertation!

To Dr. Celinda Reese-Melancon- Thank you for being a member of my committees! I appreciate all of the guidance you have given me, both over the phone and at CAC, especially in regards to professional writing. Thank you for carrying out all of the long-distance communications for committee meetings- your feedback and insights have been invaluable for improving my research projects. Thank you!

To Dr. Adam Melvin- I feel so grateful to have you as a member of my committee. I feel very fortunate that you have invested so much time in my success through my training. Your feedback regarding writing and organization have helped re-structure how I approach creating academic papers and figures. Thank you!

To Mr. Tim, Ms. Nancy, Ms. Wendy, and Ms. Ruby- Thank you for all of the help behind the scenes! Everything that you have done for me during these past six years has helped to make my time at LSU run smoother. You guys form a wonderful team that I've always felt encouraged by, even if it was as small as fixing the copy machine, talking about dogs for a small break in the day, rushing paperwork on my behalf, or allowing me to check out countless reference books for weeks at a time. Thank you for all that you have done for me, and for your positivity!

To Willie and Jessica- Thank you for your time and energy needed in getting this dissertation data collected. Thank you for your efforts in testing out all of the different iterations of E-Prime programs to find problems and glitches before we ran participants. I would likely still be collecting participants had it not been for the two of you working together to run people through the experiments all Fall! Thank you both for your enthusiasm for the research, I appreciate both of you!

To Keri- Thank you for being a friend and source of support for the past few years! I always knew at conferences that I would have a friend to share in the excitement, and I am glad we share that together. Your drive throughout grad school is inspiring. Your determination has helped push me to work harder and move through all of the various steps needed to navigate the last year of grad school successfully. While our training has been trying at times, I feel stronger and better able to get through it knowing you are right there with me. Thank you!

To Gerald Smith, Christel Sanchez, Dale Matthews, Mawmaw Gayle, and Pawpaw Floyd- Thank you for the numerous meals you shared with me, the support, love, and encouragement you offered me, and the renewed love for Louisiana that you gave to me. My training at LSU would not have been nearly as enjoyable as it has been without you guys! You have welcomed me in as part of your own family, and that means so much to me. Thank you!

To AP, Uncle Tom, Aunt Jan, Grandma Urick, Aunt Darlene, Aunt Sue, Uncle Steve, Aunt Tracy, Uncle Willy, and Aunt Teresa, and all of the rest of my extended family- Thank you for your support! My time in Louisiana has been far from family; but keeping in touch online and over the phone has helped to make me feel not as far away from everyone. Thank you for your help after the flood a few years back, for your encouragement during some of my major academic milestones, and for cheering me on throughout my life! Thank you!

To Kacie, Danielle, and Becky- I have so many positive memories of my time in Louisiana largely thanks to the three of you. You are always willing to read over my papers, listen to my ideas, grab lunch, or just hang out together. I cannot imagine what grad school would’ve been like without you three here with me. Thank you for your friendship and support! Love you all!

To Caryn, Julia, and Becca- I greatly appreciate each of you for staying so close, despite the very far geographical distance between us. It makes me so happy to catch up with each of you, even if it is just a few texts back and forth or a Facebook chat. Those small bits of communication helped me through some very exhausting days, and for that I’m thankful. You’ve supported and cheered me on throughout school and visited with me on breaks, and I feel so lucky to have friends like you guys! I cherish the memories of times spent with each of you- from Becca visiting the deep south, to meeting with Jules in Nashville, to Thanksgiving with

Caryn- each of you have shared your time, love, and friendship with me and supported me throughout my academic training. Thank you!

To Dina- Thank you for your friendship and support throughout our time in Louisiana! You have seen me through some horrible rough drafts of basically every paper I completed since we met. I know you always have my back, whatever I need(ed) you are always there to help out, lend an ear, commiserate, or celebrate. I look back at all the memories we shared in grad school fondly, knowing that we really support each other 100%. I appreciate all that you have done for me; and feel so grateful to have you as a friend!

To Lauren- Thank you for being my number one fan since forever! I always knew if I needed a boost of encouragement I could call you. I appreciate all of the love and support you and your family have given me over the years. You have always been there for me through thick and thin, and it really means the world to me to call you my friend. Thank you for allowing me to talk your ear off about school stuff, sharing your home and meals with me during long road trips home, and offering honest advice about life. Love you tons!

To Kristin, Steph, Rae, and Nat- You guys are the best supporters I could’ve ever hoped to have in my corner! I always love catching up on the phone, and even though my day to day activities were nearly always the same sort of school related work, you always asked about my progress and took an interest in what I was working on at that moment. Hearing about your lives helped me to stay sane throughout grad school. Visiting with each of you and your families was always a refreshing and invigorating break! It always brightens my day to see one of you calling, even if it is just for a bit during your commute or to catch up quick. Thank you for always keeping in touch and offering words of encouragement through the rough times and the whirlwind times. Special shout out to Josh, Kevin, Adam, Yordany, Elizabeth, Lauren, Thomas, Joey, Maggie, Willy, Andrew, Koda, and Moose for sharing all of the happy moments in your lives with me- it brings so much joy to my life! Love you all so much!

To Mom and Dad- Where do I even being to thank the two of you? You have been my support even before coming to Louisiana. You have both encouraged me each step of the way towards the completion of my Dissertation, even on the few detours in my training. You have logged so many hours of phone calls with me full of support, love, and strength to carry on and push myself harder. The help and love you have shown me throughout my life have helped build me into who I am today. Thank you from the bottom of my heart.

To Dillon- you have always been in my corner supporting me and encouraging me to carry on with my work and research. Your love and patience throughout these last few years is immeasurable, and I cannot tell you how much that means to me. I had no idea that I would meet my best friend in Louisiana, but I am so glad that I did! Thank you for always being a voice of reason, support, and encouragement throughout these last three years, it really means the world to me! You've never doubted that I could make it through this training process, even when I've doubted myself. You've witnessed the daily stress and toll the intense training program has had on me for years and continued to support me through it all. I'm so grateful for your time, energy, love, encouragement, patience, and countless meals. Thank you so much.

TABLE OF CONTENTS

ABSTRACT

Prospective memory (PM) refers to memory for future intentions (e.g. remembering to press a button when you see an animal word). Researchers classify PM intentions in the laboratory as focal or nonfocal primarily in two ways. One way, task-appropriateness, refers to how the processing for the intention relates to the processing required for an ongoing task; the PM intention and the ongoing task either match in processing (TAP; focal) or mismatch in processing (TIP; nonfocal). The second way researchers classify focality is based on cue specificity, with the PM task either being specific (focal) or general (nonfocal). Resolving this ambiguity in defining "focality" is the focus of the current research. These two experiments compared the roles of cue specificity and task-appropriateness in PM focality. Participants were randomly assigned to a control group, a focal PM condition (TAP/Specific), or one of three nonfocal conditions (TAP/General; TIP/Specific; TIP/General). Their ongoing task required them to make a semantic judgment (Experiment 1) or an orthographic judgment (Experiment 2). Cue specificity impacted PM accuracy consistently, favoring specific cues. Task-appropriateness impacted PM accuracy in Experiment 1 as predicted, but not in Experiment 2 which showed protective effects for specific, whole-word PM cues. Ongoing task performance mostly followed predicted patterns (no differences for ongoing task accuracy) and suggested ex-gaussian analyses offered a nuanced measure of RT data. These studies highlight the existing ambiguity in the operational definition of focality and provide the groundwork for a clear definition of focality centered around cue specificity and task-appropriateness.

CHAPTER 1. INTRODUCTION

Rationale

A great deal of memory research conducted in the last 30 years focuses on the formation and execution of future intentions. This area of research—prospective memory (PM)—incorporates all of the processing required to realize a delayed intention as well as the intentions' associated actions (Ellis, 1996), including the encoding of the intention, its retrieval, and its execution (Kvavilashvili & Ellis, 1996). Defined in more detail later, in essence PM involves the creation of a goal or intention that a person cannot implement immediately; the intention is delayed and the person must retrieve it at the appropriate time or place by interrupting whatever task he or she is pursuing at that time. For example, a PM intention could be remembering to email information to a colleague the following day. PM is distinct from many other memory research fields in that the retrieval is both self-initiated (Anderson & Craik, 2000; Einstein et al., 2012) and delayed after formation. During the delay period, participants complete some other task (coined the ongoing task) which occupies the delay period to prevent rehearsal (Burgess et al., 2011). PM researchers study diverse intentions by classifying them in several ways, such as by factors specific to the intention itself (e.g. tied to a time of day or tied to an event occurring), factors relating to the location of intention execution (e.g. in a laboratory or in a naturalistic setting), factors that alter a person's strategy for noticing the PM retrieval opportunity (e.g. relying on the intention to "pop" into mind or active, effortful monitoring), or factors that relate in some manner to the current cognitive processes in use for the ongoing task. This last factor—the match or mismatch in cognitive processing and whether the PM cue was in the focus of attention or not—oftentimes determined if a researcher would classify the PM task as focal or nonfocal (Einstein & McDaniel, 2005). The current research examined this classification process to clarify what makes a PM task focal and which factor(s) within focality play(s) the largest role on task performance.

In order to fully understand what focality is in PM research and uncover which features are most important for efficient and successful PM execution, the first section covers general background information about PM. The second section examines the operational definition of focality and what past work has determined about this key distinction via behavioral PM research. The third section moves into a brief discussion of key PM theories to inform how people carry out intentions. The fourth and final section reviews the two primary components involved in determining task focality—cue specificity and task-appropriateness—then moves into how researchers have studied these key features in the past and how the current study extends our understanding of these features and helps untangle what it means to be a focal or nonfocal PM intention.

PM General Background Information

One major distinction in PM research is whether the intention relates to a particular time or an event (Einstein & McDaniel, 1990). Most research teams examine event-based PM intentions (e.g. press a button when you see the word "lemon" or if you see a picture of a face with glasses). The use of time-based intentions rather than event-based intentions is less common in the laboratory (e.g., press a key after five minutes has passed; see Kvavilashvili & Ellis, 1996 for further elaboration). Generally, PM performance is better on event-based PM tasks compared to time-based tasks (Henry et al., 2004; Marsh, Hicks, & Cook, 2006; Sellen, Louie, Harris, & Wilkins, 1997). The current research primarily focused on event-based intentions and the findings that stem from those experiments.

All PM intentions have a similar structure at their core. Kliegel et al. (2011) classified this structure into four separate stages in order to uncover the underlying cognitive processes at work. The first stage is *intention formation*. During intention formation, people use encoding and planning processes to formulate their future intention, and the “what, where, and how” planning details necessary for successful execution. The second stage is *intention retention*, which involves accessing both long-term memory and working memory to store the intention. The third stage is *intention initiation* during which one begins to carry out the PM intention at the appropriate time. This step involves a variety of cognitive processes including monitoring, cognitive flexibility, and inhibition, which work together for detecting the PM cue and realizing there is an intention to carry out. The fourth stage, *intention execution*, is the last stage and also involves cognitive flexibility (Kliegel et al., 2011; Hering et al., 2014) and allows for the person to complete the PM intention. Successful PM requires completion of each stage, and any missed stage(s) results in failure to complete the PM task.

To help participants achieve PM success and complete each of Kliegel et al.’s (2011) stages, some researchers suggest increasing the depth of processing—possibly through use of imagery, elaborate encoding, or semantic processing—to strengthen the intention’s encoding (Craik & Rose, 2012; Hering et al., 2014). Increasing the depth of processing is thought to require more cognitive resources initially during encoding but can lead to more efficient use of resources long term, as deeply encoded intentions are thought to have more automatic retrieval (Hering et al., 2014).

Other factors—individual differences, PM and ongoing task demands, and the importance of the PM and ongoing task—can also determine how deeply a person decides to process and encode an intention (Craik & Rose, 2012; Hering et al., 2014). Hering et al. (2014) found that investing resources in deeper processing to encode an intention can actually lessen the overall amount of resources needed to successfully complete the PM task by reducing monitoring behavior. In contrast to this, increasing the importance of a PM task may elevate the amount of cognitive resources devoted to the PM task by increasing the level of monitoring used to detect a PM cue (Meeks & Marsh, 2010). Essentially, telling a participant that the task is highly important encourages them to monitor and possibly focus less on the ongoing task, while increasing the depth of processing for the task may give participants the perception that the task will be easy and automatic, and encourage them to reduce their monitoring and focus their efforts on the ongoing task. Thus, learning about when each cognitive process is involved during PM can provide insight into how to make PM more efficient. Individual differences in monitoring preferences, features of the PM cue, and the nature of the ongoing task can all impact the type and amount of processing people decide to use to complete their PM intention (Einstein et al., 2012).

In summary, PM tasks can be event-based or time-based. All PM tasks follow the same four stages—intention formation, intention retention, intention initiation, and intention execution—which a person must complete for PM success. PM researchers often focus on methods and strategies to improve performance at each stage of the intention, mindful of individual differences and the importance of the PM task. Another concept that researchers can use to improve PM performance is focality, described extensively in the next section.

PM Focality Background

This section defines focality, both in terms of cue specificity and task-appropriateness. Additionally, this section covers how focality interacts with different age groups and how it can impact monitoring strategies. This provides a foundation for what it means for a task to be focal or nonfocal and what factors may play a role in this pivotal categorization.

McDaniel and Einstein (2000) were the first to classify an intention based on its focality. Essentially, they define a focal PM intention as a specific PM cue that is within a participant's focus of attention during the completion of the ongoing task (Einstein et al., 2005; McDaniel & Einstein, 2000; McDaniel et al., 2015). Einstein et al. (2005) see focal tasks as those that have a high level of overlap in processing with the ongoing task (i.e. task-appropriate processing or TAP), which may lead to more automatic retrieval. For example, if a researcher asks participants to determine if a string of letters forms a word or nonword (classical lexical decision task) for the ongoing task, a focal PM task might require the participant to press a different key when they see the word "horse" on the screen. The ongoing task requires cognitive processing of the stimuli presented on the screen as words versus nonwords, so the PM task of identifying a particular word as relevant to an intention provides a high degree of processing and feature overlap. A nonfocal task requires processing outside of the ongoing task in order to detect the PM cue. For example, if the ongoing task was a lexical decision task, a nonfocal PM intention might require the participant to press a key when they see *any* animal word. The ongoing task requires processing the stimuli to determine if it is a word or nonword, but the PM intention would require further processing to classify the word as an animal or not in order to complete the PM task successfully. According to Einstein et al. (2012), (non)focality can be defined through the physical location of a stimulus (e.g. the ongoing task focuses on the center of the screen while the PM cue appears in the screen periphery) or the processing feature that the ongoing task centers on (e.g. the ongoing task centers on the meaning of a word while the PM task centers on a letter within a word). In a meta-analytic review of PM focality, Kliegel et al. (2008) found overall higher PM performance on focal tasks compared to nonfocal tasks. Kliegel et al.'s (2008) finding suggests real differences exist between focal and nonfocal PM tasks and the processing required for successful task completion. Importantly, the definition of focality presented here emphases the existing confound present between the specificity of the PM cue and the processing match/mismatch with the ongoing task. Many researchers have used cue specificity ("horse" vs. *any* animal word) as a match or mismatch in processing and have claimed this difference in specificity as a difference in focality.

Generally speaking, PM researchers view task-appropriate processing (TAP) as the hallmark for a PM cue to be considered focal, but cue specificity is commonly used as a stand in for task-appropriateness. The overlap in processing between what the PM task requires and what the ongoing task requires is another major consideration for researchers in determining whether a task is focal or nonfocal. Scullin et al. (2010) defined focality as the shared processing between the ongoing and PM tasks, highlighting the overlap in processing as the primary element of importance in determining if an intention is focal or nonfocal. Differences in PM task focality may require different cognitive processes to complete (Einstein et al., 2012). Some tasks, like a focal task, might require very few processing resources beyond that required by the ongoing task. Other tasks, primarily nonfocal tasks, might require constant monitoring to not miss the PM cue (Einstein et al., 2012).

Some experiments strengthen the focality dichotomy with findings showing that focal tasks are distinct and different from nonfocal tasks. How focality impacts PM performance in special populations, especially older adults, offers support for the distinction between focal and nonfocal tasks. Mullet et al. (2013) found no age-related differences in PM cue detection when they tested younger and older adult groups on a focal PM task, but significant differences in PM detection between the age groups when they tested participants on a nonfocal PM task. Reese-Melancon (2013) found age-related differences in both focal and nonfocal PM cue detection, with larger PM deficits for the older adult age group when tested in the nonfocal condition compared to focal. Similarly, Uttl (2011) studied focality and age-related differences in PM in a meta-analysis. He found large age-related differences in PM cue detection for both focal and nonfocal PM tasks with worse performance in the older adults; however, PM performance was not different between the focal and nonfocal conditions, possibly indicating that focality does not impact PM performance as drastically as previously thought. This finding may also suggest participants use different strategies to overcome the differences in PM task demands, possibly relying on assessments of task difficulty to adjust their monitoring strategy.

The performance differences between focal and nonfocal PM tasks have stimulated experiments focused on what strategies—or resource allocation policies—people use to complete their intentions. Generally, differences between focal and nonfocal task performance may be due to differences in the difficulty level of cue monitoring (Scullin et al., 2010), as it is more difficult to monitor for something that mismatches in processing from the ongoing task. Hicks, Franks, and Spitler (2017) demonstrated reduced, and in some cases eliminated, task interference for the ongoing task in nonfocal conditions when the PM cue from an animal category was the same exemplar repeated multiple times. This suggests that part of the difference between focal and nonfocal tasks is an attentional allocation policy that people set (Hicks, Marsh, & Cook, 2005). In standard conditions, people seem to strategically allocate more resources to nonfocal tasks and less resources to focal tasks. McDaniel et al. (2015) suggest researchers could discourage participants from monitoring for the PM cue by designing their PM study using a single, event-based, focal PM cue, with processing features directly relevant to the processing required to complete the ongoing task—which defines the best-case scenario for producing a focal PM cue. McDaniel et al. (2015) also suggest that minimizing the importance of the PM task, which guides participants away from using a monitoring strategy as much as possible. As monitoring for the PM cue to appear is not always an efficient method to complete the PM task, many researchers believe this strategy reduces performance on the ongoing task and depletes attentional resources (Smith et al., 2003; Marsh et al., 2003).

Overall, focal and nonfocal tasks place different demands on the participants, interact with the ongoing task processing requirements to differing degrees, and stimulate different strategies for PM success. Focal PM tasks can lead to different outcomes in monitoring and PM performance than nonfocal tasks. Understanding how these two halves of focality differ and how researchers have studied them previously helps to untangle what exactly it means to be focal or nonfocal, and sheds light on the possibility of gray areas in this dichotomy. The following section elaborates on the theoretical understanding of PM and how each theory relates to differences in focality.

PM Theoretical Background

PM researchers have formulated several theories to explain how a person carries out PM tasks. Each of the theories discussed explains how people go about completing PM tasks and how having the PM intention can impact performance on the ongoing task. This section covers four PM theories: Preparatory Attentional and Memory Theory (PAM theory; Smith, 2003; Smith & Bayen, 2006; Smith, 2016), the multiprocess framework (Einstein & McDaniel, 2005; McDaniel & Einstein, 2000, 2007), the dynamic multiprocess framework (Scullin et al., 2013), and the Delay Theory (Loft & Remington, 2013). The purpose of this section is not to pit these theories against one another, but to demonstrate what each theory contributes to understanding focality.

The PAM theory (Smith, 2003; Smith & Bayen, 2006; Smith, 2016) originated to explain slowing in the ongoing task reaction times (RTs) when a PM intention was also active. The amount of attentional resources spent on the PM task are therefore taken away from the resources devoted to completing the ongoing task, creating a deficit in ongoing task performance—namely, cost or the interference effect (Jäger & Kliegel, 2008; Marsh et al., 2003; Smith, 2003). The PAM theory explains ongoing task costs through the conjecture that once a person has an active intention, that person has to devote attentional resources to the intention in order to monitor for the PM cue. Smith (2003) labeled the attention now devoted to monitoring for the PM cue as preparatory attention, and it sets the person into a monitoring mindset that may be outside of conscious awareness. Importantly, the level of monitoring and attentional resources spent on the PM task can vary with the task demands, importance level of the task, etc. Marsh et al. (2003) conducted a series of experiments exploring how holding an intention creates interference for the ongoing task. Their findings suggest that costs to the ongoing task act like a metric to gauge how much monitoring is in use to detect the PM intention (Marsh et al., 2003). Since that finding, many other studies have drawn conclusions about the level of monitoring used to detect the PM cue based on the presence or absence of interference effects found on the ongoing task (see Hering et al., 2014; Smith, 2016). Measuring ongoing task costs has become a major feature of PM studies, but especially so in studies manipulating focality to estimate monitoring differences between focal and nonfocal conditions. The PAM theory would argue that monitoring differences are always present with both focal and nonfocal tasks when compared to a control condition. In a series of experiments, Smith, Hunt, McVay, and

McConnell (2007) used highly salient PM cues (e.g. participants' names) embedded in a lexical decision task, to test if costs are still present in a focal, salient condition. They found the control group without an intention had faster RTs than the focal, PM group (Smith et al., 2007). The PAM model therefore suggests that regardless of the focality of an intention, researchers should see costs in the ongoing task when compared to a control condition, possibly with differences in the magnitude of monitoring between focal and nonfocal conditions.

Some aspects of the PAM theory are also a part of the multiprocess framework (Einstein & McDaniel, 2005; McDaniel & Einstein, 2000, 2007). The multiprocess framework similarly contends that a person with an intention can actively (consciously) monitor their environment for the PM cue, and that this monitoring requires attentional resources. This framework also suggests that ongoing task costs would be observable under this situation (McDaniel & Einstein, 2005; 2007). However, the framework has a second possible process aside from active monitoring that people can use to carry out PM intentions: spontaneous retrieval (automatic, without conscious awareness). According to Einstein and McDaniel (2005), the spontaneous retrieval process is based on the Reflexive-Associative Theory, which posits that the target intention is linked with an environmental cue, and once the cue registers, the intention is automatically retrieved. Spontaneous retrieval is often described in the PM literature as the sensation of the intention "popping into mind" at the correct time to execute the task (McDaniel & Einstein, 2007). Importantly, if spontaneous retrieval is the aim, using a focal task is the best bet to avoid monitoring (McDaniel et al., 2015). In regards to focality, the multiprocess framework suggests there is a large difference in monitoring costs and performance when comparing focal to nonfocal tasks. If a task is focal, it should activate spontaneous retrieval mechanisms, leading to greater PM cue detection with no detectable monitoring costs.

Conversely, nonfocal tasks require more active monitoring for successful PM detection, leading to higher costs to the ongoing task and worse PM performance.

The dynamic multiprocess framework builds off of the multiprocess framework, and retains the same two processes, active monitoring and spontaneous retrieval, as the possible routes to successful PM execution (Scullin et al., 2013). Additionally, the dynamic multiprocess framework takes the environmental context into account. The environmental context at the time of retrieval may spontaneously pull the cue-intention pair into conscious awareness, triggering intention initiation and execution (Einstein & McDaniel, 1996). The key aspect of the dynamic multiprocess view is that people can adapt to the surrounding contexts and form expectations about the intention that inform whether they rely on active monitoring or spontaneous retrieval to complete the task (Scullin et al., 2013). This theory also allows for flexibility in which process people use at which time, postulating that people can switch between the two processes whenever they see fit to do so (Scullin et al., 2013). Scullin et al.'s (2013) work closely examines changes in ongoing task costs throughout the duration of the study and shows fluctuations in monitoring behaviors based on proximity to the onset of the task and the presence of PM cues. These patterns may differ depending on a person's expectations of the task difficulty, which is highly relevant when exploring focality. For example, in focal conditions participants are often told exactly what stimuli (e.g. the word "lemon") they should respond to in order to complete their intention. This specificity may alter participants' perceptions of the task's difficulty (i.e. this will be easy!) compared to people in a nonfocal condition who are told to respond when they see any stimuli that belongs to a certain category (e.g. when you see a fruit word). In the nonfocal condition, participants may perceive this task to be difficult—as the cue could be any fruit—leading them to increase the amount of monitoring for the PM cue so as not to miss a possible fruit word. The increased monitoring leads to greater ongoing task costs. As mentioned previously, when participants gain experience with cue detection, their perceptions of task difficulty can adjust to reduce monitoring costs (see Hicks et al., 2017; Hicks et al., 2005) reflecting the flexibility Scullin et al. (2013) promote with the dynamic multiprocess framework. In regards to focality, the dynamic multiprocess framework mirrors that of the multiprocess framework with a large difference in monitoring costs and performance when comparing focal to nonfocal tasks. Focal tasks should trigger spontaneous retrieval mechanisms, and nonfocal tasks should require more active, intentional monitoring. However, this theory also predicts that participants' expectations and experience with the tasks could drastically alter monitoring behavior.

Each of the previously mentioned theories hinges on the concept of a limited attentional capacity system in which people must split their attention between their ongoing task and the PM task (Heathcote, Loft, & Remington, 2015; Loft & Remington, 2013; Strickland, Heathcote, Remington, & Loft, 2017) such that more attention spent on one task results in a cost to the other task. The Delay Theory is different from these capacity-based theories (Loft & Remington, 2013). The Delay theory postulates that the primary reason behind increased RTs during the ongoing task is a result of an increase in the response threshold needed in order to make a decision (Loft & Remington, 2013; Heathcote et al., 2015; Strickland et al., 2017). With every decision that a person makes, he or she is accumulating evidence in support of each possible decision outcome. Once the person has accumulated enough evidence in support of one of the decisions, that is the option selected. In the case of a lexical decision ongoing task, for example, the person accumulates evidence that the string of letters they see is either a word or a nonword. Once enough evidence has accumulated for one of those options, the person makes the appropriate selection. When a person has an intention as well as an ongoing task, the Delay Theory posits that the amount of evidence needed to make the ongoing task choice increases—the person sets a higher threshold that must be met before making a choice (Heathcote et al., 2015). The rate at which that the person amasses evidence for either choice should remain stable, with only the threshold raised (Heathcote et al., 2015; Strickland et al., 2017). The RT slowing on the ongoing task in the Delay Theory is strategic in that it allows for a gain in time to respond to the PM task, rather than an attentional process (Rummel et al., 2016). Thus, it implies that people literally delay their ongoing task decision to accommodate the possibility of needed intention retrieval. If the delay is controlled (as was the case in Loft & Remington, 2013), differences between focal and nonfocal conditions should be found in any delay conditions shorter than 600ms. If the delay is longer than 600ms, focal and nonfocal conditions should result in equivalent performance. In situations when a delay is not forced, participants with an intention may adjust how much additional time they need to allow for their PM decision, which may lead to differences between focal and nonfocal tasks (though this supposition is speculative based on general patterns in PM empirical research that do not mention delay theory specifically).

In further support of the Delay Theory, Heathcote et al. (2015) reanalyzed RT data from Lourenço et al.'s (2013) study that compared cost effects when participants were told the PM cue would appear only in a word versus participants that were told the PM cue could appear in either a word or a nonword. Heathcote et al. (2015) surmise that the capacity-based theories (PAM, multiprocess framework, and dynamic multiprocess framework) would not predict item-specific effects of costs— the processing demands brought on by the addition of an intention should impact both words and nonwords. The capacity-based theories would predict either equal costs across word and nonword trials, or more costs in the nonword trials because the participant should be giving more attention to word trials to monitor for the PM cue. Lourenço et al. (2013) found contrasting RT patterns; costs were nonsignificant in the nonword trials but present in the word trials. This finding was especially pronounced in the condition when participants were told the PM cue would only appear in word trials compared to those told the cue could appear in word or nonword trials (Lourenço et al., 2013; Heathcote et al., 2015). The Delay Theory predicts thresholds are set differently for words and nonwords, as well as for focal PM intentions compared to nonfocal PM intentions (Strickland et al., 2017). Importantly, Strickland et al. (2017) concluded that when studying costs associated with intentions it is imperative to look at the shape of the RT distribution in conjunction with error rates, which would portray trading speed for accuracy. Strickland et al.'s (2017) recommendation to look at the entire RT distribution instead of just the means is elaborated on more fully in later portions of this section.

The major theories covered in this section center around costs to the ongoing tasks through RTs, and what those costs mean for monitoring strategies used to complete the PM task. Each of these theories views focality slightly differently, from having a small, minimal impact on performance and cost (PAM theory), to having large influences on performance and cost (multiprocess framework and dynamic multiprocess framework), to a functional view of increased RTs rather than a cost or detriment to the ongoing task performance (Delay Theory).

One other aspect of theoretical interest is the ex-Gaussian (exponentially modified Gaussian) distribution, useful for analyzing RT distributions. The ex-Gaussian function is a positively skewed normal distribution, with three key parameters: the mean of the Gaussian portion of the RT distribution (μ), the standard deviation (σ) of that Gaussian portion, and the mean and standard deviation of the exponential distribution (τ), or the tail skew distribution

(Brewer, 2011; Cohen & Hicks, 2017; Heathcote, Popiel, & Mewhort, 1991). Ex-Gaussian modelling does not make any predictions about the underlying cognitive processing but has previously been interpreted in PM research with regards to monitoring (Brewer, 2011). Brewer (2011) interpreted a shift in the entire distribution (μ) to mean that a person is actively monitoring for the PM cue on every single trial of the ongoing task. If there was only a change in the τ parameter, Brewer (2011) took that to mean that the participant monitored for the PM task in brief instances, reflected by an increase in trials with longer RTs. Brewer (2011) went on to say that a change in both μ and τ might indicate a general shift in attention from the ongoing task to the PM task, suggesting that these ex-Gaussian parameters may offer insight into PM processing occurring during the ongoing task (Cohen & Hicks, 2017). Analyses of RT distributions as a whole with ex-Gaussian modeling can differ from the traditional RT metric of just examining RT means (Cohen & Hicks, 2017). Analyzing RT data as a distribution with three parameters allows for a closer inspection of the costs to the ongoing task, and whether the interference is a broad overall slowing (changes in μ), or a slowing sporadically on few select trials (changes in τ). This method of analyses is not just a methodological tool but can help by providing evidence for how ongoing task interference manifests in different experimental conditions, which may in turn lead to more precise PM theories to explain changes in RTs when completing an intention.

An In-Depth Examination of Focality

Focality stems from two main principles, which are used throughout PM research, cue specificity and task-appropriateness. When researchers define a focal task, they most often narrow in on these two features—the cue must be specific or it must match in processing with the ongoing task (see McDaniel et al., 2015). Interestingly, very few studies have manipulated these two principles in a controlled experiment pitted against each other to examine the full contrast between focal and nonfocal PM tasks (see Tables 1-4). Tables 1-4 provide a representative summary of PM research and how scientists have examined focality. Table 1 shows general PM studies for comparison. These studies typically only compare an intention group to a non-intention control group. The right side of each table provides a classification for the PM intention based on whether the PM intention is specific and task-appropriate, general and task-appropriate, specific and task-inappropriate, or general and task-inappropriate. Table 2 summarizes studies that compare a focal to a nonfocal PM condition; the studies in this table define focality by whether the intention is specific or general in nature. Table 3 displays PM studies that define focality by whether the intention is task-appropriate or task-inappropriate, with no regard to specific and general cues. Finally, Table 4 is a compilation of studies pertinent to the current study—studies that compare intentions on both cue specificity and task-appropriateness. Careful manipulation of both of these critical factors in focality helps to establish if focality behaves like a dichotomy or more like a continuum in how it impacts both PM and ongoing task performance. The following section offers more detailed coverage of prior work exploring focality through manipulations of cue specificity and task appropriateness. Additionally, the section covers the levels of processing principal, which is often used to create focal and nonfocal conditions.

Table 1. PM Studies Classified by use of Cue Specificity and Task-Appropriateness.

Studies Focused on PM Performance Broadly							
Citation	Ongoing Task	PM Task	Conclusions	Specific/ TAP	General/ TAP	Specific/ TIP	General/ TIP
McDaniel & Einstein (1993).	Exp. 1: Short term memory task	Exp. 1 Press a key when you see a certain word. Cues were high and low levels of familiarity.	Exp. 1: Unfamiliar cues lead to more correct PM retrievals than familiar PM cues.	Event-based			
	Exp. 2: Short term memory task	Exp. 2: specific PM cues high or low in familiarity and high and low in distinctiveness from ongoing task.	Exp. 2: Distinctive cues lead to better performance than non-distinctive cues. Unfamiliar cues lead to better PM performance than familiar cues.	Event-based			
Maylor (1996).	Name each famous face.	Circle the trial number if a person is wearing glasses.	Compared middle aged adults to older adults. Found that middle aged adults were less likely to forget the PM task, and more likely to recover on the next PM cue than older adults.	Event-based			
West & Craik (1999).	Exp. 1: Semantic categorization matching task to words written in lowercase grey.	Exp. 1: Press a key when you see a pair of words in lowercase green font or capitalized grey font. Ignore lures where only one word in the pair is in lowercase green or capitalized grey.	Exp. 1: more older adults fell for the lures than younger adults and had worse PM performance.				Event-based
McGann et al. (2002).	Exp. 3: Rate words for pleasantness OR Rate words for readability	Exp. 3: Press a key when you see any of four specific words	Task was carried out under full or divided attention. The pleasantness rating was affected by the divided attention task but readability was not affected. Overall performance was worse under divided attention.	Event-based Event-based			

Table 1 continued.

Citation	Ongoing Task	PM Task	Conclusions	Specific/ TAP	General/ TAP	Specific/ TIP	General/ TIP
Einstein et al. (2005).	Exp. 3: Semantic Decision (does the last word fit into the given sentence?)	Exp. 3: Press a key when you see a specific word OR Press a key when you see one of six specific words	Exp. 3: PM performance was comparable in both conditions. Ongoing task performance showed slower RTs when also carrying out a PM task, but only in the 6-cue condition.	Event-based Event-based			
	Exp. 4: Semantic Decision (does the last word fit into the given sentence?)	Exp. 4: Press a key when you see a specific word	Exp. 4: Individual differences noted in costs, signifying some people use monitoring strategies more than other for the same task.	Event-based			
Hicks et al. (2005).	Lexical Decision	Press a key when you see a word in a red font OR Press a key when you see a red border surrounding a word.	Borders and letters were manipulated in salience with a large and small version. Participants in the border condition responded more often to the large border. Participants in the word color condition responded to the PM cue equally across all font sizes and border sizes.			Event-based Event-based	

Table 2. PM Studies Using Cue Specificity as the Defining Feature of Focality.

Studies Comparing General Cues to Specific Cues							
Citation	Ongoing Task	PM Task	Conclusions	Specific/ TAP	General/ TAP	Specific/ TIP	General/ TIP
Einstein et al. (1995).	Exp. 2: recall the last 10 words from a continuous memory task	Exp. 2: Press a key when you see the words Leopard, Lion, Tiger OR Press a key when you see an animal word.	Exp. 2: Specific PM cues lead to PM performance that was better than general cues	Event-based	Event-based		
Ellis & Milne (1996).	Exp. 1: Answer true/false common sense semantic knowledge questions	Exp. 1: Press a key when you see certain words (either highly related cues or low related cues) OR Press a key when you see a word in the given category (either highly related or low related)	Exp. 1: Cue specificity mattered with better overall PM performance when given specific cues. An interaction revealed that the benefit of cue specificity only matters when given highly related cues.			Event-based	Event-based
	Exp. 2: Answer true/false common sense semantic knowledge questions	Exp. 2: Press a key when you see certain words (Apple) OR Press a key when you see words belonging to a specific category (subordinate: Fruits) OR Press a key when you see words that belong to a general category (superordinate: Foods)	Exp. 2: Specific cue PM performance was roughly higher than either of the general PM cues, but only significantly higher than superordinate PM cues (Foods)			Event-based (Apple)	Event-based (subordinate Fruits) Event-based (Superordinate Foods)

Table 2 continued.

Citation	Ongoing Task	PM Task	Conclusions	Specific/ TAP	General/ TAP	Specific/ TIP	General/ TIP
Ellis & Milne (1996).	Exp. 3: Answer true/false common sense semantic knowledge questions	Exp. 3: Press a key when you see certain words (Rose) OR Press a key when you see words belonging to a specific category (subordinate: Fruit, Flowers, Trees) OR Press a key when you see words that belong to a general category (superordinate: Plants). Cues were either typical or atypical exemplars.	Exp. 3: The specific cues lead to higher PM performance than either general cue. The difference between specific and general cues was greater for atypical cues than for typical exemplars.			Event-based (Rose)	Event-based (subordinate Fruits, etc.) Event-based (Superordinate Plants)
Marsh et al. (2003).	Exp. 1: Lexical Decision	Exp. 1: Press a key when you see an animal word OR Press a key when you see the word DOG	PM performance was better with the specific cue compared to the general cue	Event-based	Event-based		
Cona et al. (2013).	Lexical Decision	Press a key when you see a specific word (focal) OR Press a key when you see a word from a category (nonfocal)	PM accuracy was high for both focal and nonfocal conditions. ERP results show increased monitoring in nonfocal condition for all participants. Only some participants recruited monitoring resources for the focal condition.	Event-based	Event-based		

Table 2 continued.

Citation	Ongoing Task	PM Task	Conclusions	Specific/ TAP	General/ TAP	Specific/ TIP	General/ TIP
Loft & Remington (2013).	Exp. 1 & 2: Lexical Decision	Exp. 1 & 2: Press a key if you see a specific word OR Press a key if you see a word from a certain category OR control condition with no intention all within-subjects factor. All conditions asked to wait to make response until after a tone.	Exp. 1: At 400ms delay, focal PM accuracy was better than nonfocal. In all other time delay conditions, no differences found in PM conditions. Exp. 2: Focal PM accuracy was greater than nonfocal PM accuracy during the 0ms, 200ms, and 400ms delay conditions, but equal in the 600ms condition.	Event-based	Event-based		
Hicks et al. (2016).	Exp. 1: Lexical Decision	Exp. 1: Press a key when you see a specific word (focal) OR when you see an animal word (nonfocal; new animal word for each exposure).	Exp. 1: PM accuracy was higher in the focal condition.	Event-based	Event-based		
	Exp. 2: Lexical Decision	Exp. 2: Press a key when you see a specific word (focal) OR Press a key when you see an animal word (nonfocal; used same animal word for all 6 exposures).	Exp. 2: PM accuracy was overall higher in the focal condition. Those who completed the focal block first, their performance on the nonfocal task was equal to focal performance. Those who completed the nonfocal block first had worse performance in nonfocal compared to focal.	Event-based	Event-based		
Strickland et al. (2017).	Lexical Decision	Press a key when you see a certain word OR Press a key when you see a word from a certain category	PM accuracy was better for the focal condition compared to the nonfocal condition. Costs indicate changes in thresholds rather than mean drift rates.	Event-based	Event-based		

Table 3. PM Studies Using Task-Appropriateness as the Defining Feature of Focality.

Studies Comparing Task-Appropriate to Task-Inappropriate							
Citation	Ongoing Task	PM Task	Conclusions	Specific/ TAP	General/ TAP	Specific/ TIP	General/ TIP
Einstein et al. (1995).	Exp. 3: Answer general knowledge and problem-solving questions	Exp. 3: Press a key when you see a question about presidents OR Press a key every 5 minute	Exp. 3: PM performance was better in the event-based task compared to the time-based task.	Event-based		Time-based	
Marsh et al. (2000).	Exp. 1: Rate items for pleasantness (sem.) OR identify words with repeated letters (ortho.)	Exp. 1: Press a key when you see an animal word (sem.) OR Press a key when you see a palindrome (ortho.)	Exp. 1: TAP lead to better PM task performance.		Event-based (sem. ongoing-sem. PM) Event-based (ortho ongoing-ortho PM)		Event-based (ortho ongoing-sem. PM) Event-based (sem. ongoing-ortho PM)
	Exp. 2: Rate items for pleasantness (sem.) AND identify words with repeated letters (ortho.)	Exp. 2: Press a key when you see an animal word (sem.) OR Press a key when you see a palindrome (ortho.)	Exp. 2: TAP lead to better PM task performance.		Event-based (sem. ongoing-sem. PM) Event-based (ortho ongoing-ortho PM)		Event-based (ortho ongoing-sem. PM) Event-based (sem. ongoing-ortho PM)

Table 3 continued.

Citation	Ongoing Task	PM Task	Conclusions	Specific/ TAP	General/ TAP	Specific/ TIP	General/ TIP
Marsh et al. (2000).	Exp. 3: Rate items for pleasantness (sem.) AND identify words with repeated letters (ortho.)	Exp. 3: Press a key when you see an animal word (sem.) OR Press a key when you see a palindrome (ortho.)	Exp. 3: PM cues were presented in brackets to increase cue salience. No significant interaction for PM, overall higher PM performance.		Event-based (sem. ongoing-sem. PM) Event-based (ortho. ongoing-ortho. PM)		Event-based (ortho. ongoing-sem. PM) Event-based (sem. ongoing-ortho. PM)
Einstein et al. (2005).	Exp. 1: Word categorization task	Exp. 1: Press a key when you see a certain word OR Press a key when you see a certain syllable	Exp. 1: PM performance was better in the focal (word cue) condition than the nonfocal (syllable condition)	Event-based		Event-based	
	Exp. 2: Word categorization task	Exp. 2: Press a key when you see a certain word OR Press a key when you see a certain syllable	Exp. 2: PM performance was better in the focal (word cue) condition than the nonfocal (syllable condition). Performance from first to last PM cue declined, especially in nonfocal condition.	Event-based		Event-based	
Marsh et al. (2005).	Exp. 1: Lexical Decision	Exp. 1: Press a key when you see an animal word OR Press a key when you see a palindrome OR no intention	Exp. 1: Both intention groups had slower RTs than the control group for the ongoing task. Under low effort, PM was better in the TAP condition. Under high effort on the ongoing task, the TAP PM cue suffered, while the palindrome (TIP) PM detection was not impacted.		Event-based (sem.)		Event-based (ortho.)

Table 3 continued.

Citation	Ongoing Task	PM Task	Conclusions	Specific/ TAP	General/ TAP	Specific/ TIP	General/ TIP
Marsh et al. (2005).	Exp. 2: Identify double contiguous letters in words (ortho.)	Exp. 2: Press a key when you see an animal word OR Press a key when you see a palindrome OR no intention	Exp. 2: Under low effort, PM was better in the TAP condition. Under high effort on the ongoing task, the TAP PM cue suffered, while the animal (TIP) PM detection was not impacted.		Event-based (ortho.)		Event-based (sem.)
	Exp. 3: Lexical Decision	Exp. 3: Press a key when you see an animal word OR Press a key when you see a palindrome (in all capitals for salience)	Exp. 3: Under low effort, PM was better in the TAP condition. Under high effort on the ongoing task, the TAP PM cue suffered, while the palindrome (TIP) PM detection was not impacted.		Event-based (sem.)		Event-based (ortho.)
Meiser & Schult (2008).	Lexical Decision	Press a key for any animal word OR press a key for any palindrome	Better PM performance for the TAP condition than the TIP condition.		Event-based		Event-based
Scullin et al. (2010).	Exp. 3: Lexical Decision	Exp. 3: Press a key when you see a specific word OR Press a key when you see a word that starts with a specific letter.	Exp. 3: more ongoing task interference in the first letter condition. No differences in PM accuracy.	Event-based		Event-based	
	Exp. 4: Long Lexical Decision	Exp. 4: Press a key when you see a specific word OR Press a key when you see a word that starts with a specific letter.	Exp. 4: monitoring interference was found in the first letter condition, but it decreased across trials. PM performance was much greater in the focal, word condition than the nonfocal, first letter condition	Event-based		Event-based	

Table 3 continued.

Citation	Ongoing Task	PM Task	Conclusions	Specific/ TAP	General/ TAP	Specific/ TIP	General/ TIP
McBride & Abney (2012).	Count the number of vowels, OR count the number of syllables, OR living/non-living judgments	Press a key for words containing repeated consecutive vowels	Vowels condition had higher PM accuracy than both other nonfocal conditions (not different from each other)		Event-based (count the vowels) Event-based (count the syllables)		Event-based (living/nonliving)
Abney et al. (2013).	Exp. 1: Lexical Decision	Exp. 1: Respond to animal words (semantic) OR respond to palindromes (orthographic)	Task-Appropriate had better PM accuracy than Task-Inappropriate (also faster RTs on the ongoing task for TAP compared to TIP)		Event-based		Event-based
	Exp. 2: Living/ Nonliving task (sem.) OR Consecutive Letter task (ortho.)	Exp. 2: Respond to animal/bird words (sem,) OR respond to palindromes (ortho.)	Semantic PM had greater accuracy compared to orthographic PM. Task-appropriate had greater performance than TIP for both ongoing task conditions.		Event-based Event-based		Event-based Event-based
Loft & Remington (2013).	Exp. 3: Lexical Decision	Exp. 3: Press a key if you see a specific word OR Press a key if you see a specific syllable OR control condition with no intention. Manipulated as within subject factor. All conditions asked to wait to make response until after a tone.	Exp. 3: Focal PM accuracy was greater than nonfocal PM accuracy during the 0ms, 600ms, and 1,000ms delay conditions, but was equal in the 1600ms condition.	Event-based		Event-based	

Table 3 continued.

Citation	Ongoing Task	PM Task	Conclusions	Specific/ TAP	General/ TAP	Specific/ TIP	General/ TIP
McDaniel et al. (2013).	Semantic categorization	Press a key when you see a certain word OR Press a key when you see a certain syllable	PM performance was equally high in both conditions. Monitoring costs occurred in both conditions but to a greater degree for the nonfocal condition. The aPFC was linked to the precuneus during nonfocal tasks and to the middle temporal gyrus during the focal task.	Event-based (word)		Event-based (syllable)	
Zuber et al. (2016).	n-back task: Is this letter the same letter that was presented 2 letters ago?	Press a key when you see a specific letter (focal) OR Press a key when you see a specific colored border (nonfocal)	PM performance was comparable for focal and nonfocal tasks. Focal and nonfocal PM are distinct but related concepts.	Event-based		Event-based	

Table 4. PM Studies using both Cue Specificity and Task-Appropriateness.

Section 4: Studies Comparing Multiple Components of Interest							
Citation	Ongoing Task	PM Task	Conclusions	Specific/ TAP	General/ TAP	Specific/ TIP	General/ TIP
Meier & Graf (2000).	Decide if word represents something natural (sem.) OR decide if word has two or more enclosed spaces (perceptual)	Press a key when you see an animal word (sem.) OR Press a key when you see a word with 3 e's (per.)	PM performance was best in the TAP conditions with processing match. No mention that PM tasks differed on general and specific domains.	Event-based (per. ongoing-per. PM)	Event-based (sem. ongoing-sem. PM)	Event-based (sem. ongoing-per. PM)	Event-based (per. ongoing-sem. PM)
West & Craik (2001).	Exp. 1: Classify a word into one of four font colors (per.) OR Classify a word into one of four semantic categories (Sem.)	Exp. 1: Press a key for blue uppercase (per.) AND Press a key for red uppercase (per.) AND Press a key for a tuber-told specific cues (Semantic) AND Press a key for a building part-told specific cues	Exp. 1: The TAP/TIP interaction was not found. There was a main effect of ongoing task with higher performance on the semantic ongoing task. There was also a main effect of PM cue with more of the perceptual cues noticed overall.	Event-based (sem.)	Event-based (per.)	Event-based (sem.)	Event-based (per.)
	Exp. 2: Classify a word into one of four different font colors (per.) OR Classify a word into one of four semantic categories (Sem.)	Press a key for blue uppercase or red uppercase (per.) AND Press a key for two specific words (sem.)	Exp. 2: The semantic ongoing task lead to greater PM performance overall compared to the perceptual ongoing task. Perceptual PM cues were responded to more often than the semantic PM cues. The interaction was significant with TAP predictions.	Event-based (sem.)	Event-based (per.)	Event-based (sem.)	Event-based (per.)

Table 4 continued.

Citation	Ongoing Task	PM Task	Conclusions	Specific/ TAP	General/ TAP	Specific/ TIP	General/ TIP
Hicks et al. (2005).	Exp. 1: Lexical Decision	Exp. 1: Press a key when you see an animal word OR press a key after 4 minutes has elapsed	Exp. 1: Participants more often remembered to fill the time-based intention rather than event-based intention.		Event-based	Time-based	
	Exp. 2: Lexical Decision	Exp.2: Press a key when you see an animal word, press a key when you see this specific animal word, press a key when 3-5 minutes and 7-9 minutes has passed, OR press a key after 4 minutes has elapsed. Focus is on Cue specificity.	Exp. 2: Time-based memory was better than event-based. No main effect of specificity. No interaction.	Event-based	Event-based	Time-based	Time-based
	Exp. 3: Lexical Decision	Exp. 3: Press a key when you see an animal word AND press a key when 3-5 minutes and 7-9 minutes has passed OR Press a key when you see this specific animal word AND press a key after 4 minutes has elapsed. Focus on cue specificity. Have 2 intentions.	Exp. 3: Some evidence that general cues might lead to worse performance.	Event-based	Event-based	Time-based	Time-based
Meeks & Marsh (2010).	Exp. 1, 2, & 3: Lexical Decision	Exp. 1: Press a key when you see an animal word Exp. 2: Press a key when you see a word with the syllable TOR in it. Exp. 3: Press a key when you see the words DEER or COW.	PM performance was at ceiling for Exp. 3, PM performance was close to 70% for the animal cues and 45% for the syllable cue. Comparisons were not made across experiments.	Event-based (Exp. 3)	Event-based (Exp. 1)	Event-based (Exp. 2)	

Exploring Cue Specificity

Many PM studies use cue specificity to distinguish a focal PM intention from a nonfocal PM intention. In these cases, the focal PM cue is specific—researchers tell participants the exact word(s), syllable, color, or pattern that they must find to complete the PM task. Nonfocal PM cues are general—researchers tell participants to respond to *any* animal word, or *any* word in a particular category (see Table 2). Cue specificity has a long history of use to study focality, starting with Einstein et al.'s (1995) study. Einstein et al. (1995) found that their specific cue led to better task performance than their general PM cue. Ellis and Milne (1996) reported similar findings in their three different experiments, mirrored by Marsh et al. (2003), Hicks et al. (2016), and Strickland et al. (2017). Cona et al. (2013) used the cue specificity distinction as their focality manipulation in their event related potential (ERP) experiment. They found PM accuracy was high for both intention conditions, but increased monitoring in the nonfocal, general PM condition. Only a portion of their participants in the focal, specific condition recruited monitoring resources to complete their PM task (Cona et al., 2013). Similarly, Loft and Remington (2013) compared PM performance on a specific, focal intention and a general, nonfocal intention. However, they introduced a delay before responses were allowed to see if this created less conflict in resource allocation between the ongoing task versus the PM task, resulting in better PM performance. In their two experiments, Loft and Remington (2013) found that the focal condition (compared to the nonfocal condition) had better PM accuracy after delays less than 400ms were introduced, but at all other longer delay lengths accuracy was equivalent across focality conditions. They took these findings as evidence that the delay freed up resources to process the nonfocal PM task, negating the focality differences present during the shortest delays. In summary, specific, focal PM cues lead to better PM performance than general, nonfocal PM cues.

Exploring Task-Appropriateness

Task-appropriateness is the other dimension PM researchers consider to determine an intention's focality. Task-appropriateness is the term to describe the overlap in processing between the ongoing task and the PM task (Meier & Graf, 2000; Maylor, 1996; Marsh et al, 2000). When the ongoing task and the PM task match in the processing they require, the PM task is said to engage task-appropriate processing (TAP) and is considered focal (Maylor, 1996; Meier & Graf, 2000). When the ongoing task and the PM task mismatch in the processing they require, the PM task uses task-inappropriate processing (TIP) and is considered nonfocal (Maylor, 1996; Meier & Graf, 2000). Researchers have considered TAP at a neurological level as a similarity in brain processes (Maylor, 1996; Marsh, Hicks, & Hancock, 2000), or a similarity in activation patterns in the frontal neocortex between PM encoding and retrieval (Craik & Rose, 2012; Mayes & Roberts, 2002); however, TAP also extends to the environmental context surrounding the PM task, which is any content aside from the PM target itself—including the ongoing task (Marsh, Hicks, & Cook, 2008; McDaniel, Robinson-Riegler, & Einstein, 1998; see also Godden & Baddeley, 1975 for importance of shared context outside of PM). Numerous PM studies have been carried out using task-appropriateness as the defining feature of focality (see Table 3). Typically, PM studies using task-appropriateness as the defining feature in determining focality use tasks that differ in their level of processing. Normally this means using one semantic PM task and one orthographic PM task. In the following section, I discuss research pertaining to levels of processing and the differences between semantic and orthographic tasks, followed by a look at experiments that have used the semantic/orthographic distinction to define TAP and TIP intentions in PM.

Levels of Processing

A difference exists between orthographic information and semantic information in how much cognitive processing people need to encode the features in question (Craik & Lockhart, 1972). Semantic processing is anything related to the meaning or definition of a word or item. Orthographic processing relates to the features of the text or image, its color, sizing, which letters are a part of the word, the shape of the letters, etc. Any information that people encode at a deeper level, helps to strengthen their eventual retrieval of that information (Craik & Lockhart, 1972). Typically, people place more value on the semantic nature of an item, thus many of the perceptual orthographic features are processed shallowly and fade in memory quickly (Craik & Lockhart, 1972; Craik & Tulving, 1975).

Importantly, Morris et al. (1977) found that the distinction between semantic and orthographic information—or the difference in the level of processing—is still subject to task-appropriateness in a retrospective memory test (Morris, Bransford, & Franks, 1977). Morris et al. (1977) had participants encode words based on the semantic meaning of the word or by determining if it rhymed with a certain word. When tested in a standard recognition test, semantic encoding lead to better memory performance, but when tested in a rhyming recognition test the rhyming encoding lead to better memory performance than the semantic encoding (Morris et al., 1977).

Seamon and Virostek (1978) went so far as to have participants order the difficulty—or depth—of mental processing required to answer questions about a word shown to them. People rated the orthographic questions as the shallowest, least effortful questions to answer; they rated the semantic questions as the most difficult and requiring the deepest processing to answer (Seamon & Virostek, 1978). The authors later found those ratings predicted memory performance with more difficult, deeper processing questions leading to better memory performance while the shallow, orthographic questions lead to worse memory performance. They concluded that it is important to think about depth of processing as a continuum, rather than discrete classifications between orthographic and semantic, as these features are elements of one item (Seamon & Virostek, 1978). Researchers use these classic levels of processing findings in their PM studies to ensure that TAP is truly different from TIP, by having PM tasks that are either semantic or orthographic relating to either a semantic or orthographic ongoing task. The next section includes further elaboration of these key findings.

All in all, PM research in focality centers around two key principles, cue specificity and task-appropriateness. Cue specificity refers to whether the PM cue is specific (e.g. “horse”) or general (e.g. *any* animal word). Task-appropriateness refers to whether the processing required to complete the PM task matches (focal) or mismatches (nonfocal) the processing required to complete the ongoing task. Oftentimes, researchers establish task-appropriateness by using differing levels of processing—namely semantic and orthographic processing—for the ongoing and PM tasks to create a focal condition and a nonfocal condition. Several researchers used the levels of processing distinction between semantic and orthographic processing to create different focality conditions, which is covered more thoroughly in the following section. Tables 1-4 classify a selection of PM experiments based on which feature(s) the researchers used to define focality in their studies.

PM Research Manipulating Task-Appropriateness through Levels of Processing

Many studies have explored the TAP/TIP differences in PM research, with some of these studies also discussing the classifications as focal or nonfocal, but most with a focus on task-appropriateness in and of itself (see Table 3). Several examples of these types of studies are described more thoroughly. The key element tying these studies together is their use of different levels of processing to create TAP and TIP conditions, which researchers often consider focal and nonfocal conditions, respectively.

In a set of three experiments, Marsh et al. (2000) compared two different ongoing tasks as a between-subjects manipulation—one semantic and one orthographic task—while participants had to complete either a semantic or an orthographic PM task. They found that whenever the processing matched between the ongoing task and the PM task, PM cue detection was higher. Einstein et al. (2005) assigned participants to either a focal, specific word PM cue—a semantic task—or a nonfocal, specific syllable PM cue— an orthographic task. They found better PM cue detection in the semantic, word task, which was TAP in relation to their ongoing task (Einstein et al., 2005). Marsh et al. (2005) carried out a series experiments exploring task-appropriateness in PM, while also manipulating the amount of effort needed to complete the ongoing task. Participants had either no intention, a TAP intention that matched the semantic ongoing task, or a TIP intention that was orthographic in nature rather than semantic. They found that the benefits of TAP for PM cue detection were only present when the demands on the ongoing task were low, rather than high. This lead Marsh et al. (2005) to conclude that when a person had a surplus of attentional resources available after meeting the demands of the ongoing task, task-appropriateness impacted PM performance with greater performance on TAP intentions rather than TIP intentions. Scullin et al. (2010) found comparable PM accuracy across their TAP and TIP conditions but did find more ongoing task costs present for the TIP condition (for similar findings see McDaniel et al., 2013 and Zuber et al., 2016). Collectively, these studies make the case that using different levels of processing to create focality conditions successfully results in conditions that are either TAP or TIP and can lead to performance differences on the PM task as well as the ongoing task RTs. Only a few studies take the levels of processing distinction and extend their work more in-line with the current exploration of focality; these studies are discussed in more detail next.

Task-Appropriateness in Concurrence with Cue Specificity

The studies discussed in this section primarily investigated task-appropriateness and focality. The majority of these studies unintentionally venture closer to examining cue specificity in conjunction with task-appropriateness through the PM tasks they selected for use in their experiments. Researchers designed many of these studies to examine other experimental manipulations as their primary goal; interestingly, these studies happen to align with the current cross between cue specificity and task-appropriateness that determine whether researchers should classify a task as focal or nonfocal. With close inspection of these experiments, it feels clear that a controlled, intentional manipulation of cue specificity and task-appropriateness is necessary, as it may help explain some of the results found in these studies discussed next.

One such study (Meier & Graf, 2000) was designed to examine PM under differing conditions in task-appropriateness. Meier and Graf (2000) carried out a PM study examining semantic and perceptual (orthographic) ongoing and PM tasks that matched or mismatched in processing (see Table 4). This study is unique in that the conditions of the study fall into the four established focality classifications based on cue specificity and task-appropriateness established previously in this document. The authors focused on manipulating the ongoing task to match in processing with a perceptual PM task (specific cues: respond when you see words with 3 letter e's) or a semantic PM task (general cues: respond when you see any animal word). The researchers told the participants exactly which letter to look out for in the perceptual instance,

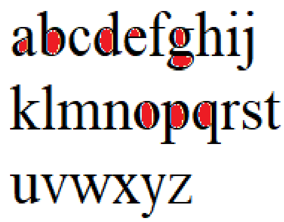

Figure 1. Enclosed Spaces used in Experiment 2, previously used by Meier and Graf (2000).

making this a specific cue, albeit a specific cue embedded within a word stimulus (similar to when researchers ask participants to look for a certain syllable). The perceptual ongoing tasks used in this experiment required participants to count the number of enclosed spaces in the letters of words and determine if there were more or fewer than 2 spaces (see Figure 1). The semantic ongoing task required participants to decide if a word was naturally occurring or manmade. Meier and Graf (2000) found that PM performance was best when the ongoing task and the PM task matched (TAP), indicated by a significant interaction between the two ongoing tasks (perceptual and semantic) and the two PM tasks (perceptual and semantic). Both of the TIP conditions had very low PM accuracy (roughly 38%). Interestingly, the interaction found in this experiment can be reclassified as an interaction between cue specificity and task-appropriateness, as shown in Figure 2. The interaction suggests that cue specificity may only impact PM accuracy under TAP conditions. Task-appropriateness in this instance was the major factor influencing performance, as both of the TIP conditions were very poor. However, the authors also manipulated cue specificity, so a closer examination of this variable in the interaction is needed

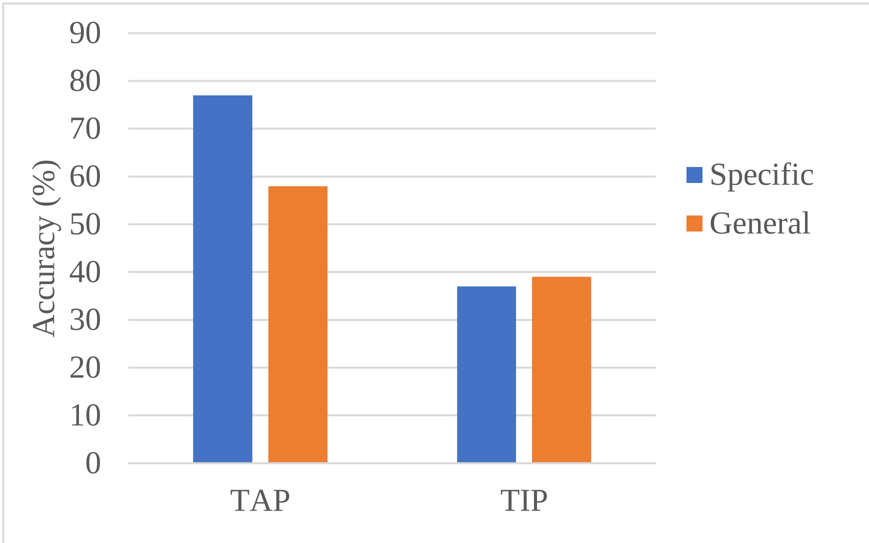

Figure 2. Meier and Graf (2000) PM Accuracy.

in order to draw further conclusions. Importantly, the TIP and TAP conditions were unbalanced in terms of cue specificity. The semantic PM task used was general, meaning that Meier and Graf's (2000) TAP/semantic condition differed on two dimensions—cue specificity and level of processing—from their TAP/perceptual condition (which had a specific PM cue); this unbalance could help to explain the drop in PM detection from the perceptual condition to the semantic condition when both matched in processing with the ongoing task. Of their two TIP conditions, one used a specific PM cue (the semantic ongoing task paired with the perceptual PM task) and the other used a general PM cue (the perceptual ongoing task paired with the semantic PM task); both TIP conditions had equally poor PM detection, which may suggest that cue specificity does not benefit PM performance under mismatching processing conditions. However, this speculation has yet to be examined further empirically.

Another set of experiments that incidentally manipulated cue specificity while studying task-appropriateness are those of West and Craik (2001). West and Craik (2001) used a similar study design comparing semantic and perceptual ongoing and PM tasks with the focus being to compare task-appropriateness in two experiments (see Table 4). Their ongoing tasks involved classifying words based on their font color or classifying words based on their semantic category, established as a between-subjects manipulation. During the first experiment, the researchers asked the participants to respond to two perceptual PM cues (press a key for a blue or red uppercase word) and two semantic PM cues (press a key for any building parts or tubers, e.g. hallway or potato) each presented twice for a total of eight PM cue presentations. They did not find the typical TAP advantage over TIP in this study, but they did find that performance was better on the semantic ongoing task, as well as better PM performance on the perceptual PM cues. The researchers may have found this result due to the extremely high salience of the perceptual PM cue (written in all capital letters) which easily stood out when compared to the lowercase ongoing task words. Marsh et al. (2000) has argued that cue salience can override task-appropriateness benefits for matching conditions. West and Craik (2001) combatted the effect that cue salience may have had on their findings by including other, capitalized lure words that were not part of the PM task.

In the second experiment, the participants again completed both the semantic tasks and the perceptual/ orthographic tasks. The researchers did find the expected benefits for the TAP conditions, but they also replicated their first experiment by finding greater overall performance on the semantic ongoing task and greater PM accuracy on the perceptual PM cues (West & Craik, 2001). These findings suggest that cue salience can impact PM performance immensely (i.e. it is important to use PM cues of equal salience across all comparison groups whenever possible) and that researchers need to consider levels of processing, as semantic ongoing tasks have an advantage compared to perceptual/ orthographic tasks in accuracy and speed. The researchers did not mention focality as a part of their experimental design, however their PM tasks differed in terms of cue specificity as the researchers provided the specific cues in the semantic task, but the perceptual capital letter words were general. This means that along with manipulating task-appropriateness, the researchers also confounded task-appropriateness with cue specificity (West & Craik, 2001). Their interaction found in their second experiment is displayed in terms of cue specificity and task-appropriateness in Figure 3. This interaction is not what one would typically expect to see in a manipulation of cue specificity and task-appropriateness, as the general cues overall (both perceptual PM in these experiments) had higher levels of PM accuracy, likely due to the high salience levels of the colored word cues.

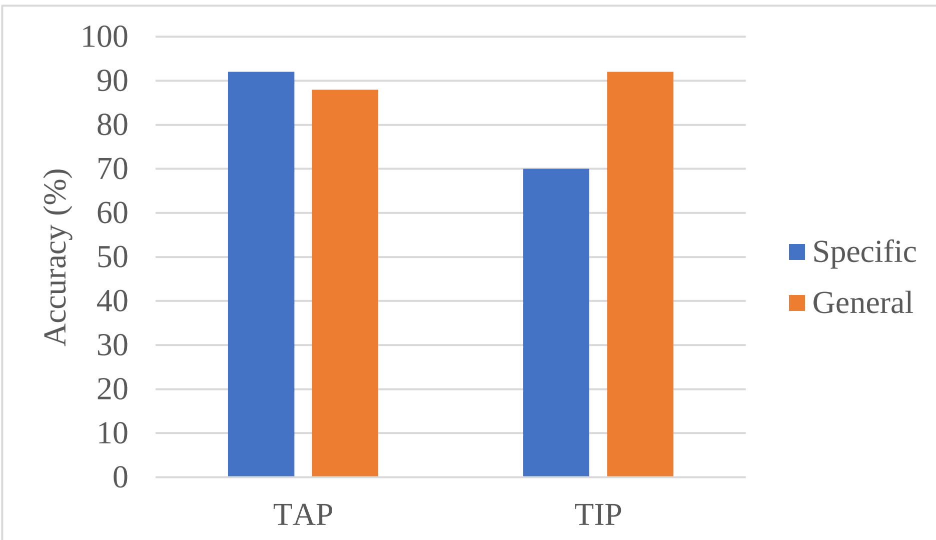

Figure 3. West and Craik (2001) PM Accuracy.

Again, this experiment shows the frequent confounding of cue specificity with task-appropriateness and the need for a clearer depiction of the role it plays in focality.

Several other studies have come close to crossing cue specificity with task-appropriateness, though most of these experiments only partially addressed cue specificity, with their primary interest in manipulating other variables. For instance, Hicks, Marsh, and Cook (2005) examined the differences between an event-based PM cue and a time-based PM cue (see

Table 4). The selected PM cues differed both in terms of cue specificity and task-appropriateness. They found better PM performance on the specific, time-based task, which may indicate that cue specificity is a stronger indicator of performance than task-appropriateness. After purposefully manipulating cue specificity, Hicks, Marsh, and Cook (2005) concluded that general cues may lead to worse performance, further strengthening the notion that cue specificity is an important factor to consider in focality. Similarly, Meeks and Marsh (2010) had a primary interest in exploring the impacts that implementation intentions (a memory strategy; see Gollwitzer, 1999) have on PM performance, especially under different focality conditions (see Table 4). Interestingly, they describe focal cues as being specific with processing matching that of the ongoing task; they describe nonfocal cues as not salient and mismatching in processing required by the ongoing task, with cues referring to an entire category (general) or a specific syllable within a word (TIP)—classifications of focality that support the current research. Meeks and Marsh classified their PM conditions based on cue specificity, but not in terms of task-appropriateness; establishing their three conditions broadly in terms of focal (specific cues) and nonfocal (general cues). While Meeks and Marsh did not make direct comparisons across their three experiments, they did convey that PM performance was at ceiling levels for the focal (specific, TAP) cue in experiment 3, near 70% for the nonfocal (general, TAP) cue in experiment 1, and close to 45% for the nonfocal (specific, TIP) cue used in experiment 2. These studies manipulated cue specificity and task-appropriateness as they sought to study other variables impacting PM performance.

A few other recent studies have purposefully attempted to examine task-appropriateness and focality. McBride and Abney (2012) examined both task-appropriateness and focality, though in a different way than the current studies, which views task-appropriateness as a defining construct within focality. McBride and Abney see task-appropriateness as different from the focal/nonfocal distinction because of how they define the overlap in processing with the ongoing and PM tasks. For these authors, task-appropriateness requires similar processing between the ongoing task and the PM task; in contrast, whether an ongoing task encourages or discourages processing of the defining features of the PM cue determines the PM cue's focality. McBride and Abney note that within the PM literature, previous work has confounded task-appropriateness with focality, so they set out to create conditions that were either focal or nonfocal and TAP or TIP (ignoring cue specificity). They used three different ongoing tasks: count the number of vowels in a word, count the number of syllables in a word, or determine if the word is living or nonliving. All participants had the same PM task—respond when they see any word with repeated consecutive vowels, but without specifying specific vowel letters. McBride and Abney classified their three tasks differently than the current classification of the study in Table 3. Namely, McBride and Abney determined that their ongoing task that asked participants to count the number of vowels was TAP and focal while their ongoing task of counting the number of syllables was TAP and nonfocal.

This view ignores cue specificity as a defining feature in determining the focality of a PM cue and instead relies on whether the ongoing task forces the participant to directly process features of the PM cue in order to complete the ongoing task (also see McDaniel & Einstein, 2000). As the key elements in focality are cue specificity and task-appropriateness based on the general consensus in the literature, the classification in Table 3 considers all three of their conditions nonfocal, due to the nature of the PM task being general: respond to *any* double vowels. The discrepancy in classification is especially important, as it sheds light on the subjective nature in PM focality classifications based on the researchers' objectives. McBride and Abney (2012) did find a difference in PM accuracy when the ongoing task was counting the vowels present in a word, which implies the idea that focal tasks may not represent one category of a dichotomy, but instead varying levels of focality or nonfocality may exist on a continuum. As a follow-up to McBride and Abney's work, Abney et al. (2013) carried out two additional studies to look at task-appropriateness and focality with orthographic and semantic manipulations. Overall, their semantic PM task had greater accuracy compared to their orthographic PM task. This finding could play into the levels of processing differences with deeper encoding leading to higher performance (Craik & Lockhart, 1972; Seamon & Virostek, 1978).

Recently, these types of studies have been taking a different methodological approach to understanding the differences between focal and nonfocal conditions. Two studies in particular (McBride & Abney, 2012; Abney et al., 2013) analyzed their work using ex-Gaussian RT parameters, which researchers use to examine the entire RT distribution, rather than just the mean. Abney et al. (2013) found μ differences only when they asked participants to put in high levels of effort. In their second experiment, they found μ differences when comparing blocks of trials with an intention to those without an intention. This finding possibly indicates that having an intention pushed participants to expend higher levels of effort to complete their tasks. McBride and Abney (2012) found differences in τ, with the general trend implying the differences between the PM conditions and the control condition were smallest for the most focal condition (their classification as TAP/focal) and largest for the most nonfocal condition (their classification as TIP/ nonfocal). Similarly, Abney et al. (2013) also found τ differences, with more lengthy RTs in the non-overlapping condition compared to the overlapping processing condition (Abney et al., 2013). τ was smaller in the control group than either of the intention comparison groups (Abney et al., 2013; for similar ex-gaussian findings see Ball, Brewer, Loft, & Bowden, 2015 and Rummel et al., 2016). Collectively, these studies show that the key features to attend to in ex-Gaussian analyses in PM research are the μ parameter—which may be a metric for sustained monitoring—and the τ parameter—which might indicate a more fleeting or transient monitoring.

The key element these studies have in common is that they all inform the outcomes of comparing cue specificity and task-appropriateness to various degrees. However, these studies do so without the intention of uncovering which aspects of focality contribute the most to PM performance. Previously, all of the studies discussed have examined focality through either task-appropriateness or cue specificity, or by confounding the two factors, with aims to manipulate the relationship between the ongoing task and PM task without much consideration to both principles in tandem. The guidelines that are commonly used to define focality stem from McDaniel and Einstein's work. Most recently, McDaniel et al. (2015) describe focal processing as the ongoing task encouraging processing of the PM cue's features with an ideal focal cue being a singular, specific cue that a person directly processes as a result of completing the ongoing task. Some researchers aim to reach this through manipulation of the cue specificity (e.g. Einstein et al., 1995; Marsh et al., 2003; Loft & Remington, 2013) while others manipulate the match in processing to achieve focal and nonfocal conditions (e.g. Marsh et al., 2000; Einstein et al., 2005; Scullin et al., 2010; McBride & Abney, 2012). This ambiguity in the literature motivated the current research.

CHAPTER 2: THE PRESENT RESEARCH

The goal of the current research was to better relate cue specificity, task-appropriateness, and focality. The working framework was that both cue specificity and task-appropriateness combine to create differences in focality. The current study purposefully compared specific and general PM intentions that required either TAP or TIP to complete. This design highlights the impacts these features have on PM performance independently and how these two components of focality interact. As discussed previously, researchers commonly define focal PM cues as specific and processed directly from the requirements of the ongoing task (see McDaniel et al., 2015); therefore, crossing cue specificity and task-appropriateness with each other gave rise to one focal condition and three nonfocal conditions (see Table 5).

Because the possible interaction between cue specificity and task-appropriateness had previously not been compared, the current study is instrumental in teasing apart how these two aspects of focality impact PM performance. Currently, studies have examined the TAP/TIP distinction using either only general cues, or only specific cues, but very few studies have used both (see Hicks, Marsh, & Cook, 2005; Meeks & Marsh, 2010; West & Craik, 2001) and only one of those studies (Hicks, Marsh, & Cook, 2005) mentioned cue specificity as a possible factor of interest. The current studies extend the past literature with a direct comparison, and provides much needed evidence for the guidelines in creating a focal task.

Table 5. Experiment design comparing Cue Specificity and Task-Appropriateness to establish focal and nonfocal conditions.

		Cue Specificity		No-Intention
		Specific	General	
Task-Appropriateness	TAP	A. Focal: Specific/ TAP	B. Nonfocal: General/ TAP	E. Control
	TIP	C. Nonfocal: Specific/ TIP	D. Nonfocal: General TIP	

Additionally, the current research allowed for direct comparisons between various nonfocal conditions. Comparing nonfocal conditions that vary in a predicable way has not often been carried out in the literature (see Abney et al., 2013; Hicks, Marsh & Cook, 2005; McBride & Abney, 2012; Meeks & Marsh, 2010). The current study allowed for direct comparisons across three nonfocal groups with various degrees of similarity with the focal comparison group (see Table 5). The comparisons between nonfocal conditions are pertinent for examining ongoing task costs and how the features of focality impact ongoing task performance. The current research addresses whether the three nonfocal conditions show equal interference effects, or if the three nonfocal conditions vary based on the features of cue specificity and task-appropriateness. This helps to clarify the nature of focality and whether it impacts PM like a dichotomy or behaves more like a continuum ranging from focal to nonfocal.

The present research also examined how participants adjust or establish their task expectations with how they perceive the difficulty of the ongoing task and PM task through their selected monitoring strategy. Additionally, the current study examined ongoing task RTs to assess differences in monitoring costs based on specific/general PM cues and TAP/TIP conditions and offers insight into how participants form task expectations in relation to cue specificity and task-appropriateness. RT data are the standard in PM research and used to measure the cost to the ongoing task (Marsh et al., 2003). RT evaluations at the onset of the ongoing tasks prior to experiencing a PM cue offers insight into participants' expectations about the degree of monitoring needed to successfully execute their PM tasks. Additionally, RTs were analyzed using the ex-Gaussian model to clearly portray the changes found in the RT distributions when completing these PM tasks with various levels of focality.

While the primary interests of this study were to examine focality and its two key factors, the current study also offered evidence for the theories mentioned previously. All of the capacity-sharing theories are interested in monitoring levels via the proxy of ongoing task costs. The current study compared costs across four main groups with various degrees of focality and an additional control group. This could support the idea that a focal task does not demonstrate monitoring or costs and that nonfocal tasks do require monitoring, possibly to differing degrees. Alternatively, this study may support the notion that certain types of nonfocal tasks do not require as much monitoring, and that participants can be flexible in the amount of monitoring they use to complete their intentions.

Predictions

The current research measures the accuracy of the PM task, accuracy of the ongoing task, and RTs during the ongoing task. Predictions about the outcomes of the dependent variables are discussed more extensively below, starting with the predictions for the PM accuracy. In Experiment 1, participants determined if a presented word was living or nonliving for the ongoing task, which was semantic in nature. In Experiment 2, participants determined if a word had more or fewer than 3 enclosed spaces inside the letters (see Figure 1), as used by Meier and Graf (2000) as the ongoing task, which was orthographic in nature. The predictions for both experiments were nearly identical. Any discrepancies in predictions between Experiments 1 and 2 are clearly indicated.

Additionally, Table 6 provides a summary of the predictions supported by each of the discussed PM theories, the central memory principles involved in the current experiments, and several empirical works of relevance. Readers should use this table in tandem with Table 5, which provides a clear picture of all anticipated group differences. Additionally, the predictions in Table 6 that are in bold font are those solidly endorsed by the theory or principle. Any predictions in regular font are inferences based on prior PM findings and logical extensions of each theory.

PM Accuracy Predictions based on Memory Principles

Based on the principle of task-appropriateness, participants in the TAP conditions were expected to perform better on the PM tasks than participants in the TIP conditions. Specific cues were expected to lead to better PM performance than general cues based on cue specificity. Levels of processing also predicted that participants would carry out semantic PM tasks more often than orthographic PM tasks regardless of the relationship between the ongoing task and the PM task, which mirrored the predictions of task-appropriateness in Experiment 1. Experiment 2 teased apart task-appropriateness and levels of processing predictions, discussed more thoroughly later.

Ongoing Task Accuracy

As with ongoing task RTs, ongoing task accuracy was measured on trials that do not contain a PM cue. As noted by Smith et al. (2007) in discussions of PAM theory, studies do not often find changes in ongoing task accuracy (see also Marsh et al., 2003). The multiprocess framework and dynamic multiprocess framework do not make solid predictions about ongoing task accuracy as a cost; they primarily focus on RT differences in the ongoing task as the interference effect. However, any potential differences were expected to resemble those found in some of the empirical neurological PM studies carried out recently (see Cona et al., 2013, McDaniel et al., 2015, and Strickland et al., 2017).

Table 6. Prediction Table based on Theories and Principles.

Predictions for Experiments 1 and 2				
	Ongoing Task RT Costs	Ex-Gaussian RTs	Ongoing Task Accuracy	PM Task Accuracy
<u>Theory</u>				
PAM Theory	**E<A** A<B=C=D B=C<D	μ and/or τ differences A=B=C=D ≠ E	**Differences are not often found.**	
Multiprocess Framework	**E=A** A<B=C=D B=C<D	μ and/or τ differences B=C=D ≠ E ?? A=E	No solid predictions	
Dynamic Multiprocess Framework	**E=A** A<B=C=D B=C<D	μ and/or τ differences B=C=D ≠ E Changes in τ, showing intermittent monitoring	No solid predictions	
Delay Theory	E<A=B=C=D A<B=C=D B=C<D?	?? Possibly changes in μ for living trials in A and B conditions. Equal Changes in μ for living/nonliving trials in C and D conditions	No solid predictions	
<u>Principle</u>				
	Ongoing Task RT Costs	Ex-Gaussian RTs	Ongoing Task Accuracy	PM Task Accuracy
Cue Specificity	**A=C < B=D**		B=D < A=C	**A=C > B=D**
Task-Appropriateness	**A=B < C=D**	Increase in τ A=B=C=D>E	C=D < A=B	**A=B > C=D**
Levels of Processing				Experiment 1: **A=B>C=D** Experiment 2: **C=D>A=B**

Table 6 continued.

Prior Research				
	Ongoing Task RT Costs	Ex-Gaussian RTs	Ongoing Task Accuracy	PM Task Accuracy
Abney et al. (2013)		τ: E=A<B=C=D		
Cona et al. (2013); McDaniel et al. (2015); Strickland et al. (2017)			Experiment 1: B and D have equal accuracy to E for both living and nonliving trials. A and C cues show decreased accuracy compared to E for living trials, and increased accuracy for nonliving trials (reflecting task specific interference) Experiment 2: B and D have equal accuracy to E for both living and nonliving trials. A and C show decreased accuracy compared to E for trials with 3 or more enclosed spaces, and increased accuracy for trials with less than 3 enclosed spaces (task specific interference)	

Notes: Table should be interpreted in tandem with Figure 5 for ease of group demarcations. Strong predictions are marked in bold font. Inference and existing empirical works guide all other predictions. Empty cells indicate no grounds to make any predictions. A = Specific/TAP condition; B= General/TAP condition; C= Specific/TIP condition; D= General/TIP condition; E= Control condition.

Collectively, the findings of Cona et al. (2013; Table 2) and Strickland et al. (2017; Table 2) predict that ongoing trial accuracy in the general, nonfocal conditions would equal that of the control group. Having encoded specific, focal cues would impact the accuracy of the ongoing task differently for the types of ongoing task trials that the cue words appear in. All of the current study's cue words were living (Experiment 1) and had 3 or more enclosed spaces (Experiment 2), so those respective ongoing task trial types were expected to show a decrease in accuracy, while the nonliving (Experiment 1) and fewer than 3 enclosed spaces (Experiment 2) trials were expected to be equal to the accuracy found in the control comparison group. McDaniel et al. (2015) also examined ongoing task accuracy and how task-appropriateness impacts it. They found that performance was high across their focal, nonfocal, and control conditions (>90%) and that all groups were equally accurate. Therefore, no main effect was predicted for task-appropriateness, when examining ongoing task accuracy.

In investigating difference scores between the first baseline block accuracy and the second experimental PM block accuracy, a slight increase in accuracy was anticipated for the control group. The control group would have more practice with the ongoing task and no intention to learn or to distract them, and therefore, should improve. The PM groups were all expected to have slightly worse ongoing task accuracy compared to their baseline blocks. A main effect of cue specificity was anticipated, in which holding general cue intentions would lead to worse ongoing task performance than specific cue intentions. A main effect of task-appropriateness was also anticipated, in which TIP conditions were expected to be worse on the ongoing task than TAP conditions.

To review, based on the general literature, no significant findings were expected for ongoing task accuracy. However, if one were to be found, predictions pointed to an interaction between cue specificity and trial type. The general PM conditions were expected to show performance equal to or slightly worse than the control condition across both the TAP and TIP conditions. The specific cues were expected to show higher accuracy on the nonliving and "less than 3 spaces" trials, and worse accuracy on the living and "3 or more" trials. Additionally, a main effect would be expected showing that overall accuracy was slightly higher for the nonliving and "less than 3" trial types, in line with Strickland et al.'s (2017) findings.

Cost Predictions

The predicted RT costs found on the ongoing tasks vary based on the theory. The PAM theory predicts that all four of the PM conditions would show monitoring costs when compared to the control condition without a PM intention. The PAM theory would also predict that the level of costs could be greater when factors that should theoretically increase monitoring define a given condition. In this way, costs should increase when moving from a specific cue to a general cue, or from a TAP condition to a TIP condition. When the two factors are layered together costs

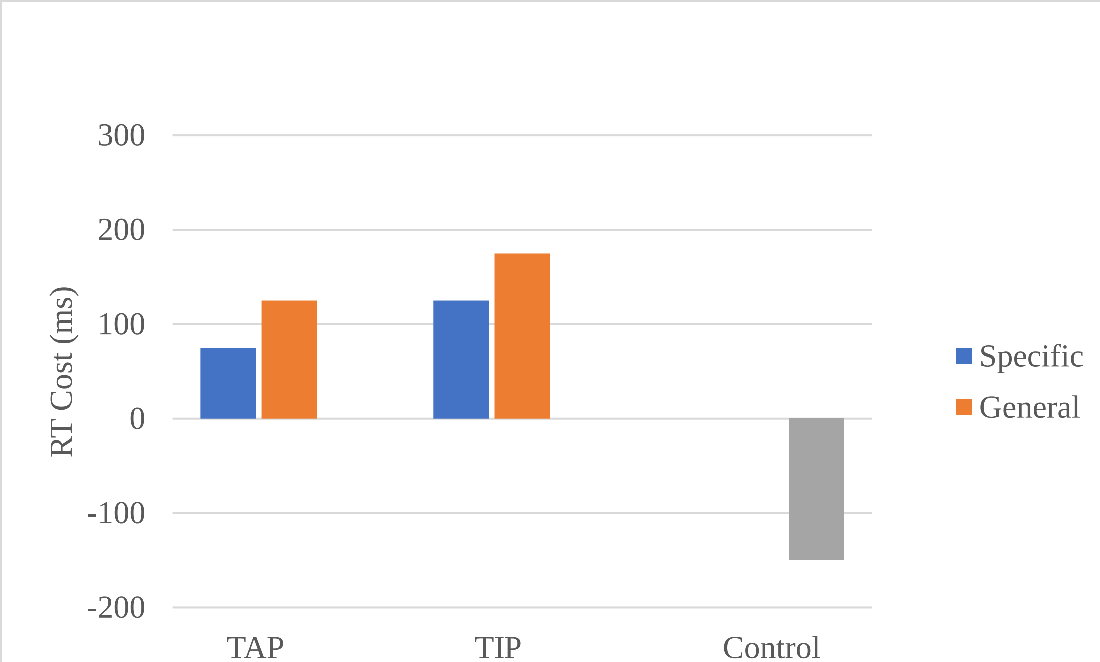

Figure 4. Experiments 1 and 2 Predicted Ongoing Task RT Difference Scores.

may work in an additive manner such that the greatest costs result from the general/TIP condition, moderate costs from the specific/TIP and general/TAP conditions, and the smallest costs from the specific/TAP condition, in a step-like manner (see Figure 4).

The multiprocess framework would predict that the focal condition (specific, TAP cue) would show little or no costs when compared to the control, no intention condition, but the nonfocal conditions should show a relative cost similar to what is depicted in Figure 4. The multiprocess framework would also consider the focal condition to rely on spontaneous retrieval. Meanwhile, the cost predictions from the dynamic multiprocess framework would be identical to the multiprocess framework predictions, aside from the possibility that monitoring costs could change throughout the task as participants gain experience with the PM cues.

Lastly, delay theory (Heathcote et al., 2015) would predict that ongoing task RTs would be comparable across all conditions but that the delay in responding to the ongoing task would increase for the conditions with a PM intention. More specifically, delay theory could predict a difference in RTs between the living and nonliving trials, as the PM task is confined to only living trials during the TAP conditions, as found by Strickland et al. (2017). The TIP conditions should encourage a cost on all trials because the participants cannot form an expectation about when the PM cue might appear.

All in all, based on previous research (see Table 2) and the theories outlined above, a main effect of cue specificity was expected, with faster ongoing task RTs in the specific conditions and slower RTs in the general conditions. A main effect of task-appropriateness was also anticipated, favoring the TAP conditions. The predictions outlined also admit the possibility of finding a difference based on trial type (living/nonliving; more than 3 spaces/less than 3 spaces) that could interact with task-appropriateness.

RT Predictions based on Ex-Gaussian Distributions

Previous research using ex-Gaussian parameters consider changes in μ to reflect relatively consistent monitoring for the PM task. Changes in τ may reflect transient monitoring in a sporadic manner across the ongoing task. Therefore, the PAM theory would predict a difference in μ and/or τ for each of the four PM groups compared to the control group. This prediction stems from PAM theory's core principle that preparatory attention must be in use at all times with an intention, whether conscious or not, which reduces available resources and leads to costs. For PAM theory (and the other capacity derived theories) it is not certain whether the changes in RT costs would be driven by μ or τ, as long as the changes were large enough to increase the overall mean RT. Similarly, the multiprocess framework would predict changes in μ/τ for the nonfocal conditions compared to the control condition, but no change in μ for the focal condition; τ might show changes in the focal condition. This prediction is based on Brewer's (2011) interpretation of μ representing active monitoring and τ representing a cost on a few trials following spontaneous processing, which should be utilized in focal conditions. The dynamic multiprocess framework may predict more changes in τ, which has been considered a metric for transient, or flexible and fleeting, monitoring, as the participants gain experience with the PM and ongoing tasks. The delay theory might predict a change in μ for living trials in the TAP conditions, but equal changes in μ for both living and nonliving trials for the TIP conditions, all compared to the control condition, which may show a similar or slightly decreased μ as a result of practice effects.

The control condition was not expected to show any differences in τ between the baseline and the experimental block (though possible to find small reductions in μ and/or σ due to practice effects); an increase in τ was anticipated for both TAP and TIP conditions, as was the case for Abney et al. (2013). Further, the tail distribution was expected to be especially large in the nonfocal conditions, with possibly no differences, or small ones, between the focal and control condition.

CHAPTER 3: EXPERIMENT 1 METHOD

The current experiment manipulated the cue specificity (specific or general) and the task-appropriateness (TAP or TIP) of a PM cue during a semantic ongoing task in order to obtain the full picture of focality and how it relates to PM accuracy and ongoing task costs. Institutional Review Board approval was obtained through Louisiana State University prior to data collection (see Appendix C).

Participants

Participants recruited for this study were Louisiana State University undergraduate students taking psychology courses. Participants were given course credit or extra credit through the Sona Systems experiment portal for their participation in this study. All participants provided informed consent before participating in the study. Estimates for the sample size needed to carry out this study came from Abney et al. (2013) who used an identical living/nonliving ongoing task and had a semantic PM task compared to an orthographic PM task. The power analyses were completed using G*Power software (Faul, Erdfelder, Lang, & Buchner, 2007) and the estimate was based on effect sizes reported by Abney et al. (2013). Abney et al. (2013) found a significant main effect of task-appropriateness, with a medium effect size of $\eta_p^2 = .08$. G*Power was used to converted this value to $f = 0.295$ for use in the sample size estimates. Using their effect size and setting the other parameters to appropriate levels ($\alpha = 0.05$, power = 0.95, the numerator df =1, and the number of groups to 4), the a priori power calculation for an ANOVA: Fixed effects, special, main effects and interactions led to a sample size computation equal to 152 total participants. This was equal to 38 people per group; and with the inclusion of the control group, the grand total of participants collected for Experiment 1 should equal 190 people for adequate power. For conservative estimates, an N=200 was set as the goal to allow for possible errors in collection with 40 people per group. Additional power analyses were conducted based on Abney et al.'s (2013) other results, including their RT ex-Gaussian analyses. However, each subsequent power analysis resulted in smaller sample size estimates, so the most conservative estimate for the proposed study was retained.

Two hundred six participants were tested in each experiment, for a grand total of 412 participants. Of these participants, two in each experiment (4 total) experienced computer glitches that prevented the full recording of their data and were excluded from all analyses.

Materials

The current experiment was programmed and conducted using E-Prime 2.0 software (Psychology Software Tools, 2012), which controlled stimulus presentation and recorded participant responses and RTs. Words used for the ongoing task were generated by the MRC Psycholinguistic Database (Wilson, 1988) and constrained to include only nouns with 3-9 letters. Words were excluded if they were an animal or a word with double letters (e.g. bull) to avoid creating additional PM cues in certain PM conditions, resulting in 348 total words (these can be found in Appendix A). These words were used for both Experiment 1 and 2, so balancing the words based on their classifications for the ongoing task was needed. The lists were balanced by animacy and the number of enclosed spaces so that exactly half of the words used were living, and the other half were nonliving (relevant for Experiment 1); additionally, half of each living and nonliving category had 3 or more enclosed spaces, and half had less than 3 enclosed spaces (relevant for the ongoing task used in Experiment 2). Each of the word lists used in the ongoing task (living words/3 or more spaces; living words/ less than 3 spaces; nonliving words/3 or more spaces; nonliving words/ less than 3 spaces) were constructed to be comparable based on measures of concreteness (*M* = 522.20 *SD* = 6.63), familiarity (*M* = 490.61, *SD* = 10.63), imaginability (*M* = 523.32, *SD* = 6.28), the number of letters (*M* = 6.20, *SD* = 0.38), and the number of syllables (*M* = 2.07, *SD* = 0.15) calculated from the MRC Psycholinguistic Database (Wilson, 1988). All words (n = 432) were normed to get a frequency of agreement in classifying the words as living or non-living. Approximately 63 individuals took part in the norming process. Words had to meet a classification ratio of .65, meaning at least 65% of the participants agree in classifying the concept as living or nonliving. The final word lists ended with an average agreement on classification above 90% for the two living word lists, and above 97% agreement for the two nonliving lists, with each list trimmed to n = 87 words.

Participants saw two PM cues selected from six possible PM cues counterbalanced across participants. The PM cues were the same for all groups, differing only by the instructions participants read, and each cue was presented twice. Use of the same cues across groups was necessary to avoid any differences in cue properties that could have impacted PM performance (Marsh et al., 2000). The PM cues used for all groups were: baboon, goose, moose, raccoon, rooster, and kangaroo. Ex-Gaussian analyses were conducted using the Quantile Maximum Probability Estimator software (QMPE; Heathcote, Brown, & Cousineau, 2004).

Procedure

Participants were tested individually. Each participant was randomly assigned to a condition and study between Experiments 1 and 2. Simultaneous collection of both experiments protected against any possible differences introduced by participants that complete experiments early or late in the semester and allows for cleaner comparisons across experiments. After providing their informed consent, participants read the instructions about the task.

The ongoing task required participants to make a living/ nonliving judgment about words shown to them on a screen by pressing either the “f” or “j” keys on a keyboard as quickly and accurately as possible. Participants first completed a short block of practice trials (20 trials total) of the ongoing task and were allowed to ask questions if needed. After the practice block of trials, participants carried out a baseline block of only the ongoing task (172 total trials). After completing the baseline block, participants read the directions for completing the PM task (excluding the control condition). The PM instructions provided were manipulated between participants. The TAP/Specific condition required participants to press the “y” key when they saw two of the following specified animal words: baboon, moose, rooster, goose, raccoon, or kangaroo (participants were told exactly which two cue words to watch for, counterbalanced across all participants). The TAP/General condition gave participants instructions to press “y” when they saw any animal word, excluding humans. The TIP conditions required participants to either press “y” when they saw two consecutive double letter o’s (TIP/Specific condition), or to press “y” when they saw any two consecutive letters, or doubles (TIP/General condition). After the participants indicated that they understood both the PM instructions and the ongoing task instructions they filled out demographic information (age and gender) that served as a distraction delay task prior to starting the PM trials. Participants were told they were allowed to ask any questions at this time before starting the task. Once participants began the ongoing task (180 trials total; 4 of which were PM trials), the 2 PM cues appeared embedded evenly spaced across the second half of the ongoing task trials (roughly every 28 trials) to allow for a longer delay between encoding the intention and the first PM cue. Once participants finished with the task, they were asked six questions about possible monitoring, as a manipulation check to insure they understood the PM instructions. The questions used can be found in Table 7. Participants were then debriefed and given credit through Sona Systems for their participation.

Table 7. Post-Experiment Questions.

Post-Experiment Questions to Assess PM Recall and Recognition		
#	Question:	Use:
1	Do you remember any additional tasks you were supposed to complete during the experiment? If so please describe the task below in as much detail as you can.	Measure Recall of PM task.
2	What did you write on the notecard you handed to the researcher? If you didn't do this part of the experiment, type "nothing".	Recall details of the PM task cue and intention specifically.
3	Which key were you supposed to press for your notecard instructions? If you didn't do this part of the experiment, type "no".	Recall the PM intention.
4	Which of these keys were you supposed to press for your notecard task? Was it T? Y? G? H? B? or N?	Recognition of the PM intention key.
5	Which of these instructions were you given earlier in the experiment? Press the number that matches your instructions. 1. Press Y when I see either of the following words: goose or moose[a] 2. Press Y when I see any animal words. 3. Press Y when I see a word with double letter o's. 4. Press Y when I see any double consecutive letters in a word. 5. I did not receive any of these directions during the experiment.	Recognition of the PM task.
6	Was there anything you did during the course of today's experiment to help you to remember to press the Y key when you saw a word that fit your notecard instructions?	Exploration of self-reported strategy use for PM task.
Notes: [a]Specific words presented matched cues presented to each participant.		

Design

This experiment manipulated two independent variables (cue specificity and task-appropriateness) with two levels each (specific and general, TAP and TIP, respectively), all between-subjects, and compared these groups to a control condition with no PM intention. Table 5 provides a schematic of the conditions included in this experiment. Additional variables of interest built into the design as within-subjects manipulations included the trial type (living and nonliving) and block (baseline block and PM block), which became import when examining ongoing task accuracy, and ongoing task RTs.

CHAPTER 4: EXPERIMENT 1 ANTICIPATED RESULTS

PM Accuracy

Participants' PM accuracy was scored as a percentage of the four possible PM cues they detected. Presenting two PM cues twice each allowed for a proportion score for PM detection. Use of two PM cues and four PM cue presentations avoided too frequent of cue presentation that could have turned the task into a vigilance task (see Graf, Uttl, & Dixon, 2002). Moreover, this number of cue presentations should still have reduced monitoring as much as possible due to the small number of cues participants in the focal condition were asked to remember (see Cohen et al., 2008; Rummel et al., 2016 Experiments 2 and 3 who used two PM cues in their focal condition; Einstein et al., 2005). Participants that did not recall the PM instructions in the post experiment questionnaire were excluded from the analysis, as their failures possibly indicated retrospective memory failure rather than PM failure. The accuracy scores were compared in a 2 (cue specificity) x 2 (task-appropriateness) ANOVA. Anticipated findings would have indicated a main effect of cue specificity in which specific cues led to higher PM performance than general cues. A main effect of task-appropriateness was also anticipated, with higher PM performance in the TAP conditions than the TIP conditions. No interaction was expected, as cue specificity and task-appropriateness have roughly the same effect sizes and should have impacted performance in a balanced manner (eliminating an interaction). However, if one was found it was expected to show focal performance as the highest, and all of the three nonfocal conditions showing worse PM performance.

Ongoing Task Accuracy

Ongoing task accuracy is affected by PM instructions inconsistently across prior research studies. The different focality groups were expected to produce very little statistical differences in ongoing task accuracy, with the noted possibility that the following differences could be obtained assuming that intentions did produce a cost on ongoing task performance. Importantly, ongoing task accuracy does not include PM cue trial accuracy. Ongoing task accuracy was examined using a 2 (cue specificity: specific/general) x 2 (task-appropriateness: TAP/TIP) x 2 (trial type: living/nonliving) x 2 (block: baseline/PM block) mixed model Analysis of Variance (ANOVA), here block and trial type were within-subjects variables and cue specificity and task-appropriateness were between-subjects variables. The inspection of trial type as a separate factor is due to the predictions of the Delay theory (Heathcote et al., 2015), which predicts task specific interference on trials that match with the PM cue (living trials in this case). The results are displayed in Figures 5 and 6.

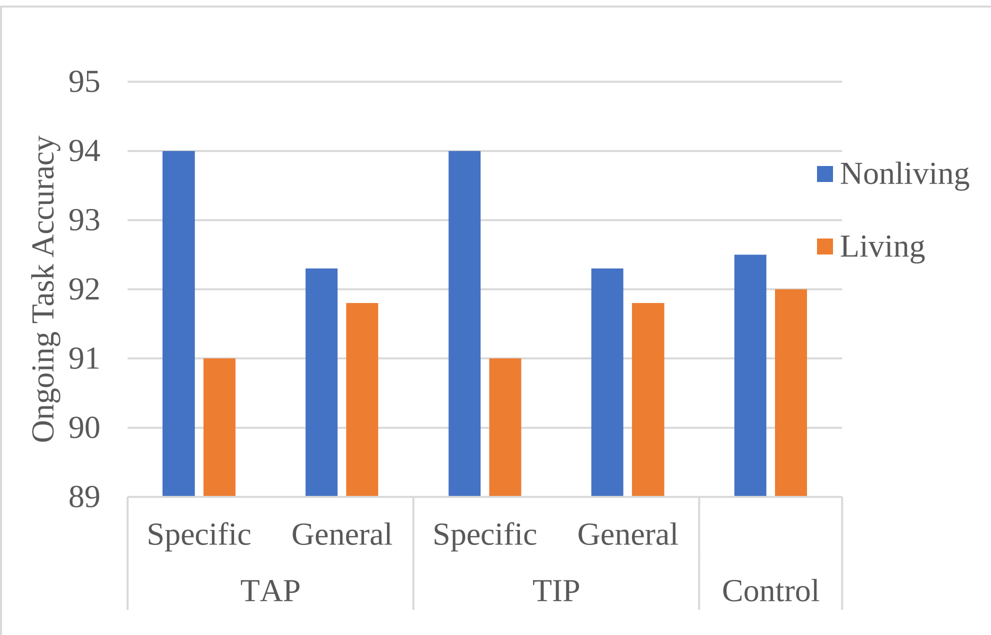

Figure 5. Predicted Experiment 1 Ongoing Task Accuracy in the PM Block.
Note: Classified by Living/ Nonliving trials to allow for Delay Theory Predictions (Heathcote et al., 2015) with task specific interference, as found by Strickland et al. (2017).

A difference score was calculated by taking the average ongoing task accuracy percentage from the first baseline block and subtracting it from the ongoing task accuracy percentage from the PM (second) block. A main effect of task-appropriateness was expected, indicating a larger decrease in ongoing task accuracy for the TIP conditions compared to the TAP conditions (see Figure 6). A main effect of cue specificity was anticipated, with worse

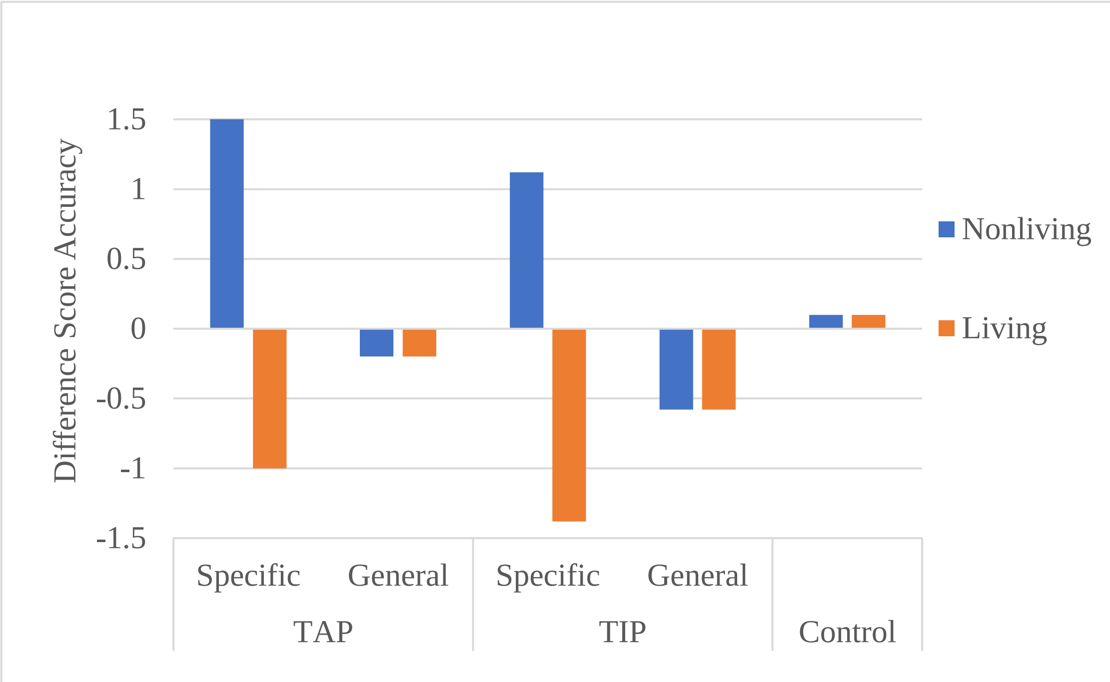

Figure 6. Predicted Experiment 1 Ongoing Task Accuracy Difference Scores.
Note. Block 2-Block 1. Classified by Living/ Nonliving trials to allow for Delay Theory Predictions (Heathcote et al., 2015) with task specific interference, as found by Strickland et al. (2017).

accuracy in the general, nonfocal conditions (see Figures 5 and 6). A main effect of trial type was expected, with living trials predicted to lead to worse accuracy than nonliving trials (see Figures 5 and 6). This prediction was based on the Delay Theory expectations of task specific interference disrupting the trials that were similar to the PM cues (see Strickland et al., 2017).

A significant interaction between block and cue specificity was expected, where the PM block impacts the general condition more than the specific condition. Similarly, an interaction between block and task-appropriateness was anticipated, where the PM block impacted the TIP conditions more than the TAP conditions. An interaction between block and trial type was not expected (see Figure 6). An interaction between trial type and cue specificity was expected, in that nonliving trials should have shown higher accuracy, especially for the specific condition. The living trials should have shown worse performance, especially for the specific condition. These predictions were based on the principles of cue specificity, task-appropriateness, as well as cue specific interference which was prevalent in the Delay Theory.

No interaction was expected between task-appropriateness and cue specificity for ongoing task accuracy. And finally, three-way or four-way interactions were not expected.

Planned comparisons between the control condition and each of the conditions in the 2 (cue specificity) x 2 (task-appropriateness) analysis were also carried out. Differences were expected in the difference scores demonstrating the overall impact of holding an intention, with each PM condition expected to show worse performance than the control, with the least difference found in the focal (specific/TAP) condition. When examining living/nonliving performance, an interaction with cue specificity was expected showing no differences between the control group and the general conditions, but the specific conditions were expected to have worse living trial accuracy compared to the control group, yet better accuracy compared to the control group for the nonliving trials (based on Strickland et al., 2017; see Figure 5).

Ongoing task RTs

RTs for the ongoing task were prepared in a similar manner to Brewer (2011). RTs during the baseline block and the PM block were measured separately. Any trials with incorrect decisions from the RT analysis were excluded. Additionally, RTs on the PM trials and the three trials following the PM cue were excluded as per Brewer’s (2011) recommendations. Each participant’s mean RT for each block of trials were calculated, and any RT score outside of 2.5 standard deviations were excluded from analyses. Each participant had two RT means calculated through this process: their baseline RTs and their PM block RTs. These RTs were also subjected to ex-Gaussian modeling as described later.

A difference score between the two RT means was calculated for each participant and subjected to a 2 (cue specificity) x 2 (task-appropriateness) ANOVA (see Figure 4). A main effect of cue specificity was expected, with more slowing for the general conditions compared to the specific conditions. A main effect of task-appropriateness was also expected, with more slowing on the TIP conditions compared to the TAP conditions. No interaction was expected for ongoing task RTs. However, this prediction came with the supposition that if an interaction was found, it would likely have shown the fastest RT in the focal condition (Specific/TAP), with the other three nonfocal conditions (Specific/TIP; General/TAP; General/TIP) all equally slow. This finding aligns with focality acting like a dichotomy between focal and nonfocal conditions.

Planned comparisons between the control condition and each of the four conditions of the focality manipulations were conducted. All of the PM groups were expected to show slowing in the ongoing task compared to the control condition. The RT slowing was expected to be least extreme for the focal (specific/ TAP) condition, and most extreme for the nonfocal (general/ TIP) condition. The other two nonfocal conditions (General/ TAP and Specific/ TIP) were expected to show moderate amounts of slowing compared to the control condition. These predictions would have supported the PAM theory and the delay theory, with each intention condition still showing some slowing compared to the control group. If the findings had shown no cost differences between the focal condition and the control condition, this would have supported the multiprocess framework and the dynamic multiprocess framework.

Ex-Gaussian Analyses

Similar to the RT preparations suggested by Brewer (2011), each participant's RT scores were entered into the QMPE software (Heathcote et al., 2004) separately in order to calculate estimates for μ, τ, and σ[1]. Fixed quantiles were established (0, .2, .4, .6, .8, 1) to estimate each participants' parameters. RTs were separated for living trials and nonliving trials.

Each model parameter was subjected to a 2 (block: baseline/ PM block) x 2 (trial type: living/ nonliving) x 5 (Group: 4 PM groups and control group) mixed methods ANOVA to see how the parameters changed with the inclusion of an intention.

The first parameter of interest, μ, is the RT distribution mean. If μ is indeed influenced by active monitoring (e.g., a target checking strategy or relatively consistent monitoring; Guynn, 2003), a main effect of group would be expected with the nonfocal PM conditions showing higher μ values, with lower μ values for the focal and control condition. An interaction between group and block was expected, as the control group (and possibly the focal group) was expected to have similar μ values across both blocks, and all of the PM nonfocal groups were expected to increase in the PM block. A three-way interaction between trial type, block, and group was also expected. If this interaction was present, it was predicted to show similar μ values for living and nonliving trials for the two TIP conditions that increased during the PM block, similar μ values for the living and nonliving trials for the control condition that may have decreased slightly in the PM block due to practice effects, and an increased μ value for only the living trials in the two TAP groups when in the PM block. This interaction would support the delay theory.

The second parameter of interest was τ, which affects the tail of the full RT distribution. An interaction between group and block was anticipated. The control condition was not expected

[1] Special thanks to Hunter Ball for resources and guidance in using the QMPE software.

to show any differences in τ between blocks, but τ was expected to increase for all PM conditions. It was predicted that the tail distribution would be especially large in the nonfocal conditions, with possibly no or small differences between the focal and control condition. The differences among the three nonfocal conditions was of particular interest, as this comparison had not been carried out in prior work.

No differences in the third parameter, σ, or the RT standard deviation, were anticipated among any of the groups.

CHAPTER 5: EXPERIMENT 1 ACTUAL FINDINGS

Data Cleaning

In both experiments, trials with RT data less than 200 ms were removed, resulting in a loss of .09% of the data from Block 1 and .22% of the data in Block 2 for an overall loss across both experiments of 0.15% of the data. For each participant's correct responses, extremely long RT responses were examined. Any RT exceeding 3 standard deviations above a participant's own condition mean was excluded from analyses, resulting in a loss of 1.80% of the total RT data across both experiments. Six participants were excluded from all analyses (2 from Experiment 1, 4 from Experiment 2) for high frequencies (>10) of false alarm responses to non-PM cues indicating they misunderstood the task. RT responses were then compared at a group level to identify participants that exceeded 3 standard deviations above their respective group means. This resulted in excluding 4 more participants in Experiment 1 and excluding 1 participant from Experiment 2. Finally, ongoing task performance was analyzed to insure all participants had at least 90 correct trials for each block of trials, roughly 55% accuracy in each block. Participants that had poor ongoing task performance in both blocks were excluded (loss of 8 participants), as well as those that performed poorly in only Block 1 (loss of 5 participants), or only Block 2 (loss of 5 participants).

After data cleaning, Experiment 1 retained 196 participants, and Experiment 2 retained 183 participants. Demographic information for each group of participants can be found in Table 8. No significant differences were found between groups when comparing ages, F (9,336) = 1.39, p = .19. A Chi-square test indicated no significant association between group and gender across both experiments, $\chi2$ (9) = 7.56, p = .58, Cramer's V = .14.

Table 8. Summary Demographic Information for Experiments 1 and 2.

Demographic Information for Experiments 1 and 2					
	Group	Initial N	Final N	Average Age (std. dev)	Gender: Percent Female (%)
Experiment 1					
TAP/ Specific	1	41	39	19.43 (1.32)	68.42
TAP/General	2	42	40	18.61 (1.09)	66.67
TIP/Specific	3	41	40	19.11 (1.24)	79.49
TIP/General	4	41	38	18.83 (1.04)	78.95
Control Group	5	41	39	19.18 (1.72)	69.23
Experiment 2					
TAP/ Specific	6	41	38	19.06 (1.14)	78.38
TAP/General	7	42	35	18.82 (1.24)	82.35
TIP/Specific	8	41	34	18.86 (1.11)	78.79
TIP/General	9	41	38	18.82 (0.67)	63.89
Control Group	10	41	38	19.40 (2.24)	68.42

PM Accuracy

In order to get the full picture of PM Accuracy, two PM proportions were calculated. First was the proportion of PM cues (out of 4) that participants correctly identified. Second was a proportion calculated to include correct responses on the PM trial or on either of the 2 trials immediately following the PM trial. These two proportions created a strict PM response window versus a lax response window criterion. The post-experiment questions (Question 2 and Question 5) were used to classify whether participants were able to freely recall the PM task independently (Question 2) as well as recognize the PM intention correctly out of a list of plausible PM instructions (Question 5). These two questions were used to exclude participants that failed to recall and/or recognize the PM task after completing the experiment. Excluding the control group, 31 participants failed to free recall their PM task (successful participant n = 126) and 13 participants failed to recognize their PM instructions correctly (successful participant n = 144).

I first carried out a 2 (appropriateness) x 2 (specificity) ANOVA on PM proportion correct using a strict PM response window, and only included participants that correctly recalled the PM intention in the post-experiment questions. There was a significant main effect of task-appropriateness: $F(1,122) = 14.36$, $p < .001$, $\eta_p^2 = .11$, with PM accuracy significantly higher in the TAP condition ($M = .67$, $SE = .04$) compared to the TIP condition ($M = .44$; $SE = .04$). I also found a significant main effect of cue specificity: $F(1,122) = 25.69$, $p < .001$, $\eta_p^2 = .17$. PM accuracy was significantly higher in the specific condition ($M = .71$, $SE = .04$) compared to the general condition ($M = .40$; $SE = .04$). There was no interaction [$F(1,122) = .001$, $p = .98$, $\eta_p^2 < .001$]. See Figure 7 for a visualization of these findings.

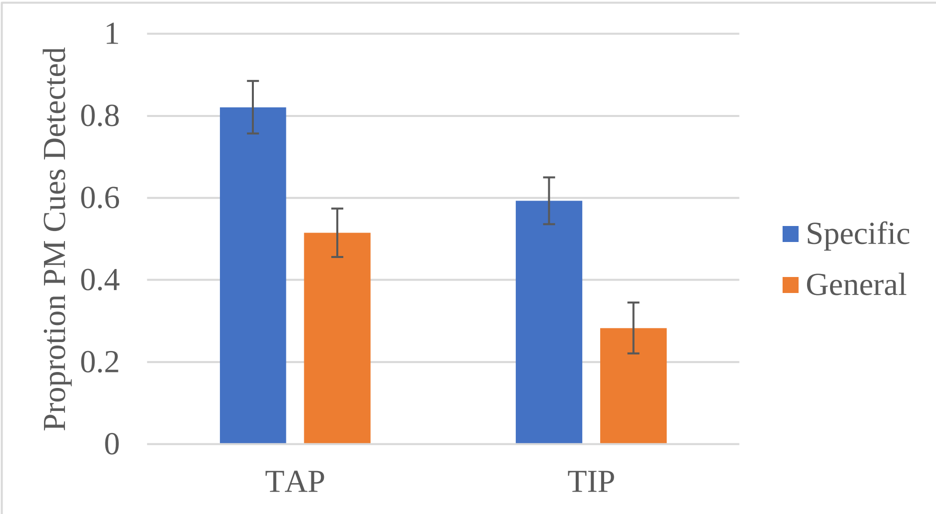

Figure 7. Experiment 1 PM Accuracy Under Strict PM Window and Recall Only Criterion.

This analysis was repeated using the lax PM response window, then both analyses repeated again but including all participants that were able to successfully recognize their PM instructions. The pattern observed for the strict response window and strict recall analyses remained in all other combinations of strict and lax response windows and recall vs. recognition inclusion. For a full report of these analyses, see Table 9. The two main effects remained, and no interaction was present in any analyses for PM accuracy.

Table 9. Corroborative Analyses to Examine PM Accuracy.

Experiment 1						
	df	*F* value	*p* value	η_p^2	Power	Result
Lax PM Window with Strict Recall						
Appropriate	(1,122)	9.91	**.002**	.08	.88	TAP > TIP
Specificity	(1,122)	20.37	**< .001**	.14	.99	Specific > General
Interaction	(1,122)	.048	.828	< .001	.06	None
Strict PM Window with Lax Recognition						
Appropriate	(1,140)	10.06	**.002**	.07	.88	TAP > TIP
Specificity	(1,140)	25.91	**< .001**	.16	1.00	Specific > General
Interaction	(1,140)	2.96	.088	.02	.40	None
Lax PM Window with Lax Recognition						
Appropriate	(1,140)	6.74	**.010**	.05	.73	TAP > TIP
Specificity	(1,140)	20.38	**< .001**	.13	.99	Specific > General
Interaction	(1,140)	1.99	.161	.01	.29	None

Ongoing Task Accuracy

Accuracy for the ongoing task was calculated as a proportion score of the number of trials the participant correctly identified out of their total number of valid trials in each block. The number of valid trials for each participant varied slightly as the first two trials of each block were excluded from analyses as well as the two trials immediately following each of the PM cues.

In order to examine differences in ongoing task performance, I carried out a 2 (cue specificity) x 2 (appropriateness) 2 x (trial type: living and nonliving) x 2 (block: baseline and PM block) mixed model ANOVA on the proportion of correct decisions participants made for the ongoing task to see if their accuracy changed at all with the addition of a PM intention. A main effect of trial type was present, $F(1,133) = 114.67$, $p < .001$, $\eta_p^2 = .46$, indicating that participants had better accuracy on the nonliving trials ($M = .93$, $SE = .005$) compared to the living trials ($M = .86$, $SE = .01$). Additionally, I found a main effect of Block, $F(1,133) = 6.78$, p

= .01, η_p^2 = .05. This main effect shows slight but significant differences in accuracy with Block 1 having more accurate ongoing task performance (*M* = .90, *SE* = .01) compared to the performance in Block 2 (*M* = .89, *SE* = .01). No significant 2-way, 3-way, or 4-way interactions were present (all *p*'s > .10).

For closer examination of possible differences between the two blocks, a difference score was calculated by taking the accuracy proportion score for Block 1 away from Block 2. A 2(cue specificity) x 2(appropriateness) x 2(trial type) mixed model ANOVA was then carried out on the ongoing task accuracy difference scores. There was no longer a significant main effect of trial type, $F(1,133) = .95$, $p = .33$, $\eta_p^2 = .01$. Additionally, there were no main effects for either cue specificity or task-appropriateness, and no interactions among any of the variables.

Planned comparisons were carried out to examine each experimental condition against the control group that had no PM intention for the duration of the experiment (see Table 10).

Table 10. Ongoing Task Accuracy Group Comparison Summary.

	Mean Difference	*SE*	Dunnet t-value	*p* value	Cohen's d (small = .02, medium = .05, large = .08)
Block 1 Accuracy					
TAP/Specific vs. Control	0.0006	0.0129	0.047	1.00	0.01
TAP/General vs. Control	-0.0082	0.0128	-0.640	0.92	-0.15
TIP/Specific vs. Control	0.0016	0.0126	0.124	1.00	0.03
TIP/General vs. Control	-0.0138	0.0133	-1.038	0.68	-0.23
Block 2 Accuracy					
TAP/Specific vs. Control	0.0136	0.0193	0.706	0.89	0.17
TAP/General vs. Control	-0.0141	0.0192	-0.735	0.88	-0.17
TIP/Specific vs. Control	0.0069	0.0192	0.359	0.99	0.09
TIP/General vs. Control	-0.0188	0.0195	-0.966	0.73	-0.21
Difference Score Accuracy					
TAP/Specific vs. Control	0.0068	0.0093	0.726	0.88	0.16
TAP/General vs. Control	-0.0106	0.0092	-1.153	0.61	-0.25
TIP/Specific vs. Control	-0.0020	0.0092	-0.213	1.00	-0.05
TIP/General vs. Control	-0.0088	0.0096	-0.918	0.77	-0.25

There were no significant differences between any of the experimental groups and the control group.

Ongoing Task RTs

Analyzing Baseline Performance

The ongoing task RT analyses were carried out using both the RT means and the RT medians for comparison purposes. This section reports the findings for the RT means, and any differences found in the median analyses are noted to highlight discrepant findings.

Analyzing Block 1 RT means on their own allows for a look into baseline differences between groups and provides a clearer picture of costs that can be attributed to the addition of an intention. I carried out a 2 (specificity) x 2 (appropriate) x 2 (trial type) mixed model ANOVA on the mean RTs for Block 1. There was a main effect of trial type, $F(1,153) = 151.80$, $p < .001$, $\eta_p^2 = .50$, with faster responses for the living trials ($M = 780.09$, $SE = 9.38$) compared to the nonliving trials ($M = 875.83$, $SE = 13.93$). There were no other main effects or interactions, signifying overall steady baseline performance across all experimental groups.

Planned comparisons were carried out to compare the experimental groups to the control condition for both the living and nonliving trial types. No significant differences were found in comparison to the control group (see Table 11).

Analyzing the PM Block

The same analyses were repeated for Block 2 alone to see changes resulting from the added PM intention. A 2 (specificity) x 2 (appropriate) x 2 (trial type) mixed model ANOVA was carried out on the RT means of Block 2. The main effect of trial type was again significant, $F(1,153) = 140.37$, $p < .001$, $\eta_p^2 = .48$, showing faster RTs for the living trials ($M = 853.98$, SE =

12.18) compared to the nonliving trials (*M* = 960.33, *SE* = 16.13). A main effect of appropriateness was also found, $F(1,153) = 8.97$, $p = .003$, $\eta_p^2 = .06$, showing faster RTs for the TAP conditions (*M* = 866.52, *SE* = 19.12) compared to the TIP conditions (*M* = 947.78, *SE* = 19.25).

Table 11. Planned Comparisons for Block 1 RTs.

	Mean Difference	*SE*	Dunnet t value	*p* value	Cohen's d (small = .2, medium = .5, large = .8)
Nonliving Mean RTs					
TAP/Specific vs. Control	36.23	40.39	0.90	0.78	0.19
TAP/General vs Control	13.70	40.13	0.34	0.99	0.07
TIP/Specific vs. Control	37.46	40.13	0.93	0.76	0.22
TIP/General vs Control	21.01	40.65	0.52	0.96	0.12
Living Mean RTs					
TAP/Specific vs. Control	18.24	27.21	0.67	0.91	0.13
TAP/General vs Control	-7.45	27.04	-0.28	1.00	-0.06
TIP/Specific vs. Control	8.16	27.04	0.30	0.99	0.07
TIP/General vs Control	2.48	27.39	0.09	1.00	0.02

These main effects were qualified by a significant interaction between trial type and appropriateness, $F(1,153) = 5.87$, $p = .02$, $\eta_p^2 = .04$, and an unexpected 3-way interaction between trial type, specificity, and appropriateness, $F(1,153) = 4.43$, $p = .04$, $\eta_p^2 = .03$. Figures 8 and 9 display each of these interactions. The same analyses carried out on median scores

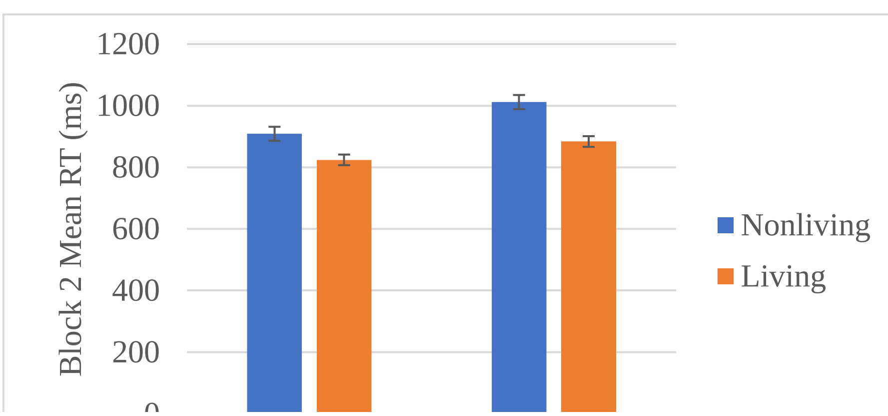

Figure 8. Block 2 Interaction between Trial Type and Appropriateness.

duplicated the main effects for trial type and appropriateness, however both of the interactions were no longer significant (trial type and appropriateness: $F(1,153) = 3.69$, $p = .057$, $\eta_p^2 = .02$; trial type, specificity, and appropriateness: $F(1,153) = 2.85$, $p = .09$, $\eta_p^2 = .02$).

Pairwise comparisons for the interaction between appropriateness and trial type show significant differences between nonliving and living trials for both TAP [$F(1,153) = 44.71$, $p <$

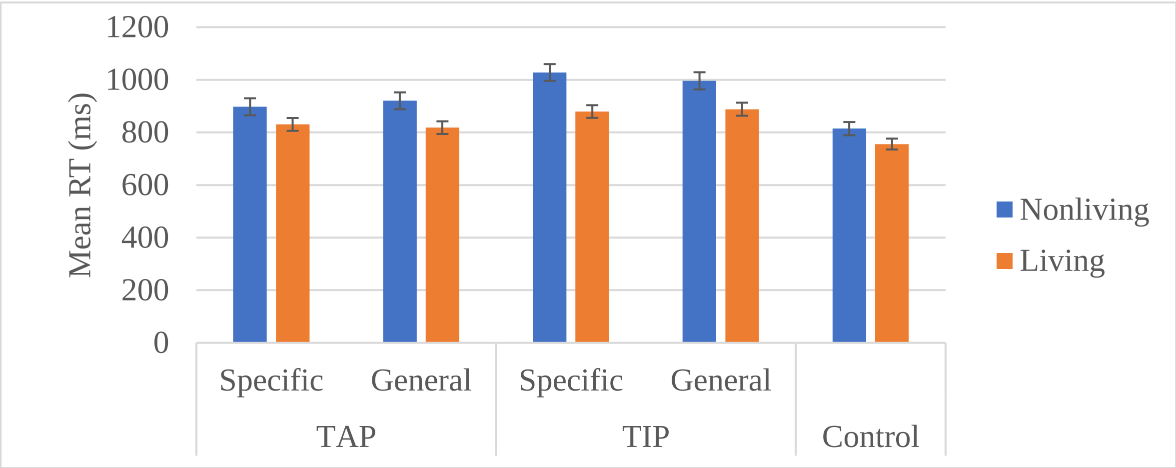

Figure 9. Block 2 Mean RT Interaction between Trial Type, Specificity, Appropriateness. Note. The control group was not analyzed as part of the 3-way interaction but is shown for comparison.

$.001$, $\eta_p^2 = .23$] and TIP conditions [$F(1,153) = 101.15$, $p < .001$, $\eta_p^2 = .40$], with the interaction seemingly driven by the difference in magnitude of these comparisons.

Pairwise comparisons were also conducted to investigate the factors driving the 3-way interaction. Comparisons between each trial type for all four experimental groups were significant (all $p < .001$), but with varying magnitudes in their effect sizes with the TAP/Specific group having the smallest difference between living and nonliving trials ($\eta_p^2 = .08$), the TAP/General and TIP/General groups having moderate effect sizes ($\eta_p^2 = .18$ and $\eta_p^2 = .19$, respectively) and the largest difference between trial types present for the TIP/Specific group ($\eta_p^2 = .31$).

All experimental groups were compared to the control group in planned comparisons. Both the nonliving trials [$F(4,191) = 7.42$, $p < .001$] and the living trials [$F(4,191) = 5.05$, $p = .001$] showed significant differences between the control group and TIP experimental groups, but not the TAP groups. The details of the follow up analyses are presented in Table 12.

Table 12. Planned Comparisons with the Control Group for Nonliving and Living Mean RTs.

	Mean Difference	*SE*	Dunnet t value	*p* value	Cohen's d (small = .2, medium = .5, large = .8)
Nonliving Mean RTs					
TAP/Specific vs. Control	83.19	43.90	1.90	0.18	0.48
TAP/General vs Control	105.96	43.62	2.43	0.054	0.59
TIP/Specific vs. Control	213.44	43.62	4.89	**<.001**	1.21
TIP/General vs Control	181.71	44.19	4.11	**<.001**	0.96
Living Mean RTs					
TAP/Specific vs. Control	74.71	33.58	2.22	0.09	0.55
TAP/General vs Control	62.45	33.37	1.87	0.19	0.47
TIP/Specific vs. Control	123.68	33.37	3.71	**0.001**	0.92
TIP/General vs Control	132.51	33.80	3.92	**<.001**	0.83

Difference Score Analyses

A difference score was calculated by subtracting the mean RTs of Block 1 from Block 2; the difference score was then analyzed in a 2 (specificity) x 2 (appropriateness) x 2 (trial type) mixed model ANOVA. The main effect of trial type that was significant in each block measured alone was washed out using the difference score, showing that the overall pattern of RT speed on the two trial types was consistent across both blocks, $F(1,153) = 1.93$, $p = .17$, $\eta_p^2 = .01$. The rest of the significant findings map onto the PM Block analyses with a significant main effect of appropriateness ($F(1,153) = 22.32$, $p < .001$, $\eta_p^2 = .37$), a significant interaction between trial type and appropriateness [$F(1,153) = 6.57$, $p = .01$, $\eta_p^2 = .04$], and a significant 3-way interaction between trial type, appropriateness, and specificity ($F(1,153) = 4.08$, $p = .045$, $\eta_p^2 = .03$)[2]. These interactions are visible in Figure 10.

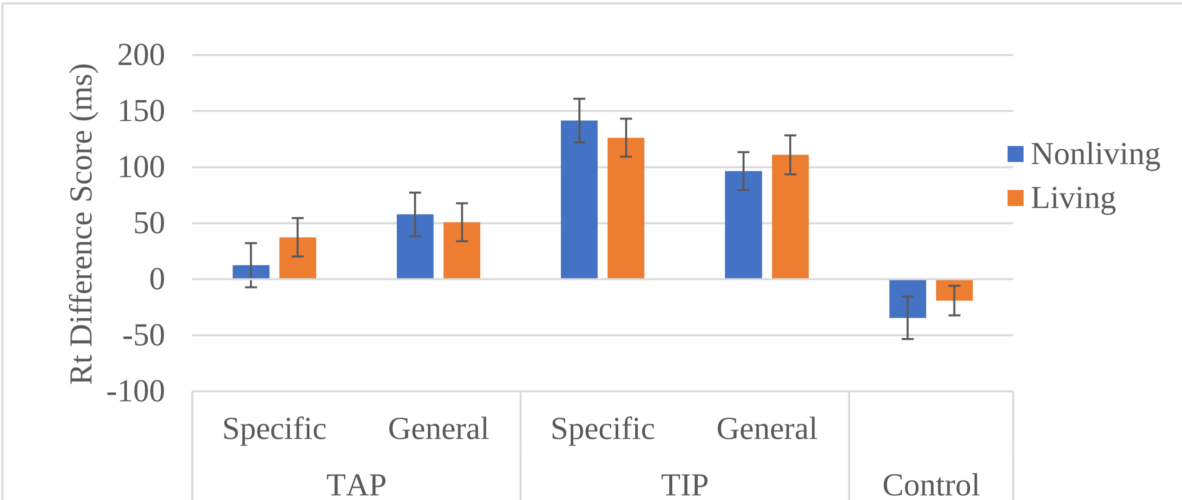

Figure 10. Mean RT Difference Scores by Trial Type for all groups.

Pairwise comparisons on the 3-way interaction show a similar pattern at the Block 2 analyses, but with the difference scores, the only comparison that reached significance was between the two trial types for the TIP/Specific group [$F(1,153) = 8.88$, $p = .003$, $\eta_p^2 = .06$], including the nonsignificant difference between living and nonliving trials for the control group.

The planned comparison follow-up test revealed significant differences between the experimental groups and the control group for both living ($F(4,191) = 9.83$, $p < .001$) and nonliving trials ($F(4,191) = 14.58$, $p < .001$; see Table 13).

[2] A 2 (specificity) x 2 (appropriateness) x 2 (trial type) Mixed model ANCOVA was carried out examining mean RTs in the PM Block while adjusting for Block 1 performance across both trial types. A significant main effect of trial type was present, $F(1,152) = 11.63$, $p < .001$, $\eta_p^2 = .07$, with faster performance on the living trials ($M = 853.97$, SE = 8.42) compared to the nonliving trials ($M = 960.31$, $SE = 10.14$). All other findings from the difference score analyses were replicated in the ANCOVA.

Table 13. Planned Comparisons with the Control Group for Mean RT Difference Scores.

	Mean Difference	*SE*	Dunnet t value	*p* value	Cohen's d (small = .2, medium = .5, large = .8)
Nonliving					
TAP/Specific vs. Control	46.96	27.64	1.70	0.26	0.48
TAP/General vs. Control	92.26	27.47	3.36	**0.004**	0.74
TIP/Specific vs. Control	175.99	27.47	6.41	**<.001**	1.41
TIP/General vs. Control	160.70	27.82	5.78	**<.001**	1.24
Living					
TAP/Specific vs. Control	56.48	23.25	2.43	0.054	0.75
TAP/General vs. Control	69.90	23.10	3.03	**0.01**	0.73
TIP/Specific vs. Control	115.52	23.10	5.00	**<.001**	1.18
TIP/General vs. Control	130.04	23.40	5.56	**<.001**	1.23

These findings show the focal group (Tap/Specific) is the only group that is not different from the control group for both the nonliving and living trial analyses. When examining median scores instead, the pattern remains the same for the nonliving trials, with only the TAP/Specific group showing no difference from the control group. However, in the living trials, all experimental groups are significantly different from the control group, showing evidence of cost for all intention groups in the living trials.

Ex-Gaussian Analyses

The RT data was also analyzed using ex-Gaussian parameters to examine the different parameters of each participant’s RT distribution and look for changes that would reflect different strategy use in different conditions. This section of the results looks first at μ—a possible proxy for monitoring based strategies—followed by τ—thought to represent transient or spontaneous retrieval—and concludes with σ—not associated with any PM strategies and not expected to show any differences among groups.

Analyses for μ

I carried out a 2 (specificity) x 2 (appropriateness) x 2 (trial type) x 2 (block) mixed model ANOVA on the μ parameter. A main effect of trial type was present, $F(1,153) = 116.80$, $p < .001$, $\eta_p^2 = .43$, with living trials ($M = 555.83$, $SE = 4.43$) showing a smaller μ than nonliving trials ($M = 605.75$, $SE = 6.61$). This finding parallels the traditional RT findings with an average faster RT score for the living trials compared to the nonliving trials. There was also a significant interaction between block and appropriateness: $F(1,153) = 8.52$, $p = .004$, $\eta_p^2 = .05$. Pairwise comparisons show no difference between the two blocks for TAP conditions [$F(1,153) = .60$, $p = .44$, $\eta_p^2 = .004$] but a significant difference between them for TIP conditions [$F(1,153) = 11.18$, $p = .001$, $\eta_p^2 = .06$]. The pattern of findings is displayed in Figure 11. No other main effects or interactions reached a significant level (all p's > .07).

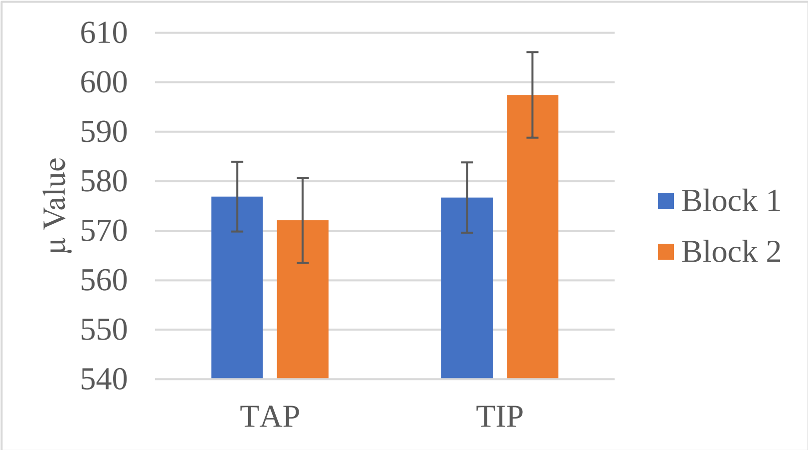

Figure 11. μ interaction between block and appropriateness.

Comparisons to the control group were carried out for both blocks for each trial type. There were no significant differences for either trial type in Block 1 for any of the experimental groups when compared to the control group. However, both trial types showed significant differences in Block 2 when compared to the control group. The details of the Block 2 comparisons can be found in Table 14.

Table 14. Block 2 planned comparisons with control group for μ values.

	Mean Difference	*SE*	Dunnet t value	*p* value	Cohen's d (small = .2, medium = .5, large = .8)
Block 2 Nonliving μ					
TAP/Specific vs. Control	55.32	21.32	2.60	**0.04**	0.68
TAP/General vs. Control	17.52	21.18	0.83	0.82	0.24
TIP/Specific vs. Control	70.08	21.18	3.31	**0.004**	0.89
TIP/General vs. Control	60.23	21.45	2.81	**0.02**	0.65
Block 2 Living μ					
TAP/Specific vs. Control	20.00	14.62	1.37	0.45	0.33
TAP/General vs. Control	13.56	14.53	0.93	0.76	0.26
TIP/Specific vs. Control	38.24	14.53	2.63	**0.03**	0.62
TIP/General vs. Control	39.23	14.72	2.67	**0.03**	0.60

Baseline Block Comparisons for μ

To further ensure the groups were equivalent during the baseline block, a 2 (specificity) x 2 (appropriateness) x 2 (trial type) mixed model ANOVA was carried out on just Block 1 μ values. A main effect of trial type was present, F 1,153) = 109.62, $p < .001$, $\eta_p^2 = .42$, showing the same pattern of smaller μ values for the living trials (M = 553.07, SE = 4.46) compared to the nonliving trials (M = 600.53, SE = 3.36). Additionally, there was an unexpected main effect of specificity, $F(1,153) = 4.96$, $p = .03$, $\eta_p^2 = .03$, indicating larger μ values in the specific conditions (M = 587.94, SE = 7.05) compared to the general conditions (M = 565.66, SE = 7.10). No other main effects or interactions were found.

Analyzing the PM Block for μ

Due to the differences found in the baseline block, a difference score was calculated by subtracting the μ values found in Block 1 from the values found for Block 2. A 2 (specificity) x 2 (appropriateness) x 2 (trial type) mixed model ANOVA was carried on to examine the μ difference scores. A main effect of appropriateness was found, $F(1,153) = 8.52$, $p = .004$, $\eta_p^2 = .05$, signifying a difference between TAP conditions ($M = -4.78$, $SE = 6.16$) and TIP conditions ($M = 20.75$, $SE = 6.21$). This difference shows a trend for a slight decrease in μ with the addition of an intention for the TAP conditions, but an increase in μ for the TIP conditions after receiving their PM intention. No other main effects or interactions were found[3].

In follow-up tests comparing the experimental conditions to the control condition, significant differences were present for both living and nonliving trials. The pattern of findings (shown in Table 15) shows no differences between the TAP conditions and the control group for the nonliving trials, and no differences between the TAP/Specific condition and the control condition for the living trials. Significant differences were found for the TIP conditions in nonliving trials, and for all three nonfocal conditions for the living trials.

Table 15. μ difference score planned comparisons with the control group.

	Mean Difference	*SE*	Dunnet *t* value	*p* value	Cohen's d (small = .2, medium = .5, large = .8)
Nonliving Difference Score μ					
TAP/Specific vs. Control	26.88	15.92	1.69	0.27	0.52
TAP/General vs. Control	34.21	15.82	2.16	0.10	0.50
TIP/Specific vs. Control	63.27	15.82	4.00	**<.001**	0.95
TIP/General vs. Control	63.66	16.02	3.97	**<.001**	0.93
Living Difference Score μ					
TAP/Specific vs. Control	17.21	11.67	1.47	0.39	0.40
TAP/General vs. Control	34.13	11.59	2.94	**0.013**	0.80
TIP/Specific vs. Control	38.11	11.59	3.29	**0.004**	0.75
TIP/General vs. Control	49.48	11.74	4.21	**<.001**	0.94

[3] The analyses for Block 2 were carried out using Block 1 performance on living trials as a covariate. The pattern of findings was replicated with only a main effect of appropriateness present. When carried out using the nonliving trials as a covariate, the main effect of appropriateness was still present, and a main effect of trial type reappeared.

Group Analyses for μ

A 2 (block) x 2 (trial type) x 5 (group) mixed methods ANOVA was carried out for μ to get a complete sense of group differences. There was a main effect of trial type [$F(1,191) = 143.95$, $p < .001$, $\eta_p^2 = .43$], replicating the pattern of a smaller μ for the living trials ($M = 553.75$, $SE = 3.97$) compared to the nonliving trials ($M = 600.30$, $SE = 5.69$).

An interaction was also found between block and group, $F(1,191) = 7.45$, $p < .001$, $\eta_p^2 = .14$. Figure 12 displays the interaction. Pairwise comparisons show that all five groups were

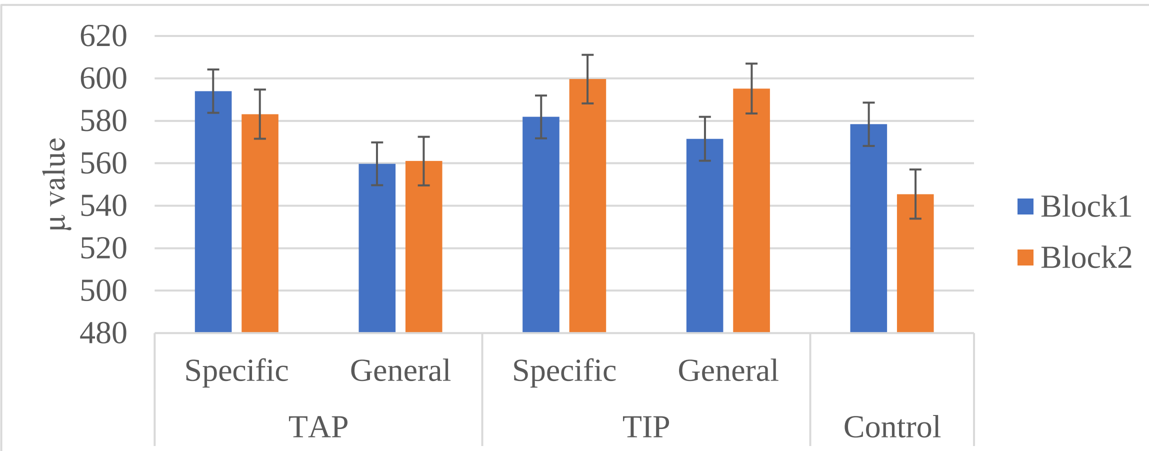

Figure 12. μ group differences across blocks.

comparable during Block 1 [$F(4,191) = 1.56$), $p = .19$, $\eta_p^2 = .03$] but differed in Block 2 [$F(4,191) = 3.98$), $p = .004$, $\eta_p^2 = .08$]. The TAP/Specific, TIP/Specific, and TIP/General groups were all significantly different from the control group for Block 2. The TAP/General condition was not significantly different from the control group but was significantly different from both TIP conditions.

Analyses for τ

The analyses for τ began with a 2 (specificity) x 2 (appropriateness) x 2 (trial type) x 2 (block) mixed model ANOVA on the τ parameter to see overall patterns. A main effect of trial type was present, $F(1,153) = 53.16$, $p < .001$, $\eta_p^2 = .26$, with living trials ($M = 262.98$, $SE = 7.43$) showing a smaller τ than nonliving trials ($M = 315.19$, $SE = 11.28$). Additionally, a main effect of Block was present, $F(1,153) = 111.75$, $p < .001$, $\eta_p^2 = .42$, showing much larger τ values in Block 2 ($M = 324.90$, $SE = 10.41$) compared to Block 1 ($M = 253.28$, $SE = 8.45$). Similar to the μ findings, there was a significant interaction between block and appropriateness: $F(1,153) = 14.97$, $p < .001$, $\eta_p^2 = .09$. The pattern of findings is displayed in Figure 13. Pairwise

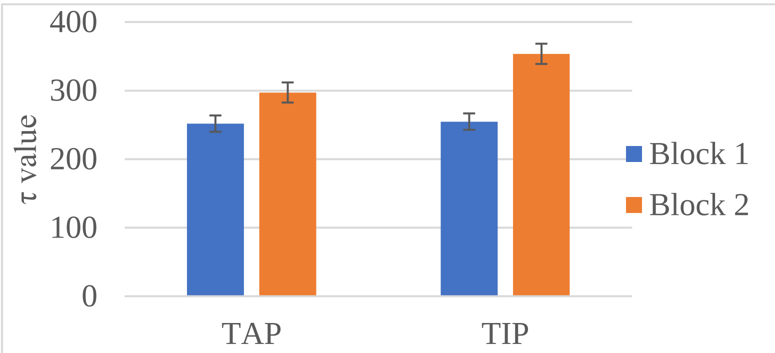

Figure 13. τ interaction between block and appropriateness.

comparisons of this interaction show a significant difference between the two blocks for both the TAP conditions [$F(1,153) = 22.61$, $p < .001$, $\eta_p^2 = .13$] as well as the TIP conditions [$F(1,153) = 103.58$, $p < .001$, $\eta_p^2 = .40$], with the magnitude of the effects driving the interaction.

There was also a significant 3-way interaction found between trial type, specificity, and appropriateness, $F(1,153) = 5.43$, $p = .02$, $\eta_p^2 = .03$. This interaction is presented in Figure 14. Pairwise comparisons show no difference between trial types for the TAP/Specific group ($p = .224$), but significant differences for all other experimental groups (all p's $<.001$) and the control group ($p = .022$).

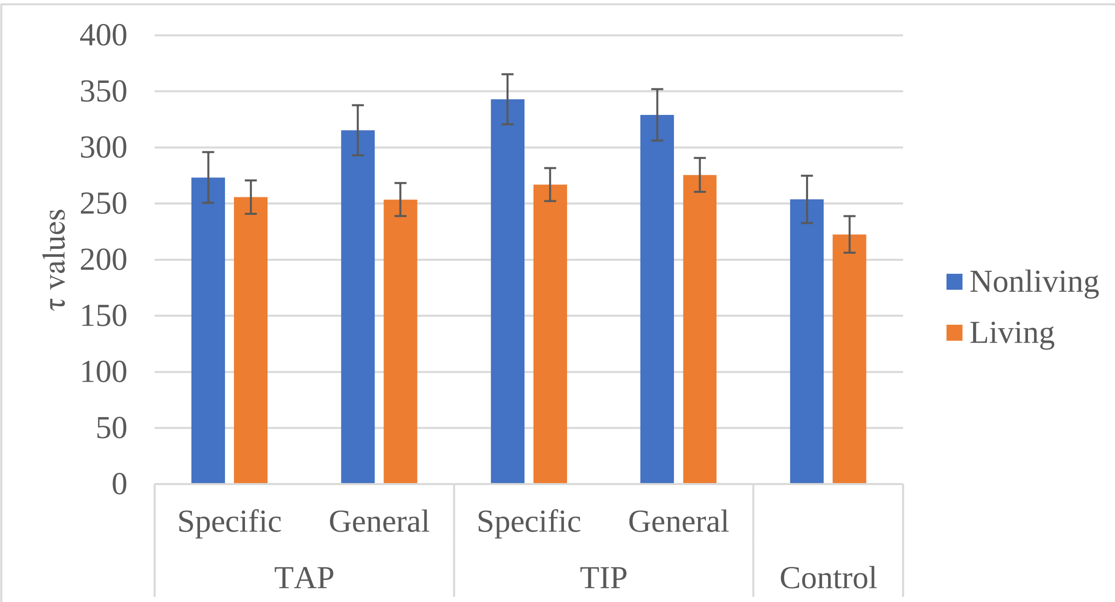

Figure 14. τ 3-way interaction between trial type, specificity, and appropriateness.
Note. The control group is provided just for comparison.

In follow-up tests to compare each experimental group to the control group, no significant findings were present for either trial type in Block 1, showing comparable baseline performance across all groups. Block 2 performance did show differences between the experimental groups and the control group. The three nonfocal groups were significantly different from the control group in the nonliving trials, and the TIP conditions were significantly different from the control group for the living trials. These outcomes are presented in Table 16.

Table 16. Block 2 τ values compared to the control group.

	Mean Difference	*SE*	Dunnet *t* value	*p* value	Cohen's d (small = .2, medium = .5, large = .8)
Block 2 Nonliving τ					
TAP/Specific vs. Control	28.88	36.10	0.80	0.84	0.24
TAP/General vs. Control	90.33	35.88	2.52	**0.04**	0.56
TIP/Specific vs. Control	144.71	35.88	4.03	**<.001**	1.03
TIP/General vs. Control	123.69	36.34	3.40	**0.003**	0.81
Block 2 Living τ					
TAP/Specific vs. Control	53.52	25.64	2.09	0.12	0.52
TAP/General vs. Control	50.92	25.48	2.00	0.15	0.45
TIP/Specific vs. Control	82.05	25.48	3.22	**0.006**	0.84
TIP/General vs. Control	94.62	25.81	3.67	**0.001**	0.77

Baseline Block Comparisons for τ

A 2 (specificity) x 2 (appropriateness) x 2 (trial type) mixed model ANOVA was carried out on just Block 1 for the τ parameter to insure baseline performance across groups was comparable. A main effect of trial type was present, $F(1,153) = 48.70$, $p < .001$, $\eta_p^2 = .24$, with smaller τ values for the living trials (*M* = 228.63, *SE* = 7.42) compared to the nonliving trials (*M* = 277.93, *SE* = 10.61). No other main effects or interactions were present.

Difference Scores for τ

A difference score was calculated to account for the performance in Block 1 when looking at Block 2 performance (Block 2 – Block 1). A 2 (specificity) x 2 (appropriateness) x 2 (trial type) mixed model ANOVA was carried out on τ difference scores. A main effect of appropriateness was found, $F(1,153) = 14.97$, $p < .001$, $\eta_p^2 = .09$. The difference between blocks was larger in the TIP conditions (*M* = 97.83, *SE* = 9.61) compared to the TAP conditions (*M* = 45.40, *SE* = 9.55). As both difference scores were positive, this suggests a definite trend for higher τ values in Block 2 compared to Block 1.

Planned comparisons between the experimental groups and the control group for the difference scores were carried out for each trial type. The overall ANOVA tests for each trial type were significant indicating significant differences between the experimental groups and the control. For the nonliving trials, the TIP conditions differed from the control group (the TAP/General nonfocal condition was very close to significantly different from the control at p = .051). The TIP conditions, but not the TAP conditions, were significantly different from the control group for the living trials. These comparisons are displayed in Table 17 below.

Table 17. Comparisons to the control group for τ difference scores.

	Mean Difference	*SE*	Dunnet *t* value	*p* value	Cohen's d (small = .2, medium = .5, large = .8)
Less than 3 Difference Score τ					
TAP/Specific vs. Control	18.74	23.56	0.80	0.84	0.22
TAP/General vs. Control	57.47	23.42	2.45	0.051	0.54
TIP/Specific vs. Control	111.03	23.42	4.74	**< .001**	1.07
TIP/General vs. Control	96.71	23.72	4.08	**< .001**	0.94
3 or more Difference Score τ					
TAP/Specific vs. Control	40.56	20.94	1.94	0.17	0.52
TAP/General vs. Control	39.69	20.81	1.91	0.18	0.43
TIP/Specific vs. Control	75.28	20.81	3.62	**0.001**	0.96
TIP/General vs. Control	83.16	21.08	3.95	**< .001**	0.90

Using Block 1 Performance as a Covariate for Block 2 τ Analyses

The pattern of findings using Block 1 τ values as a covariate for Block 2 performance deviated from the difference score analyses. Using Block 1 living trial τ values as a covariate, I carried out a 2 (specificity) x 2 (appropriateness) x 2 (trial type) mixed model ANCOVA for Block 2 τ values. The significant main effect of appropriateness, $F(1,152) = 12.878$, $p < .001$, $\eta_p^2 = .078$, was still present showing larger τ values in the TIP conditions (M = 353.873, SE = 11.465) compared to the TAP conditions (M = 295.875, SE = 11.390). This main effect was qualified with a significant interaction between trial type and appropriateness [$F(1,152) = 4.434$,

$p = .037$, $\eta_p^2 = .028$], presented in Figure 15. Pairwise comparisons for the interaction between trial type and appropriateness revealed significant differences between TAP and TIP for both trial types, but at a larger magnitude for the nonliving trials [$F(1,152) = 12.715$, $p < .001$, $\eta_p^2 = .077$] compared to the living trials [$F(1,152) = 6.587$, $p = .011$, $\eta_p^2 = .042$].

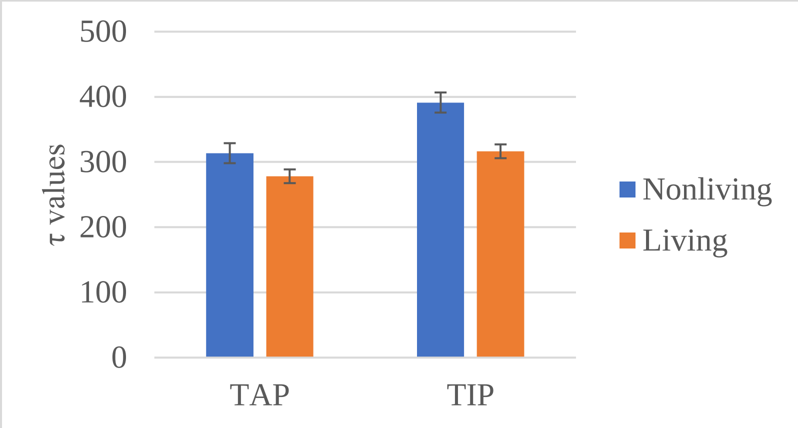

Figure 15. Block 2 τ interaction between trial type and appropriateness.
Note. This analysis used Block 1 performance as a covariate.

Additionally, a 3-way interaction between trial type, appropriateness, and specificity [$F(1,152) = .09$, $p = .01$, $\eta_p^2 = .05$] was present. This interaction is displayed in Figure 16. In the 3-way pairwise comparisons, only the TAP, nonliving comparison between specific and general

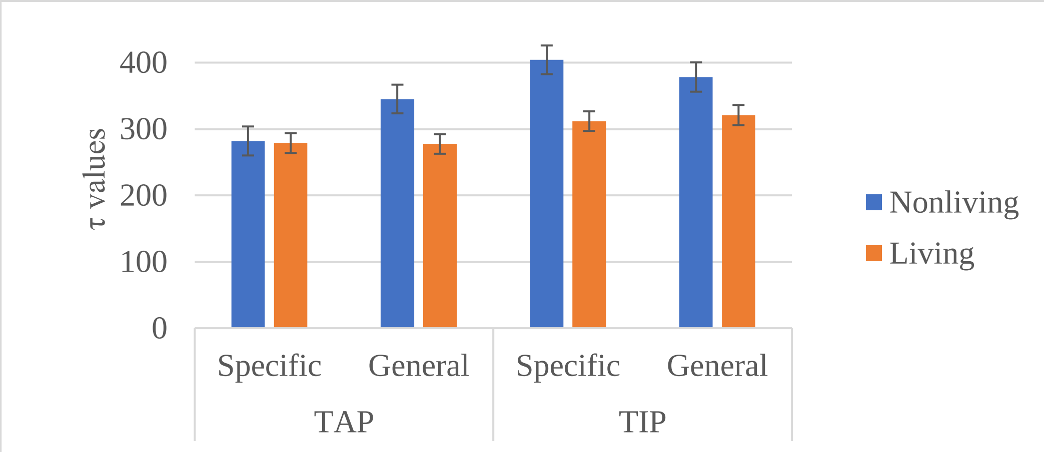

Figure 16. Block 2 τ interaction between trial type, specificity, and appropriateness.
Note. This analysis used Block 1 performance as a covariate.

led to a significant difference [$F(1,152) = 4.24$, $p = .04$, $\eta_p^2 = .03$], but all other comparisons were non-significant (p's > .40).

This pattern of findings was replicated when using the nonliving trial τ values in Block 1 as the covariate with the addition of a main effect of trial type, mirroring other main effect of trial type findings with smaller τ values for the living trials ($M = 297.39$, $SE = 7.26$) compared to the nonliving trials ($M = 352.54$, $SE = 8.52$).

Group Analyses for τ

To further analyze group differences and comparisons to the control group, I carried out a 2 (block) x 2 (trial type) x 5 (group) mixed model ANOVA for the τ parameter. A main effect of Block was present, $F(1,191) = 97.25$, $p < .001$, $\eta_p^2 = .34$, with larger τ values in Block 2 ($M = 308.18$, $SE = 9.03$) compared to Block 1 ($M = 249.63$, $SE = 7.71$). There was also a main effect of group present, $F(4,191) = 2.53$, $p = .04$, $\eta_p^2 = .05$. The means and standard errors for each group are presented in Table 18.

Table 18. Group means for τ.

	Block 1 Mean (*SE*)	Block 2 Mean (*SE*)
TAP/Specific	246.57 (17.27)	282.50 (20.25)
TAP/General	257.07 (17.06)	311.93 (19.99)
TIP/Specific	255.25 (17.06)	354.69 (19.99)
TIP/General	254.24 (17.50)	350.46 (20.51)
Control	235.02 (17.27)	241.31 (20.25)

Additionally, there was a main effect of trial type, $F(1,191) = 63.57$, $p < .001$, $\eta_p^2 = .25$, with smaller τ values in the living trials compared to the nonliving trials. These main effects were qualified by an interaction between trial type and group [$F(1,191) = 3.072$, $p = .018$, $\eta_p^2 = .060$] and an interaction between block and group [F $(1,191) = 8.970$, $p < .001$, $\eta_p^2 = .158$]. These interactions are displayed in Figures 17 and 18 below.

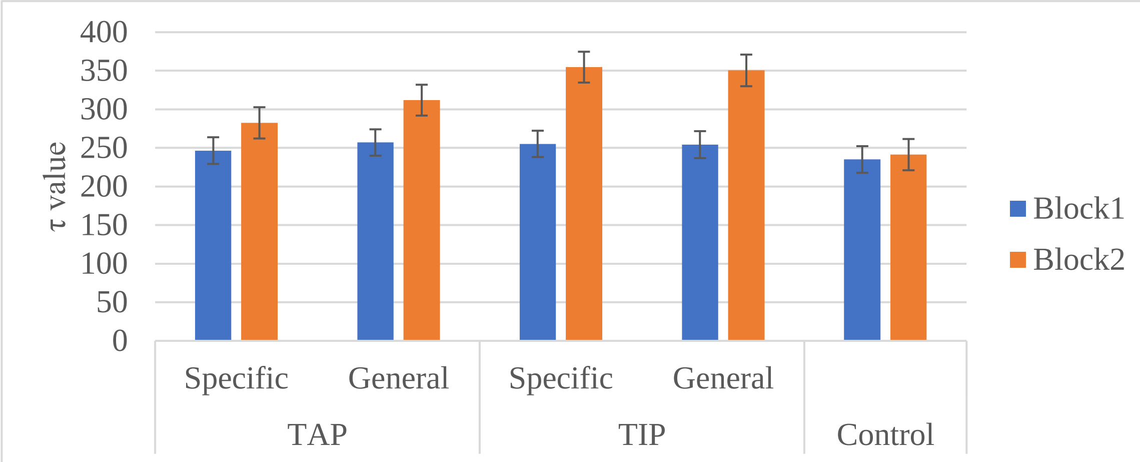

Figure 17. τ Interaction between block and group.

The simple main effects analyses for the trial type and group interaction show no overall differences for the living trials across groups [$F(4,191) = 1.81$, $p = .13$, $\eta_p^2 = .04$], but significant differences for the nonliving trials $F(4,191) = 2.96$, $p = .02$, $\eta_p^2 = .06$] with significant differences between the three nonfocal groups and the control comparison group, as well as a significant difference between the focal TAP/Specific condition and the TIP/Specific condition.

In comparing the two trial types for each group, all nonfocal groups and the control group

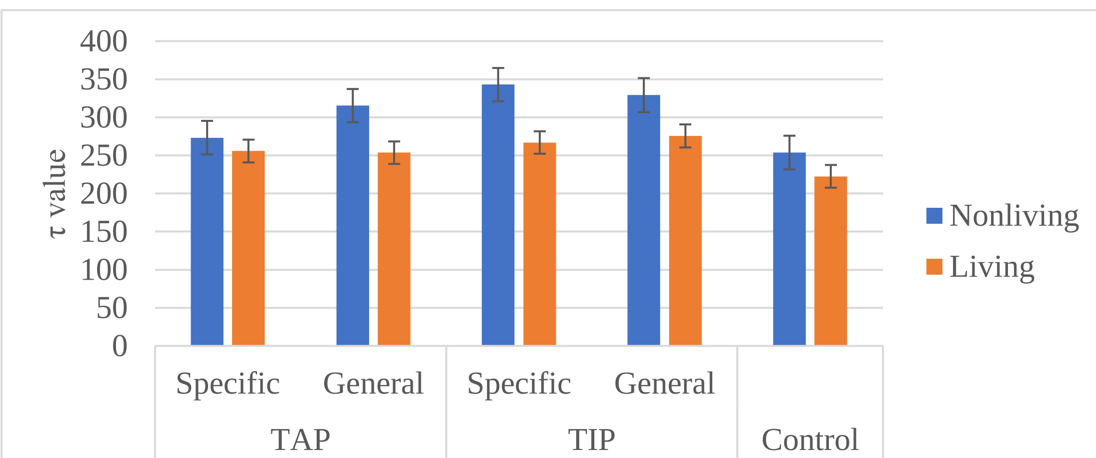

Figure 18. Interaction between trial type and group for τ.

showed significant differences between trial types (all p's < .02), but the TAP/Specific group did not show a significant difference between nonliving and living trials for τ.

The simple main effects analyses for the group x block interaction show no overall differences across groups for Block 1 [$F(4,191) = .28$, $p = .89$, $\eta_p^2 = .01$] but significant differences between groups for Block 2 [$F(4,191) = 5.55$, $p < .001$, $\eta_p^2 = .10$], with significant differences between the control group and each of the three nonfocal groups, and significant differences between the TAP/Specific group and both of the TIP groups.

Analyses for σ

As no major differences were anticipated for the σ parameter, I ran a 2 (specificity) x 2 (appropriateness) x 2 (trial type) x 2 (block) mixed model ANOVA on the σ parameter. There was a main effect of trial type, consistent with the other parameters [$F(1,153) = 35.10$, $p < .001$, $\eta_p^2 = .19$] which found larger σ values for the nonliving trials (*M* = 63.87, *SE* = 2.37) compared to the living trials (*M* = 48.87, *SE* = 1.76). Aside from the differences in trial type, no other main effects or interactions reached significance (all *p*'s > .10).

Group Analyses for σ

To further insure full examination of the σ parameter, a 2 (trial type) x 2 (block) x 5 (group) was carried out on the σ parameter. The main effect of trial type persisted [$F(1,191) = 35.46$, $p < .001$, $\eta_p^2 = .16$], but no other main effects or interactions were found, including no main effect of group ($p = .25$) signifying no overall impact on σ when given a PM instruction.

Correlations

Correlations between PM Accuracy, Ongoing Task RTs, μ values, and τ values were conducted in order to look at the functionality of costs and changes in the ongoing task on PM performance, and to examine whether any relationships existed between PM accuracy and RT costs. Strength of the relationship is based on the guidelines from Evans (1996).

TAP/Specific Group

PM accuracy was not significantly correlated with any RT measures for the TAP/Specific group. The traditional RT measures were very strongly correlated with τ difference scores. μ and τ show a trend for a negative relationship with one another. See Table 19.

Table 19. TAP/Specific Correlations

	Lax PM Acc.	Nonliving Mean RT D.S.	Living Mean RT D.S.	Nonliving μ D.S.	Living μ D.S.	Nonliving τ D.S.	Living τ D.S.
Relaxed PM Accuracy		.03 p = .86	.21 p = .21	< .001 p = 1.00	.31 p = .06	.03 p = .86	.001 p = .99
Nonliving Mean RT D. S.			**.42 p = .01**	**.33 p = .04**	.27 p = .10	**.81 p < .001**	.17 p = .29
Living Mean RT D. S.				-.17 p = .29	-.08 p = .64	**.54 p < .001**	**0.830 p < .001**
Nonliving μ D. S.					**.36 p = .02**	-.29 p = .08	-.25 p = .12
Living μ D. S.						.05 p = .78	**-.60 p < .001**
Nonliving τ D. S.							**.33 p = .04**
Living τ D. S.							

Notes: Difference score is abbreviated D. S. here. Correlation strength from Evans (1996): Very weak = .00-.19; Weak = .20-.39; Moderate = .40-.59; Strong = .60-.79; Very Strong = .80-1.0.

TAP/General Group

PM Accuracy was not significantly correlated traditional measures of RT or μ values. PM accuracy does have a weak correlation with the τ parameter. Traditional RT measures have a moderate relationship with μ values, and a strong relationship with τ values. The μ and τ parameters were not correlated with each other.

This group shows an overall similar pattern of findings to the focal group discussed previously with some of the relationships slightly increasing in strength. See Table 20.

Table 20. TAP/General Correlations

	Lax PM Acc.	Nonliving Mean RT D.S.	Living Mean RT D.S.	Nonliving μ D.S.	Living μ D.S.	Nonliving τ D.S.	Living τ D.S.
Relaxed PM Accuracy		.31 $p = .052$	.29 $p = .07$	.05 $p = .77$	.12 $p = .47$	**.33** **$p = .04$**	.24 $p = .13$
Nonliving Mean RT D. S.			**.64** **$p < .001$**	**.54** **$p < .001$**	.25 $p = .12$	**.78** **$p < .001$**	**.57** **$p < .001$**
Living Mean RT D. S.				**.41** **$p = .01$**	**.39** **$p = .01$**	**.47** **$p = .002$**	**0.89** **$p < .001$**
Nonliving μ D. S.					**.32** **$p = .047$**	-.10 $p = .54$	.29 $p = .07$
Living μ D. S.						.07 $p = .68$	-.04 $p = .83$
Nonliving τ D. S.							**.48** **$p = .002$**
Living τ D. S.							

Note. Difference score is abbreviated D. S. here. Correlation strength from Evans (1996): Very weak = .00-.19; Weak = .20-.39; Moderate = .40-.59; Strong = .60-.79; Very Strong = .80-1.0.

TIP/Specific Group

PM accuracy in the TIP/Specific condition had a moderate-weak correlation with traditional RT measures, and weak relationships with both μ and τ measures, but only for one trial type for each parameter. The traditional RT measures showed a moderate strength correlation with μ values, and very strong correlations with the τ parameter. The μ and τ parameters do not show a consistent relationship with each other for this group. The pattern of correlations in this group show relationships forming were there were none (or very weak) for the TAP conditions. See Table 21 for correlations.

Table 21. TIP/Specific Correlations

	Lax PM Acc.	Nonliving Mean RT D.S.	Living Mean RT D.S.	Nonliving μ D.S.	Living μ D.S.	Nonliving τ D.S.	Living τ D.S.
Relaxed PM Accuracy		**.44** **p = .004**	**.33** **p = .04**	.30 p = .06	**.37** **p = .02**	**.31** **p = .05**	.17 p = .29
Nonliving Mean RT D. S.			**.69** **p < .001**	**.53** **p < .001**	**.50** **p = .001**	**.81** **p < .001**	.50 **p = .001**
Living Mean RT D. S.				**.57** **p < .001**	**.64** **p < .001**	**.42** **p = .01**	**0.85** **p < .001**
Nonliving μ D. S.					**.52** **p = .001**	-.08 p = .64	**.40** **p = .01**
Living μ D. S.						.22 p = .17	.16 p = .32
Nonliving τ D. S.							**.32** **p = .045**
Living τ D. S.							

Note. Difference score is abbreviated D. S. here. Correlation strength from Evans (1996): Very weak = .00-.19; Weak = .20-.39; Moderate = .40-.59; Strong = .60-.79; Very Strong = .80-1.0.

TIP/General Group

PM accuracy was moderately correlated with traditional RT measures, suggesting a functional effect of slowing down ongoing task processing at the expense of PM accuracy. PM accuracy was also significantly related to μ difference scores (moderate strength) as well as τ difference scores (weak strength). The traditional RT measures were moderately related to the μ values, and strongly related to the τ values, however μ and τ values were only very weakly related to each other, and not significantly so. See Table 22.

Table 22. TIP/General Correlations

	Lax PM Acc.	Nonliving Mean RT D.S.	Living Mean RT D.S.	Nonliving μ D.S.	Living μ D.S.	Nonliving τ D.S.	Living τ D.S.
Relaxed PM Accuracy		**.50 $p = .002$**	**.53 $p = .001$**	**.62 $p < .001$**	**.35 $p = .03$**	**.17 $p = .30$**	**.44 $p = .01$**
Nonliving Mean RT Difference Score			**.73 $p < .001$**	**.62 $p < .001$**	**.45 $p = .004$**	**.81 $p < .001$**	.61 **$p < .001$**
Living Mean RT Difference Score				**.69 $p < .001$**	**.56 $p < .001$**	**.42 $p = .01$**	**0.87 $p < .001$**
Nonliving μ Difference Score					**.61 $p < .001$**	.04 $p = .80$	**.47 $p = .003$**
Living μ Difference Score						.12 $p = .46$	.09 $p = .58$
Nonliving τ Difference Score							**.43 $p = .01$**
Living τ Difference Score							

Notes: Correlation strength from Evans (1996): Very weak = .00-.19; Weak = .20-.39; Moderate = .40-.59; Strong = .60-.79; Very Strong = .80-1.0.

CHAPTER 6: EXPERIMENT 1 DISCUSSION

Summary of Experiment 1 Findings

I anticipated finding a main effect of specificity and a main effect of task-appropriateness in PM accuracy, and the actual findings mapped onto that prediction (see Figure 7). Overall, participants carried out the PM task more often in the TAP conditions, and more often when given specific cues to watch for compared to general cues. When given a TIP intention that is general, cue detection was very low, supporting the notion that specificity and task-appropriateness are additive in their interference with the intention.

Overall, no differences were anticipated for ongoing task accuracy, as most costs generally present as RT differences. My analyses found a main effect of trial type, which would persist across most of the subsequent analyses. Participants had higher accuracy rates for their nonliving decisions than their living decisions, which was possibly driven by the nature of the task excluding animals. Initial analyses showed a slight drop in accuracy for the PM block, but the effect was not significant when analyzed with difference scores. This suggests, as predicted, no major cost or changes to the accuracy of the ongoing task decision across any of the groups, as all groups were not significantly different from the control group.

The more sensitive indicator of changes in ongoing task interference—RT differences--were anticipated. I expected to find a main effect of specificity, as well as a main effect of task-appropriateness, with the slowest RTs in the TIP/General condition. Based on the PAM and Delay Theories, all groups were expected to differ from the control, showing task interference with the introduction of the PM task. Conversely, for the multiprocess framework, only the nonfocal conditions were expected to differ from the control. The difference between trial types was again present, with faster RTs in the living trials compared to the nonliving trials. This finding paired with the accuracy finding suggests on the nonliving trials, participants had to take slightly longer to make a decision about animacy, but were more often correct about those decisions, showing the added decision time was functional in their ongoing task performance. A main effect for task-appropriateness was present, showing slightly faster, but significantly different times in the TAP conditions. The TAP conditions did not show elevated RTs compared to the control group, but the TIP conditions did. In comparing both trial types to the control group, the distinction for the focal group became evident—with no RT differences found between the TAP/Specific group and the control, but significant slowing in all three nonfocal groups.

The ex-gaussian analyses again revealed the main effect of trial type, mirroring the finding from the traditional RT measures with a smaller μ in the living trials. An interaction between block and appropriateness also mirrored the traditional findings, with no differences in μ for Block 1, but an increase in μ for the TIP conditions in Block 2, suggesting a greater reliance on a monitoring strategy for the TIP conditions in Block 2. When compared to the control group, the TIP conditions differed for both living and nonliving trials in Block 2. No difference with the control group was found for either trial type for the TAP/General condition (unexpectedly), and the TAP/Specific group only showed a difference with the control group for the nonliving trials. This finding may relate to the task-specific interference principle with the TAP conditions being able to form a more efficient strategy for detecting their cue as they knew it would appear in a living trial. This strategy may have been more obvious for the *any* animal group (TAP/General) as they are already thinking of entire groups of living things for their task (humans, animals, plants).

Interestingly in the baseline block analysis for μ, a main effect of specificity was found, most likely driven by the TAP/General group having very low μ values for both blocks. In looking at the difference scores between blocks, only a main effect of task-appropriateness was present, again reflecting the slower responses from the TIP condition after receiving their PM intention. The TAP/Specific condition showed no differences from the control, suggesting no ongoing task interference in this focal condition. The TAP/General group showed a difference from the control only in living trials, while the two TIP conditions showed differences in μ for both trial types indicating significant changes in the RT distributions—and likely greater overall monitoring for the PM cue—for these groups compared to the control.

In the analyses for the τ parameter, the main effect of trial type was present again, with smaller τ values for the living trials, again reflecting the pattern that the nonliving trials are taking more time before a decision is made. The values for τ were much larger in Block 2 when looking at just the PM conditions, suggesting endorsement for sporadic retrieval after receiving a PM intention. Figure 14 shows no differences between living and nonliving τ values for the TAP/Specific group, which suggests that their intention mitigated the prevalent trial type effect and eliminated their need to attend to nonliving trials, even intermittently, which is especially interesting considering the control group still showed the trial type effect. Having the TAP/Specific intention did not lead to differences in τ compared to the control group for either trial type in Block 2, while the three nonfocal conditions differed from the control group in the nonliving trials (and living trials for the TIP conditions). When difference scores for τ were compared to the control condition, the TIP conditions were significantly different. Interestingly, in group comparisons, all three nonfocal groups showed a significant difference compared to the control group for nonliving trials. The τ value did not change across blocks for the control group, strengthening the idea that τ could be a metric for the PM strategy relying on spontaneous retrieval. All of the nonfocal groups showed an increase in τ only after receiving their intention.

These findings paired with the findings from the μ analyses display consistent increases in μ and τ for the TIP conditions, but little changes in the TAP groups—especially the TAP/Specific group—for either parameter. These overall patterns were mostly driven by the performance on the nonliving trials.

Both of the RT analyses—traditional methods as well as ex-gaussian parameters—showed a significant change for the TIP groups after receiving an intention. The distinction between the two TAP conditions was only evident in the traditional RT analyses when comparing to the control group. However, in the ex-gaussian parameters, the perhaps more nuanced differences between the two TAP conditions became more evident.

The general lack of main effect for specificity in RT analyses was not anticipated. The impact specificity has on PM accuracy was not reflected in the RT analyses, possibly suggesting this impact on focality is only evident in PM accuracy, and that the ongoing task performance is not as impacted by differences in cue specificity. This could be a source of some of the discrepant findings and theoretical motivations related to focal versus nonfocal costs occasionally being found or absent in previous PM work when using cue specificity as the sole distinction between focal and nonfocal conditions.

Correlations between the PM accuracy and RT cost measures show no overall relationship between costs and PM performance for the TAP conditions, reflecting the patterns presented in the analyses showing virtually no costs for the TAP conditions but higher PM accuracy overall. Any costs present in these conditions are nonfunctional in improving PM detection. For the TIP conditions however, and especially true for the TIP/General condition, correlations were present. This suggests the costs incurred on the ongoing task led to greater PM detection, demonstrating functionality to the increased RTs.

Collectively, these findings support the overall purpose of this experiment—showing that both task-appropriateness and cue specificity are important to consider for PM accuracy. The cost findings are consistent with the existing literature in that no costs were found in ongoing task accuracy, and that there was slowing evident in the RT analyses. The cost findings more closely resemble the predictions established by the multi-process framework, with no differences found between the focal condition and the control condition but slowing observed in all nonfocal conditions.

Study Conclusions

The results of this study indicate that PM researchers must be more careful in how they define focality in future work. Explicit classification of both cue specificity and task-appropriateness are warranted in future studies of focality, as the patterns of findings, especially in ongoing task costs, differ (e.g., costs are more prominent in task-appropriateness differences rather than cue specificity conditions). How researchers define focality impacts the differences expected between the nonfocal and focal group, including RT measures of cost that may further relate to people's metacognitive appreciation of task difficulty. As shown by this first experiment, nonfocal tasks can vary in their relationship to the ongoing task, and that variability in nonfocal tasks lead to different levels of PM deficits and costs to the ongoing task.

The experiment also speaks to the general nature of focality. Understanding the degree of influence each component of focality—cue specificity and task-appropriateness—holds on PM performance and ongoing task costs can move the PM field towards more nuanced experimental designs capable of more detailed predictions. As demonstrated by the RT parameters especially in the TAP conditions, ex-Gaussian modeling may be the most sensitive to differences in ongoing task RTs.

One issue with this first experiment lies in the use of semantic and orthographic tasks to form the TAP/TIP distinction. Prior work examining different levels of processing show that semantic processing often leads to better performance overall. Therefore, by using a semantic ongoing task, we are possibly finding the benefits of TAP due to the deeper level of processing, rather than the match between the PM task and the ongoing task. In order to tease apart the influence of task-appropriateness and levels of processing, a second experiment was conducted to generalize these effects in a different ongoing task. The task was orthographic in nature, so that a TIP condition could use semantic PM cues. If the same task-appropriateness main effect was replicated in the second experiment (showing that TAP yields better performance than TIP), then the findings of both experiments are due to task-appropriateness rather than a deeper, semantic level of processing. If TIP leads to greater performance than TAP in the second experiment, then levels of processing influenced PM performance, perhaps even more than task-appropriateness. Additionally, the second experiment allowed for a chance to replicate the findings from Experiment 1 using a different ongoing task, but with all of the same PM cues and ongoing task stimuli.

CHAPTER 7: EXPERIMENT 2 METHOD

The second experiment manipulated the cue specificity (specific or general) and the task-appropriateness (TAP or TIP) of a PM cue during an orthographic ongoing task in order to develop a full picture of how focality relates to PM accuracy and ongoing task costs. Institutional Review Board approval was obtained through Louisiana State University prior to data collection (see Appendix C). The methods used in Experiment 2 were identical to Experiment 1, so an abbreviated version of the method section is included here with any differences from Experiment 1 mentioned explicitly. Please see Appendix B for the full method section for Experiment 2. Data cleaning occurred simultaneously with Experiment 1 and was described in Chapter 5 (see also Table 8).

Participants

Participants recruited for this study were selected in the same manner as Experiment 1. The sample size estimations were identical to the first experiment, with a goal set at 40 people per group totaling 200 participants for Experiment 2, with the total n = 206.

Materials

The materials used for experiment 2 were identical to those used in Experiment 1.

Procedure

The procedure for the second experiment was identical to Experiment 1, except on the following points. The orthographic ongoing task asked participants to classify the ongoing task words based on how many enclosed spaces were in each word: equal to and greater than 3 or fewer than 3. For example, the word “rooster” has 3 enclosed spaces—one in each letter “o” and one in the letter “e” (see Figure 1). The TAP and TIP conditions in this second experiment were reversed from Experiment 1 in order to match the orthographic processing used in the ongoing task. Now, the TAP conditions required participants to either press “y” when they saw two consecutive double letter o’s (specific condition), or to press “y” when they saw any two consecutive letters, or doubles (general condition). The TIP/Specific condition required participants to either press “y” when they saw either of two specific cue words (counterbalanced across moose, goose, baboon, rooster, raccoon, or kangaroo). The TIP/General instructions directed participants to press “y” when they saw any animal word, excluding humans.

Design

Again, identical to Experiment 1, this experiment manipulated two independent variables (cue specificity and task-appropriateness) with two levels each (specific and general, TAP and TIP, respectively), all between-subjects, and compared those groups to a control condition with no PM intention. Table 5 provides a schematic of the conditions included in this experiment. Additional variables of interest built into the design as within-subjects manipulations included the trial type (living and nonliving) and block (baseline block and PM block), which became important when studying ongoing task accuracy, and ongoing task RTs.

CHAPTER 8: EXPERIMENT 2 ANTICIPATED RESULTS

The anticipated results of the second experiment were nearly identical to those of the first experiment. As the only difference was the ongoing task used, the same principles and expectations held for the predictions in the second experiment.

PM Accuracy

Identical to Experiment 1, I scored participants on their PM accuracy as a percentage of the 4 possible PM cues they detected. Using just two PM cues presented twice each allowed for a proportion for PM detection while also avoiding too frequent of cues that could have turned the task into a vigilance task (see Graf, Uttl, & Dixon, 2002). Additionally, the use of two cues reduced monitoring as much as possible as the number of cues participants in the focal condition were asked to remember was minimal yet still offered cue variability (see Rummel et al., 2016 Experiments 2 and 3; Einstein et al., 2005). I distinguished between participants that failed to recall the PM instructions in the post experiment questionnaire and those that were only able to recognize their instructions (excluding participants that failed both recall and recognition checks), as their failures indicated retrospective memory failure rather than PM failure. I compared their accuracy scores in a 2 (cue specificity) x 2 (task-appropriateness) ANOVA. Anticipated findings indicated a main effect of cue specificity in which specific cues led to higher PM performance than general cues. I also expected a main effect of task-appropriateness, with higher PM performance in the TAP conditions than the TIP conditions. No interaction was expected, though if an interaction was present, I anticipated it would indicate that focal performance was highest, with all of the three nonfocal conditions showing worse PM performance. If the predicted results were found, they would offer support for the notion that task-appropriateness has a stronger impact on PM performance than levels of processing.

Importantly, if the level of processing was the stronger factor influencing PM performance, the TIP conditions would show higher performance than the TAP conditions in this experiment, signifying that semantic processing was more advantageous than orthographic, regardless of the ongoing task demands.

Ongoing Task Accuracy

Again, PM instructions affected ongoing task accuracy inconsistently across prior research. I suspected that very few statistical differences would exist, but the following predictions would result assuming that intentions produced a cost on the ongoing task performance. The ongoing task accuracy excluded PM trials. I examined ongoing task accuracy using a 2 (cue specificity: specific/general) x 2 (task-appropriateness: TAP/TIP) x 2 (trial type: 3 or more/ less than 3) x 2 (block: baseline/PM block) mixed model ANOVA, where block and trial type are within-subjects variables and cue specificity and task-appropriateness are between-subjects variables. The expected results are displayed in Figures 19 and 20.

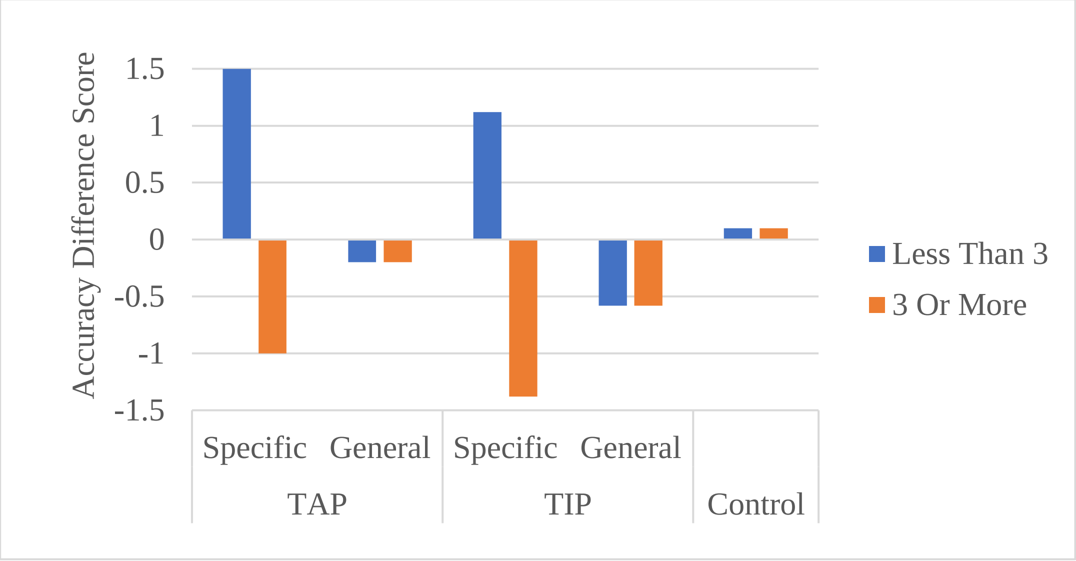

Figure 19. Predicted Experiment 2 Ongoing Task Accuracy Difference Scores.
Note. Classified by trial type to allow for Delay Theory Predictions (Heathcote et al., 2015) with task specific interference, as found by Strickland et al. (2017)

I calculated a difference score for each participant by taking the average ongoing task accuracy percentage from the first baseline block and subtracting the ongoing task accuracy percentage from the PM (second) block. I anticipated finding a main effect of task-appropriateness, indicating a larger decrease in ongoing task accuracy for the TIP conditions compared to the TAP conditions when in the PM block as compared to the baseline block (see Figure 20).

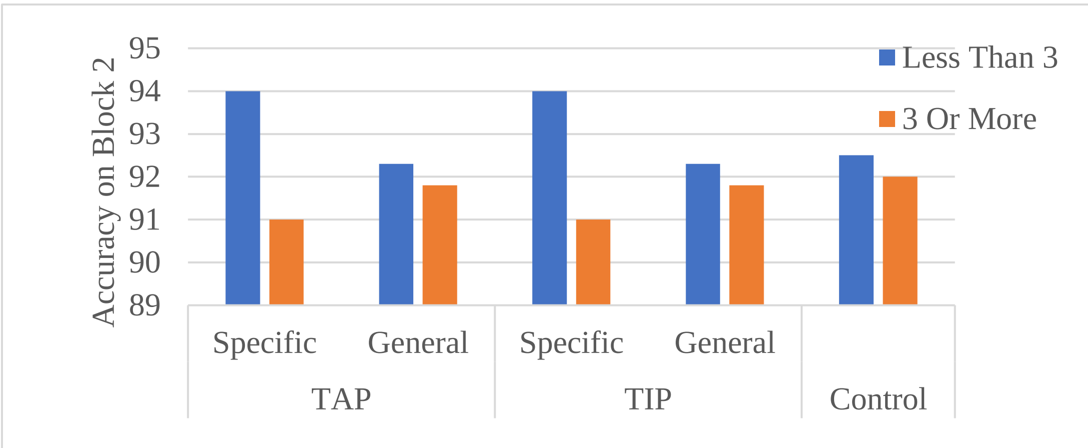

Figure 20. Predicted Experiment 2 Ongoing Task Accuracy in the PM Experimental Block. Note. Classified by trial type to allow for Delay Theory Predictions (Heathcote et al., 2015) with task specific interference, as found by Strickland et al. (2017).

I also expected to find a main effect of cue specificity, with worse accuracy in the general, nonfocal conditions (see Figures 19 and 20). A main effect of trial type was expected with "3 or more" trials leading to worse accuracy than "less than 3" trials (see Figure 19). This prediction was based on the Delay Theory expectations of task specific interference disrupting the trials that are similar to the PM cues (see Strickland et al., 2017). Additionally, I expected to see a main effect of block type, with worse accuracy overall during the second block with the inclusion of the PM intention, compared to the baseline.

I expected to find a significant interaction between block and cue specificity, where the PM block impacted the general condition more than the specific condition. Similarly, an interaction between block and task-appropriateness was expected, where the PM block impacted the TIP conditions more than the TAP conditions. I did not expect to find an interaction between block and trial type (see Figure 20). I did expect to see an interaction between trial type and cue specificity, in that “less than 3” trials would show higher accuracy, especially for the specific condition. The “3 or more” trials were expected to show worse performance, especially for the specific condition.

No interaction was expected between task-appropriateness and cue specificity for ongoing task accuracy. And finally, I did not anticipate finding any three-way or four-way interactions.

Importantly, I also carried out planned comparisons between the control condition and each of the four PM conditions. Differences were expected in the difference scores demonstrating the overall impact of holding an intention, with each PM condition expected to show worse performance than the control, with the least difference found in the focal (specific/TAP) condition. When examining the “3 or more”/ “less than 3” trial type performance, an interaction with cue specificity was expected to show no differences between the control group and the general conditions. The specific conditions were expected to have worse “3 or more” trial accuracy compared to the control group, yet better accuracy compared to the control group for the “less than 3” trials (based on Strickland et al., 2017; see Figure 19). If these differences were not present, then I hypothesized it might be a result of participants not consciously recognizing how their PM cue would fit with the ongoing task directions. Additionally, I suspected if this would appear, it would most likely have been in this second experiment as the task was orthographic rather than semantic, and the number of enclosed spaces a word has is not as common of a classifying method as animacy.

Ongoing task RTs

I prepared the RTs for the ongoing task in a similar manner to Brewer (2011). I compiled the RTs during the baseline block and the PM block separately. Any trials with incorrect decisions were excluded from the RT analysis. Additionally, RTs on the PM trials and the three trials following the PM cue were excluded as per Brewer's (2011) recommendations. Each participant's mean RT was calculated for each block of trials, and any RT score outside of 2.5 standard deviations was excluded from analyses. Each participant had two RT means calculated through this process: their baseline RTs and their PM block RTs. These RT values were also subjected to ex-Gaussian modeling as described later.

A difference score between the two RT means was calculated for each participant and subjected to a 2 (cue specificity) x 2 (task-appropriateness) ANOVA (see Figure 4). I expected a main effect of cue specificity, with more slowing for the general conditions compared to the specific conditions. I also expected to find a main effect of task-appropriateness, with more slowing on the TIP conditions compared to the TAP conditions. I did not anticipate finding an interaction for ongoing task RTs. However, I recognized the possibility that an interaction could have been present, in that the focal condition (Specific/TAP) was the fastest RT, with the other three nonfocal conditions (Specific/TIP; General/TAP; General/TIP) all equally slow. This finding would have aligned with the notion that focality behaves like a dichotomy between focal and nonfocal conditions.

I conducted planned comparisons between the control condition and each of the four conditions of the focality manipulations. All of the PM groups were expected to show slowing in the ongoing task compared to the control condition. I expected to see that the RT slowing was least extreme for the focal (specific/ TAP) condition, and most extreme for the nonfocal (general/

TIP) condition. I anticipated that the other two nonfocal conditions (General/ TAP and Specific/ TIP) would show moderate amounts of slowing compared to the control condition. These hypothetical findings would support the PAM theory and Delay Theory, with each intention condition still showing some slowing compared to the control group. Alternatively, if the results showed no cost differences between the focal condition and the control condition, this would support the multiprocess framework and the dynamic multiprocess framework.

Ex-Gaussian Analyses

Similar to the RT preparations suggested by Brewer (2011), each participant's RT scores were entered into the QMPE software (Heathcote et al., 2004) separately in order to calculate estimates for μ, τ, and σ. Fixed quantiles were established (0, .2, .4, .6, .8, 1) to estimate each participants' parameters. RTs were separated for "3 or more" trials and "less than 3" trials.

Each model parameter was subjected to a 2 (block: baseline/ PM block) x 2 (trial type: 3 or more/ less than 3) x 5 (Group: 4 PM groups and control group) mixed methods ANOVA to see how the parameters changed with the inclusion of an intention.

The first parameter of interest, μ, is the RT distribution mean. Acting on the supposition that μ is influenced by active monitoring (e.g., a target checking strategy; Guynn, 2003), a main effect of group was expected with the nonfocal PM conditions predicted to show higher μ values, with lower μ values for the focal and control condition. I expected an interaction between group and block, as the control group (and possibly the focal group) was expected to have similar μ values across both blocks, and all of the PM nonfocal groups were expected to increase in the PM block. I also expected to find a three-way interaction between trial type, block, and group. The predicted outcome of this interaction was that the μ values for "3 or more" and "less than 3" trials for the two TIP conditions would increase during the PM block, similar μ values for the "3 or more" and "less than 3" trials for the control condition would decrease slightly in the PM block due to practice effects, and the μ value for the "3 or more" trials would increase in the two TAP groups when in the PM block. This interaction would have supported the Delay Theory.

The second parameter of interest was τ, which describes the tail distribution. I anticipated an interaction between group and block. I did not expect that the control condition would show any differences in τ between block, but τ was expected to increase for all PM conditions. The tail distribution was expected to be especially large in the nonfocal conditions, with possibly no differences, or small differences between the focal and control condition. The differences among the three nonfocal conditions were highly interesting, as this comparison was not present in the PM literature.

I did not expect to find any differences in the third parameter, σ, or the RT deviation, among any of the groups.

Additionally, the ex-Gaussian parameters and the overall ongoing task RTs were correlated with PM accuracy as a metric for functionality of the costs (see Loft, Bowden, Ball & Brewer, 2014). If the costs were positively correlated with PM accuracy, then those that completed the ongoing task more slowly had better PM accuracy, demonstrating that the slowing served as a benefit. However, greater costs could also be dysfunctional, and be related to worse PM performance, as was the case for Ihle et al.'s (2016) nonfocal condition (which used a TAP/ TIP distinction to form conditions; Ihle, Ghisletta, & Kliegel, 2016).

CHAPTER 9: EXPERIMENT 2 ACTUAL FINDINGS

PM Accuracy

Identical to the analyses for Experiment 1, two PM proportions were calculated as either strict or lax, with the lax criterion allowing for late PM responses. The post-experiment questions (Question 2 and Question 5) were used to classify participants as able to freely recall the PM task independently as well as recognize the PM intention correctly out of a list of plausible PM instructions. Twenty-one participants failed to recall their PM intention (n=124 for successful participants) and 16 participants failed to recognize their PM intention instructions (successful participants n = 129).

I first carried out a 2 (specificity) x 2 (appropriateness) ANOVA on PM proportion correct strictly on PM trials, and only included participants that correctly recalled the PM intention in the post-experiment questions. In contrast to Experiment 1, there was no main effect of appropriateness: $F(1,120) = 1.62$, $p = .21$, $\eta_p^2 = .01$. However, congruent with Experiment 1's findings, I found a significant main effect of cue specificity: $F(1,120) = 10.31$, $p = .002$, $\eta_p^2 = .08$. PM accuracy was also significantly higher in the specific condition (M =.76, SE = .05) compared to the general condition (M =.56; SE = .04). There was no significant interaction, although it was approaching significance: $F(1,120) = 3.66$, $p = .058$, $\eta_p^2 = .03$. See Figure 21 for a visualization of these findings.

This pattern of findings remained stable when looking at the lax trial response window and task recall in post instructions, as well as analyses using PM responses in the lax trial response window and relaxed recognition criteria based on post-experiment question responses. However, when PM accuracy was measured using the strict trial response window (only PM trials) and all participants with successful recognition of the PM instructions were included in the analyses, the interaction between appropriateness and cue specificity became significant [$F(1,125) = 4.05$, p = .046, $\eta_p^2 = .03$]. For a full report of these corroborative analyses, please see

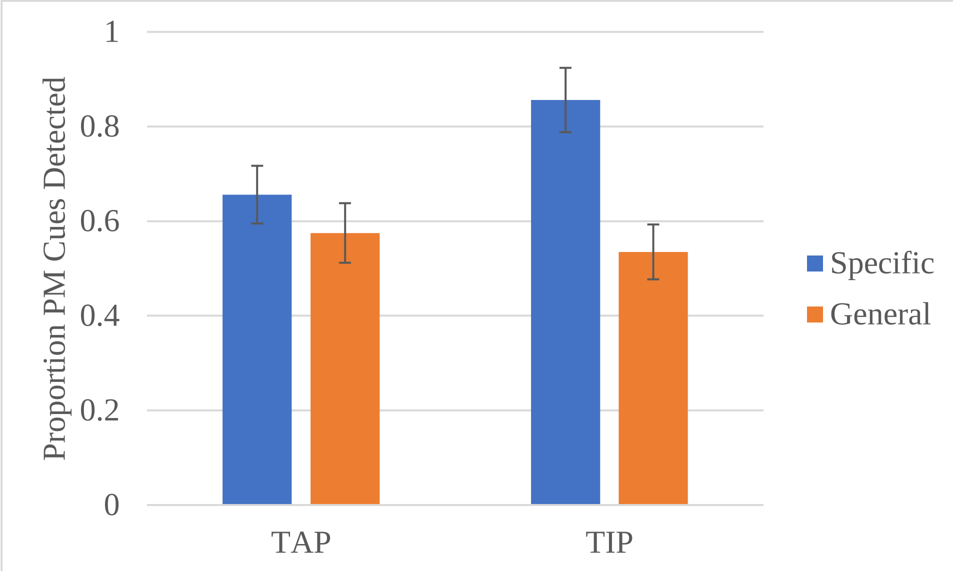

Figure 21. PM Accuracy using a strict response window and participants with correct recall.

Table 23. Based on the stability of the effect size, and the consistency of the *p*-values hovering around the .05 mark, the pattern appears consistent across all combinations of strict and lax criterion for measuring PM Accuracy- namely consisting of a significant effect of cue specificity, a washout of task-appropriateness, and a borderline significant interaction.

Post-hoc contrasts of the interaction examined the differences between levels of task-appropriateness for both the specific conditions and the general conditions. In the specific comparison, TAP and TIP were significantly different [$F(1,122) = 5.64$, $p = .02$, $\eta_p^2 = .04$]. The two general conditions were also significantly different from each other [$F(1,122) = 4.32$, $p = .04$, $\eta_p^2 = .03$], though to a smaller magnitude.

Table 23. Corroborative Analyses to Examine PM Accuracy

Experiment 2						
	df	*F* value	*p* value	η_p^2	Power	Result
Lax PM Window with Strict Recall						
Appropriate	(1,120)	1.59	.21	.01	.24	TAP = TIP
Specificity	(1,120)	6.08	**.02**	.05	.69	Specific > General
Interaction	(1,120)	2.92	.09	.02	.40	None
Strict PM Window with Lax Recognition						
Appropriate	(1,125)	1.37	.24	.01	.21	TAP = TIP
Specificity	(1,125)	8.16	**.005**	.06	.81	Specific > General
Interaction	(1,125)	4.05	**.046**	.03	.52	Specific > General: especially so in TIP conditions
Lax PM Window with Lax Recognition						
Appropriate	(1,125)	1.51	.22	.01	.23	TAP = TIP
Specificity	(1,125)	4.96	**.03**	.04	.60	Specific > General
Interaction	(1,125)	3.56	.06	.03	.47	none

Ongoing Task Accuracy

In similar fashion as Experiment 1, I carried out a 2 (cue specificity) x 2 (appropriateness) 2 x (trial type: living and nonliving) x 2 (block: baseline and PM block) mixed model ANOVA on the number of correct decisions participants made for the ongoing task to see if their accuracy changed with the addition of a PM intention. A main effect of trial type was present, $F(1,122) = 52.29$, $p < .001$, $\eta_p^2 = .30$, indicating that participants had better accuracy on the "less than 3" trials ($M = .90$, $SE = .01$) compared to the "3 or more" trials ($M = .83$, $SE = .01$). There was also a main effect of Block [$F(1,122) = 11.81$, $p = .001$, $\eta_p^2 = .09$] with accuracy being slightly higher in Block 2 ($M = .87$, $SE = .01$) compared to Block 1 ($M = .86$, $SE = .01$). Additionally, I found an interaction between trial type and block, $F(1,122) = 11.31$, $p = .001$, $\eta_p^2 = .09$. This interaction is displayed in Figure 22. Pairwise comparisons show a significant difference between Block 1 and 2 for the "3 or more" trial type ($p < .001$), but no significant difference between Blocks for the “less than 3” trial type (*p* = .767) No other main effects, 2-way, 3-way, or 4-way interactions were present (all *p*’s > .09).

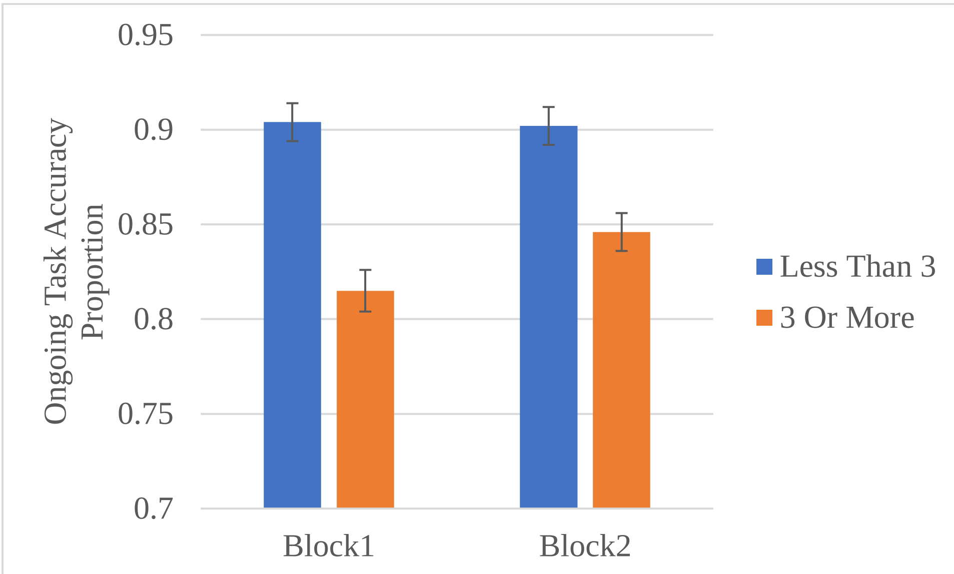

Figure 22. Ongoing task accuracy interaction between trial type and block.

For a closer examination of possible differences between the two blocks, a difference score was calculated by subtracting the accuracy proportion score for Block 1 from Block 2. A 2 (specificity) x 2 (appropriateness) x 2 (trial type) mixed model ANOVA was then carried out on the ongoing task accuracy difference proportions. There was a significant main effect of trial type, $F(1,122) = 11.31$, $p = .001$, $\eta_p^2 = .09$. This main effect revealed an insignificant drop in accuracy for the ‘less than 3’ trials from Block 1 to Block 2 (*M* = -0.002, *SE* = .007) and a slight gain in accuracy from Block 1 to Block 2 for the ‘3 or more’ trials (*M* = .031, *SE* = .006). No other main effects or interactions reached significance.

Planned comparisons were carried out to look at each experimental condition against the control group that had no PM intention for the duration of the experiment. There were no significant differences between any of the experimental groups and the control group. Details for the planned comparisons can be found in Table 24.

Table 24. Ongoing task accuracy comparisons with the control group.

	Mean Difference	*SE*	Dunnet *t* value	*p* value	Cohen's d (small = .02, medium = .05, large = .08)
Block 1 Accuracy					
TAP/Specific vs. Control	0.043	0.02	1.86	0.20	0.45
TAP/General vs. Control	0.029	0.02	1.22	0.56	0.29
TIP/Specific vs. Control	0.033	0.02	1.39	0.44	0.32
TIP/General vs. Control	0.027	0.02	1.17	0.59	0.29
Block 2 Accuracy					
TAP/Specific vs. Control	0.042	0.02	1.71	0.26	0.39
TAP/General vs. Control	0.041	0.03	1.60	0.32	0.38
TIP/Specific vs. Control	0.032	0.03	1.25	0.54	0.30
TIP/General vs. Control	0.026	0.02	1.04	0.69	0.24
Difference Score Accuracy					
TAP/Specific vs. Control	0.006	0.01	0.52	0.96	0.13
TAP/General vs. Control	-0.008	0.01	-0.66	0.92	-0.17
TIP/Specific vs. Control	0.003	0.01	0.29	1.00	0.09
TIP/General vs. Control	0.008	0.00	6.80	0.91	0.15

Ongoing Task RTs

Analyzing Baseline Performance

The ongoing task RT analyses were carried out in an identical manner to those of Experiment 1, using both the RT means and the RT medians. This section reports the findings for the RT means and any differences found in the median analyses are noted to highlight discrepant findings.

Analyzing Block 1 RT means allows for an examination of the baseline differences between groups and provides a clearer picture of the costs that can be attributed to the addition of an intention. I carried out a 2 (specificity) x 2 (appropriate) x 2 (trial type) mixed model ANOVA on the mean RTs for Block 1. There was a main effect of trial type, $F(1,141) = 54.10$, $p < .001$, $\eta_p^2 = .28$, with faster responses for the "3 or more" trials ($M = 1608.88$, $SE = 40.21$) compared to the "less than 3" trials ($M = 1757.69$, $SE = 46.65$). There were no other main effects or interactions, signifying overall steady baseline performance across all experimental groups. Additional comparisons were carried out to compare the experimental groups to the control condition for both the "3 or more" and "less than 3" trial types. Again, no significant differences were found in comparison to the control group for either trial type.

Analyzing the PM Block

The same analyses as above were repeated looking at Block 2 means to examine changes resulting from the added PM intention. A 2 (specificity) x 2 (appropriate) x 2 (trial type) mixed model ANOVA was carried out on the RT means of Block 2. The main effect of trial type reappeared, $F(1,141) = 57.80$, $p < .001$, $\eta_p^2 = .29$, again showing faster RTs for the "3 or more" trials ($M = 1622.04$, $SE = 35.17$) compared to the "less than 3" trials ($M = 1771.86$, $SE = 44.51$). In contrast to Experiment 1, there was no main effect of appropriateness, $F(1,141) = 2.70$, $p = .10$, $\eta_p^2 = .02$. As another contrast to Experiment 1, no interactions were significant in these analyses. The overall null pattern of findings is presented in Figure 23.

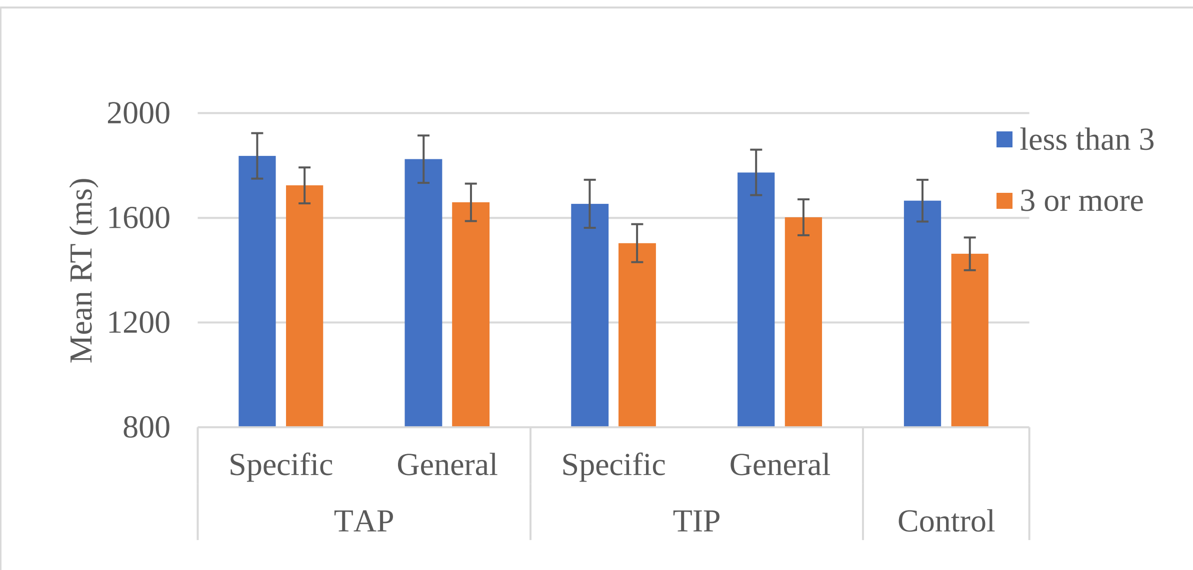

Figure 23. Main effect of trial type across groups.

All experimental groups were compared to the control group in planned comparisons. The ‘less than 3” trial type showed no difference between the control group and any of the experimental groups [$F(4,178) = .98$, $p = .42$]. However, the “3 or more” trial type showed an overall significant difference between the control group and the experimental group [$F(4,178) = 2.50$, $p = .04$], with only the TAP/Specific group showing significant differences from the control group. Details of the follow-up analyses are displayed in Table 25. The pattern of findings was comparable when looking at median RT values with just a minor difference in the outcome for the TAP/Specific condition’s comparison—the group failed to reach significance when compared to the control group for the “3 or more” trial type analyses ($p = .056$). However, Cohen’s *d* when comparing the TAP/Specific to the control condition for the “3 or more” trials still reached a medium effect size (d = .59), suggesting the pattern still persists in the analyses of the median RTs.

Table 25. Planned comparisons with the control group in the PM block.

	Mean Difference	*SE*	Dunnet *t* value	*p* value	Cohen's d (small = .2, medium = .5, large = .8)
less than 3 Mean RTs					
TAP/Specific vs. Control	171.07	120.78	1.42	0.43	0.34
TAP/General vs. Control	158.48	123.34	1.28	0.51	0.29
TIP/Specific vs. Control	-12.01	124.28	-0.10	1.00	-0.02
TIP/General vs. Control	108.02	120.78	0.89	0.79	0.22
3 or more Mean RTs					
TAP/Specific vs. Control	261.44	95.30	2.74	**0.02**	0.64
TAP/General vs. Control	196.66	97.32	2.02	0.14	0.44
TIP/Specific vs. Control	40.85	98.06	0.42	0.98	0.11
TIP/General vs. Control	139.60	95.30	1.46	0.40	0.37

Difference Score Analyses

A difference score was calculated by subtracting the mean RTs of Block 1 from Block 2; this difference score was then analyzed with a 2 (specificity) x 2 (appropriateness) x 2 (trial type) mixed model ANOVA. Identical to Experiment 1, the main effect of trial type that was significant in each block measured alone was washed out using the difference score, showing that the overall pattern of RT speed on the two trial types was consistent across both blocks, $F(1,141) = 0.003$, $p = .95$, $\eta_p^2 < .001$. In contrast to Experiment 1, there was no main effect of appropriateness, no interaction between trial type and appropriateness, and no 3-way interaction between trial type, appropriateness, and specificity. However, an interaction between appropriateness and specificity was present, $F(1,141) = 4.85$, $p = .03$, $\eta_p^2 = .03$. Pairwise comparisons found a significant difference between the two specific conditions [F(1,141) = 8.03, $p = .005$, $\eta_p^2 = .05$], but no significant differences between the two general conditions ($p = .79$)[4]. When analyzed using medians, this interaction lost significance ($p = .056$), but retained a similar effect size ($\eta_p^2 = .03$). This interaction is visible in Figure 24.

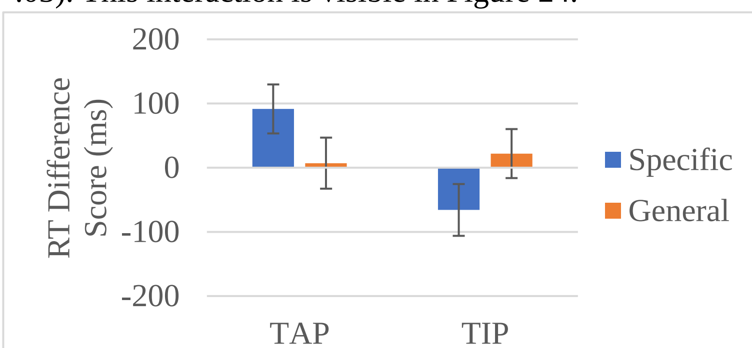

Figure 24. Interaction between appropriateness and specificity for RT difference scores.

---

[4] The pattern of findings differed when Block 1 performance collapsed across trial type was used as a covariate to account for individual differences. A main effect of trial type was present, ($F(1,140) = 13.32$, $p < .001$, $\eta_p^2 = .09$, with faster RTs in the "3 or more" trials ($M = 1622.11$, $SE = 18.78$) compared to the "less than 3" trials ($M = 1771.95$, $SE = 21.20$). There was a main effect of appropriateness[$F(1,140) = 4.64$, $p = .03$, $\eta_p^2 = .03$] with faster RTs in the TIP conditions ($M = 1658.11$, $SE = 25.63$) compared to the TAP conditions ($M = 1735.95$, $SE = 25.44$). The interaction between specificity and appropriateness was replicated in this ANCOVA [F(1,140) = 5.15, $p = .03$, $\eta_p^2 = .04$] following the same pattern presented from the difference scores.

The planned comparison follow-up test revealed significant differences between the experimental groups and the control group, but only for the "3 or more" trials [$F(4,178) = 6.13$, $p < .001$] and not for the "less than 3" trials [$F(4,178) = 1.68$, $p = .16$]. These findings show the nonfocal group (TIP/Specific) is the only group that is not different from the control group for the "3 or more" trial analyses (see Table 26). Interestingly, this is the experimental group that received the specific animal names (i.e. goose).

Table 26. Mean RT difference score follow up tests with control group.

	Mean Difference	*SE*	Dunnet *t* value	*p* value	Cohen's d (small = .2, medium = .5, large = .8)
LessThan3					
TAP/Specific vs. Control	142.04	58.98	2.41	0.06	0.55
TAP/General vs. Control	66.22	60.23	1.10	0.65	0.25
TIP/Specific vs. Control	23.11	60.69	0.38	0.99	0.09
TIP/General vs. Control	69.83	58.98	1.18	0.58	0.27
3 Or More					
TAP/Specific vs. Control	253.89	57.79	4.39	**<.001**	1.03
TAP/General vs. Control	160.85	59.01	2.73	**0.03**	0.61
TIP/Specific vs. Control	58.11	59.46	0.99	0.73	0.26
TIP/General vs. Control	187.02	57.79	3.24	**0.005**	0.80

Ex-Gaussian Analyses

Identical to Experiment 1, the RT data was also analyzed using ex-Gaussian parameters to examine the shape of each participant's RT distribution and look for changes that reflected different strategy use in different conditions. This section of the results looks at μ first, then τ, and σ.

Analyses for μ

I carried out a 2 (specificity) x 2 (appropriateness) x 2 (trial type) x 2 (block) mixed model ANOVA on the μ parameter. A main effect of trial type was present, $F(1,141) = 10.84$, *p*

< .001, η_p^2 = .07, with “3 or more” trials (*M* = 1065.11, *SE* = 24.99) showing a smaller μ than “less than 3” trials (*M* = 1120.37, *SE* = 30.83). This finding parallels the traditional RT findings with an average faster RT score for the “3 or more” trials compared to the “less than 3” trials. There was also an unexpected main effect of block: *F*(1,141) = 4.78, *p* = .03, η_p^2 = .03, with larger μ values in Block 1 (*M* = 1112.53, *SE* = 30.22) compared to Block 2 (*M* = 1072.94, *SE* = 26.18).

Comparisons with the control group show no overall differences for either trial type during either block, showing that performance across groups was not different when compared to the control.

Baseline Analyses for μ

A 2 (specificity) x 2 (appropriateness) x 2 (trial type) mixed model ANOVA was carried out to look at the μ parameter only in Block 1. There was a main effect of trial type, *F*(1,141) = 4.07, *p* = .046, η_p^2 = .03, again with smaller μ values for the “3 or more” trials (*M* = 1090.33, *SE* = 29.29) compared to the “less than 3” trials (*M* = 1134.74, *SE* = 34.79). An interaction between trial type and specificity was also present, *F*(1,141) = 4.41, *p* = .04, η_p^2 = .03. The interaction is displayed in Figure 25. Pairwise comparisons found no differences between the two trial types in

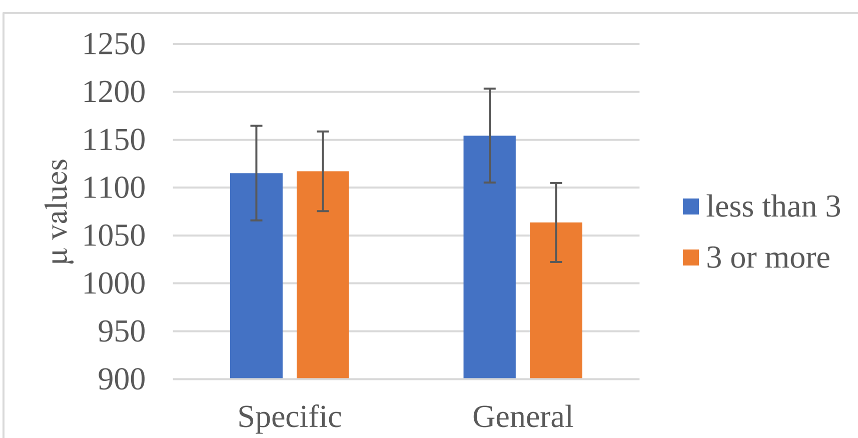

Figure 25. Baseline block interaction between specificity and trial type for μ.

the specific conditions (p = .95) but a significant difference between trial types in the general conditions [$F(1,141) = 8.54$, $p = .004$, $\eta_p^2 = .06$].

No differences were found when comparing the experimental groups to the control group for the "less than 3" trials ($F(4,178) = .38$, $p = .82$) or for the "3 or more" trials ($F(4,178) = .42$, $p = .80$).

Difference Scores for μ

Difference scores were created by subtracting the μ for Block 1 from the μ parameter for Block 2 (Block 2 – Block 1). The resulting score was analyzed in a 2 (specificity) x 2 (appropriateness) x 2 (trial type) mixed model ANOVA. There were no main effects or interactions for the μ parameter (all p's > .09)[5].

In comparison with the control group, none of the experimental groups reached significance for the "less than 3" trials; but for the "3 or more" trials, the TIP/General condition did reach significance indicating that this group differed from the control group. These values are presented in Table 27.

Group Analyses for μ

To fully explore group differences, I conducted a 2 (block) x 2 (trial type) x 5 (group) mixed model ANOVA on μ. There was a main effect of trial type [$F(1,178) = 11.62$, $p = .001$, $\eta_p^2 = .06$] that showed smaller μ values for the "3 or more" trials (M = 1055.40, SE = 22.03) compared to the "less than 3" trials (M = 1109.40, SE = 28.00). Additionally, a main effect of block was present, $F(1,178) = 12.80$, $p < .001$, $\eta_p^2 = .07$. Block 2 had smaller μ values (M =

[5] When the μ parameters from Block 1 were averaged across trial type and used as a covariate, a main effect of trial type was present [$F(1,152) = 10.77$, $p = .001$, $\eta_p^2 = .07$] with smaller μ values found in the living trials (M = 558.56, SE = 4.22) compared to the nonliving trials (M = 610.93, SE = 6.02). There was also a main effect of appropriateness, $F(1,152) = 8.67$, $p = .004$, $\eta_p^2 = .05$, with larger μ parameters in the TIP conditions (M = 597.50, SE = 6.15) compared to the TAP conditions (M = 572.00, SE = 6.11).

1054.33, *SE* = 22.93) compared to Block 1 (*M* = 1110.47, *SE* = 27.23).

Table 27. μ Difference score comparisons with the control group.

	Mean Difference	*SE*	Dunnet *t* value	*p* value	Cohen's d (small = .2, medium = .5, large = .8)
Less than 3 μ Difference Score					
TAP/Specific vs. Control	58.39	60.58	0.96	0.74	0.22
TAP/General vs. Control	27.19	61.87	0.44	0.98	0.10
TIP/Specific vs. Control	49.99	62.34	0.80	0.84	0.20
TIP/General vs. Control	65.32	60.58	1.08	0.66	0.25
3 or more μ Difference Score					
TAP/Specific vs. Control	117.81	56.59	2.08	0.12	0.52
TAP/General vs. Control	113.77	57.79	1.97	0.16	0.43
TIP/Specific vs. Control	42.54	58.23	0.73	0.88	0.17
TIP/General vs. Control	187.01	56.59	3.30	**0.004**	0.82

Analyses for τ

I carried out a 2 (specificity) x 2 (appropriateness) x 2 (trial type) x 2 (block) mixed model ANOVA on the τ parameter to look at differences in the tail portions of the RT distributions. A significant main effect of trial type was present, $F(1,141) = 44.21$, $p < .001$, $\eta_p^2 = .24$, with smaller τ values for the "3 or more" trial type (*M* = 554.77, *SE* = 18.28) compared to the "less than 3" trial type (*M* = 648.15, *SE* = 22.04). There was also a main effect of block, $F(1,141) = 11.47$, $p = .001$, $\eta_p^2 = .08$, with lower τ values in Block 1 (*M* = 575.42, *SE* = 20.29) compared to Block 2 (*M* = 627.50, *SE* = 20.69). Additionally, there was a significant interaction between block and appropriateness, $F(1,141) = 6.74$, $p = .01$, $\eta_p^2 = .05$. The interaction is shown in Figure 26 below. Pairwise comparisons found no differences in the TIP conditions across blocks ($p = .58$), but a significant difference between blocks for the TAP conditions, $F(1,141) = 18.03$, $p < .001$, $\eta_p^2 = .11$.

Comparisons with the control group across both trial type and block were also conducted. The overall ANOVA tests were all non-significant (p's > .11) in comparing the experimental groups to the control group for both trial types and blocks.

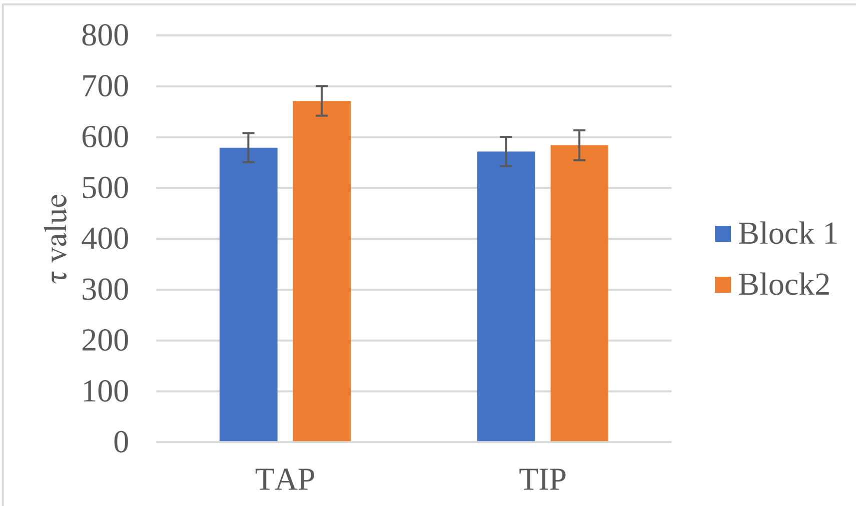

Figure 26. Interaction between trial type and appropriateness for τ.

Baseline performance for τ

The Block 1 baseline analyses compared the experimental groups to each other to insure all participants were beginning the task in the same manner prior to receiving their PM instructions. A 2 (specificity) x 2 (appropriateness) x 2 (trial type) mixed model ANOVA was carried out on the τ values for Block 1. A main effect of trial type was present, $F(1,141) = 31.27$, $p < .001$, $\eta_p^2 = .18$, reflecting the pattern found previously with smaller τ values for the "3 or more" trials ($M = 523.98$, $SE = 20.00$) compared to the "less than 3" trials ($M = 626.86$, $SE = 24.35$). All other main effects and interactions showed no differences among groups.

Difference Score for PM Block τ Analyses

To account for the main effect of trial type found in the baseline block, a difference score was calculated by subtracting the τ values from Block 1 from those of Block 2 and analyzing this difference. A 2 (specificity) x 2 (appropriateness) x 2 (trial type) mixed Model ANOVA was carried out on the τ difference scores. There was a main effect of appropriateness, $F(1,141) = 6.74$, $p = .01$, $\eta_p^2 = .05$, which showed a larger difference score (τ increase in Block 2) in the TAP conditions ($M = 91.98$, $SE = 21.66$) compared to the TIP conditions ($M = 12.18$, $SE = 21.83$). These values are comparable to the difference scores found in Experiment 1 (Experiment 1 TAP $M = 45.40$, Experiment 1 TIP $M = 97.83$), simply reversed for TAP and TIP. The difference scores for the experimental groups were not significantly different from the control group difference scores for the "less than 3" trial type. However, for the "3 or more" trial type, the overall ANOVA showed a significant difference between the experimental groups and the control group ($F(4,178) = 2.88$, $p = .02$) with only the TAP/Specific group reaching significance[6]. These outcomes can be found in Table 28.

Table 28. Comparisons with the control group for the difference score in τ.

	Mean Difference	*SE*	Dunnet *t* value	*p* value	Cohen's d (small = .2, medium = .5, large = .8)
Less than 3 Difference Score τ					
TAP/Specific vs. Control	18.93	59.20	0.32	0.45	0.08
TAP/General vs. Control	39.36	60.45	0.65	0.92	0.16
TIP/Specific vs. Control	-35.27	60.91	-0.58	0.91	-0.15
TIP/General vs. Control	-0.31	59.20	-0.01	1.00	-0.00
3 or more Difference Score τ					
TAP/Specific vs. Control	134.97	47.29	2.85	**0.02**	0.67
TAP/General vs. Control	41.82	48.29	0.87	0.80	0.19
TIP/Specific vs. Control	16.33	48.66	0.34	0.99	0.08
TIP/General vs. Control	-1.91	47.29	-0.04	1.00	-0.01

[6] These results were replicated using both trial types for Block 1 τ values as a covariate.

Group Analyses for τ

To further analyze group differences, I conducted a 2 (trial type) x 2 (block) x 5 (group) mixed model ANOVA for the τ parameter. A main effect of trial type was found, $F(1,178) = 64.49$, $p < .001$, $\eta_p^2 = .27$, which followed the consistent pattern of larger τ values for the "less than 3" trial type ($M = 645.55$, $SE = 19.25$) compared to the "3 or more" trial type ($M = 549.31$, $SE = 16.69$). There was also a main effect of block, $F(1,178) = 10.99$, $p = .001$, $\eta_p^2 = .06$, with larger τ values for Block 2 ($M = 620.01$, $SE = 18.10$) compared to Block 1 ($M = 574.85$, $SE = 18.51$). This main effect was qualified by an interaction between block and group, $F(4,178) = 2.67$, $p = .03$, $\eta_p^2 = .06$, which is displayed in Figure 27. Pairwise comparisons of the interaction

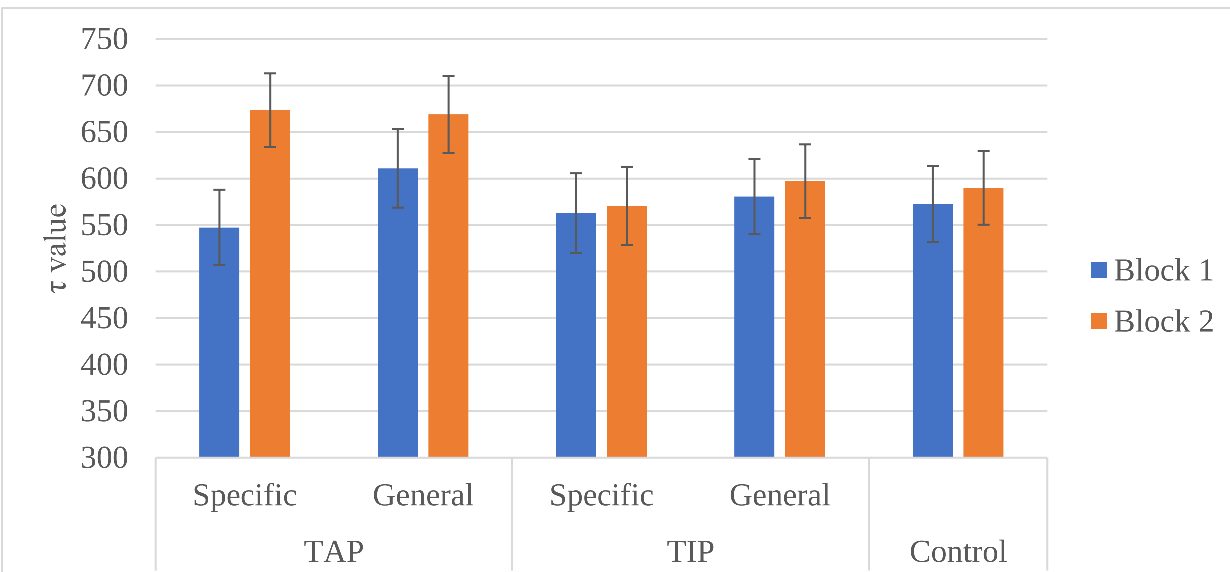

Figure 27. τ Interaction between block and group.

found a significant difference between blocks for the TAP/Specific group [$F(1,178) = 17.78$, $p < .001$, $\eta_p^2 = .09$]. The TAP/General group was approaching significance [$F(1,178) = 3.48$, $p = .06$, $\eta_p^2 = .02$], and the TIP conditions and control condition showed no differences in τ across blocks (all p's > .56).

Analyses for σ

A 2 (specificity) x 2 (appropriateness) x 2 (trial type) x 2(block) mixed model ANOVA was carried out on the σ parameter. There was a significant main effect of trial type, $F(1,141) = 43.41$, $p < .001$, $\eta_p^2 = .24$, with a smaller σ value for the "3 or more" trial type ($M = 206.30$, $SE = 8.42$) compared to the "less than 3" trial type ($M = 275.60$, $SE = 12.15$). No other main effects of interactions were found. Additionally, there were no differences found when comparing the experimental groups to the control group for either block or trial type.

Group Analyses for σ

In order to include the control group in the overall analysis, I carried out a 2 (trial type) x 2 (block) x 5 (group) mixed model ANOVA and found a main effect of trial type replicating the pattern found in the previous analyses. I also found a main effect of Block, $F(1,178) = 4.45$, $p = .04$, $\eta_p^2 = .24$, showing smaller σ values for Block 2 ($M = 230.06$, $SE = 8.63$) compared to Block 1 ($M = 248.47$, $SE = 9.58$).

Correlations

Correlations between PM Accuracy, Ongoing Task RTs, μ values, and τ values were also carried out for Experiment 2 in order to look at the functionality of costs and changes in the ongoing task on PM performance. These correlations helped to examine whether any relationships existed between PM accuracy and RT costs. Strength of the relationship is based on the guidelines from Evans (1996).

TAP/Specific Group

For the TAP/Specific group, no significant correlations were present for the TAP/Specific group, indicating any RT differences were not functional for the PM intention. The traditional mean RT measures were strongly related to measures of μ, and moderately related to τ measures. The μ and τ parameters show a trend for a negative relationship with each other. See Table 29.

Table 29. TAP/Specific Correlations

	Lax PM Acc.	Nonliving Mean RT D.S.	Living Mean RT D.S.	Nonliving μ D.S.	Living μ D.S.	Nonliving τ D.S.	Living τ D.S.
Relaxed PM Accuracy		.04 p = .82	.07 p = .67	.14 p = .41	.18 p = .27	-.12 p = .46	-.19 p = .48
Nonliving Mean RT D. S.			**.76 p < .001**	**.64 p < .001**	**.53 p = .001**	**.35 p = .03**	**.42 p = .01**
Living Mean RT D. S.				**.57 p < .001**	**.68 p < .001**	.17 p = .31	**0.59 p < .001**
Nonliving μ D. S.					**.54 p < .001**	**-.49 p = .002**	.18 p = .29
Living μ D. S.						-.06 p = .73	-.19 p = .25
Nonliving τ D. S.							.26 p = .11
Living τ D. S.							

Note. Difference score is abbreviated D. S. here. Correlation strength from Evans (1996): Very weak = .00-.19; Weak = .20-.39; Moderate = .40-.59; Strong = .60-.79; Very Strong = .80-1.0.

TAP/General Group

PM accuracy was not significantly related to any of the RT measures. The traditional RT measures had a moderate correlation with μ, and a weak relationship with τ.

The two Ex-Gaussian parameters showed a moderate negative relationship with each other. This relationship most likely suggests that when a large difference score is present for μ—indicating active monitoring/general slowing—reliance on spontaneous retrieval (τ) was low. Inversely, when the difference score for μ is lowest, or even negative—indicating practice effects and increased average speed in the PM block—reliance on spontaneous retrieval was increased. See Table 30.

Table 30. TAP/General Correlations

	Lax PM Acc.	Nonliving Mean RT D.S.	Living Mean RT D.S.	Nonliving μ D.S.	Living μ D.S.	Nonliving τ D.S.	Living τ D.S.
Relaxed PM Accuracy		.09 p = .60	.19 p = .27	-.06 p = .75	-.02 p = .91	.16 p = .36	.26 p = .14
Nonliving Mean RT D. S.			**.83 p < .001**	**.56 p < .001**	**.58 p < .001**	**.46 p = .005**	.33 p = .06
Living Mean RT D. S.				**.53 p = .001**	**.65 p < .001**	.32 p = .06	**0.45 p = .01**
Nonliving μ D. S.					**.55 p = .001**	**-.47 p = .005**	.03 p = .85
Living μ D. S.						.03 p = .87	**-.38 p = .03**
Nonliving τ D. S.							.31 p = .07
Living τ D. S.							

Notes: Difference score is abbreviated D. S. here. Correlation strength from Evans (1996): Very weak = .00-.19; Weak = .20-.39; Moderate = .40-.59; Strong = .60-.79; Very Strong = .80-1.0.

TIP/Specific Group

In this group, the PM accuracy was not significantly related to any of the RT metrics, suggesting no functional use of differences in RT after having an intention. There was a moderate relationship between traditional measures of RTs and μ values, which was not present with τ values. Similar to the TAP/General group, the μ and τ values showed a negative, moderate relationship with each other, most likely suggesting that when a large difference score was present for μ—indicating active monitoring/general slowing—reliance on spontaneous retrieval (τ) was low. Inversely, when the difference score for μ was lowest, or even negative—indicating practice effects and increased average speed in the PM block—reliance on spontaneous retrieval increased. See Table 31 for correlations.

Table 31. TIP/Specific Correlations

	Lax PM Acc.	Nonliving Mean RT D.S.	Living Mean RT D.S.	Nonliving μ D.S.	Living μ D.S.	Nonliving τ D.S.	Living τ D.S.
Relaxed PM Accuracy		-.12 p = .49	.18 p = .30	.01 p = .95	.23 p = .19	-.10 p = .58	-.10 p = .48
Nonliving Mean RT D. S.			**.50 p = .003**	**.48 p = .004**	**.46 p = .01**	**.44 p = .01**	-.05 p = .79
Living Mean RT D. S.				.22 p = .22	**.66 p < .001**	.24 p = .18	0.24 p = .17
Nonliving μ D. S.					.32 p = .06	**-.58 p < .001**	-.17 p = .33
Living μ D. S.						.09 p = .61	**-.57 p < .001**
Nonliving τ D. S.							.14 p = .45
Living τ D. S.							

Notes: Difference score is abbreviated D. S. here. Correlation strength from Evans (1996): Very weak = .00-.19; Weak = .20-.39; Moderate = .40-.59; Strong = .60-.79; Very Strong = .80-1.0.

TIP/General Group

All of the RTs measures showed no relationship to PM accuracy. A moderately strong correlation was present between the traditional RT measures and both μ and τ values. Only the nonliving μ and τ difference scores showed a strong, negative relationship with each other. See Table 32 for the correlations.

Table 32. TIP/General Correlations

	Lax PM Acc.	Nonliving Mean RT D.S.	Living Mean RT D.S.	Nonliving μ D.S.	Living μ D.S.	Nonliving τ D.S.	Living τ D.S.
Relaxed PM Accuracy		.11 p = .51	.003 p = .99	-.10 p = .54	.04 p = .81	.19 p = .26	-.04 p = .82
Nonliving Mean RT D. S.			**.54 p < .001**	.30 p = .07	**.42 p = .01**	**.58 p < .001**	.21 p = .20
Living Mean RT D. S.				.25 p = .13	**.71 p < .001**	.23 p = .16	**0.44 p = .01**
Nonliving μ D. S.					.31 p = .06	**-.61 p < .001**	-.06 p = .70
Living μ D. S.						.08 p = .64	-.31 p = .06
Nonliving τ D. S.							.23 p = .16
Living τ D. S.							

Notes: Difference score is abbreviated D. S. here. Correlation strength from Evans (1996): Very weak = .00-.19; Weak = .20-.39; Moderate = .40-.59; Strong = .60-.79; Very Strong = .80-1.0.

CHAPTER 10: EXPERIMENT 2 DISCUSSION

Summary of Experiment 2 Findings

I anticipated replicating the findings from Experiment 1, with a main effect of cue specificity and a main effect of task-appropriateness for PM accuracy. Interestingly, while the main effect of specificity was present as anticipated, the main effect for appropriateness was disrupted in this second experiment (see Figure 21). PM accuracy was especially elevated in the TIP/Specific condition (The focal condition from Experiment 1), suggesting the specific, semantic information offers a protective effect on PM performance when carried out in unrelated ongoing tasks. The benefit for specific cues was still evident in this ongoing task, strengthening the idea that consideration for cue specificity is critical when designing focal and nonfocal conditions.

For ongoing task accuracy, overall performance was expected to be slightly worse than that of Experiment 1 as the task is more difficult. Other than that, no differences were anticipated for ongoing accuracy across groups aside from a main effect of trial type, favoring the "less than 3" trial types as they would not be relevant to the specific PM intention and could largely be ignored (see Strickland et al. 2017, for task specific interference). The main effect of trial type was present but reversed from expectations in that better performance was found for the "less than 3" trials compared to the "3 or more" trials. This mirrors the trial type effect found in Experiment 1 in that accuracy was better on the trial type not related to the PM cues. A slight effect of block was also present, showing an increase in accuracy during the second block—possibly evidence for practice effects. Comparisons with the control group showed no differences between any of the groups, mapping onto the predicted null findings for ongoing task accuracy.

The main effect of trial type was again present in the traditional ongoing task RT analyses, with faster RTs in the “3 or more” trials. This finding matches that of Experiment 1, with faster decisions for the trial type that the PM cues belong to. This finding paired with the accuracy findings suggest functional slowing for the “less than 3” trials which had better accuracy but slower RTs. In analyzing the PM block, no main effect was present for appropriateness, again suggesting the cost effects were disrupted by benefits from the levels of processing principle. The TAP/Specific group actually showed significant differences from the control group in the “3 or more” trials, while all other groups were not significantly different for either trial type. This finding contradicts predictions set forth by task-appropriateness in which this group should be least likely to show differences from the control group due to its focal nature.

In looking at the RT difference scores for the experimental conditions compared to the control condition, only the “3 or more” trials showed significant differences. The TIP/Specific condition was not significantly different from the control group, where the other three comparison groups (the focal condition and two nonfocal conditions) all showed increases in RT. The focal condition in this second experiment was showing a cost to the ongoing task RT, while one nonfocal condition showed no such impact.

In the ex-gaussian analyses, the main effect of trial type was again present with smaller μ values for the “3 or more” trials, corroborating the traditional RT findings. There was also an overall decrease in μ from Block 1 to Block 2, indicative of practice effects on the task. An unexpected interaction was found in the baseline block between trial type and specificity (slightly resembling the μ baseline block findings for specificity). When accounting for these random discrepancies in Block 1 with a difference score, all main effects and interactions were non-significant, including comparisons to the control group aside from the TIP/General condition on the "3 or more" trials for μ. Overall, μ did not show any major differences among groups suggesting none of the groups strongly endorsed extreme active monitoring strategies in the context of this ongoing task.

The trial type main effect was present in the τ analyses, with smaller tail distributions in the "3 or more" trials, in line with the other trial type findings. A main effect of Block was found for τ values, but this time the Block 2 values were larger, indicative of more reliance on transient retrieval of the PM intention, especially in the TAP/General condition (see Figure 26). A main effect of appropriateness was found in the difference score analyses, but with the TIP conditions showing overall smaller differences in τ than the TAP conditions. This finding demonstrates the strength of impact that levels of processing has on PM intentions, as the semantic TIP conditions show patterns uncharacteristic of nonfocal intentions, and the TAP/Specific condition fails to align with traditional expectations for a focal PM intention. In fact, the focal condition in Experiment 2 was the only group to show a significant difference from the control group in the "3 or more" trials as well as the only group to show a significant change in τ from Block 1 to Block 2 (though the TAP/General condition was close to significance).

The relationships between PM accuracy and RT costs were explored in correlations split by group. The overall finding was that the RT costs found (which were not very prominent and were disrupted by the ongoing task switch) had no significant relationship with PM accuracy for any of the experimental groups. This finding is in contrast to the correlations found in Experiment 1, further supporting the importance of the ongoing task context surrounding the PM task in forming performance predictions.

Study Conclusions

One of the primary findings of this second experiment was the major disruption caused by switching the ongoing task's processing. Most of the ongoing tasks carried out in PM research are semantically driven, and these findings suggest that the benefits of semantic cues in semantic ongoing tasks is unrivaled even with a match in processing in an orthographic task. Careful consideration for the cue specificity, task-appropriateness, *and* processing needed for the ongoing task greatly impact PM performance. This study shows stability in cue specificity impacting PM cue detection, and predictable impacts of task-appropriateness on cue detection when also mindful of the processing required to carry out the intention.

The second experiment primarily helped to establish the role task-appropriateness plays in PM performance compared to level of processing. My predictions were in favor of task-appropriateness being the major factor impacting PM performance, in congruence with Meier and Graf's (2000) research and Abney et al.'s (2013) findings (also see Morris et al., 1977). My predictions did not pan out, however, and levels of processing fully disordered the predictions for ongoing task RT costs and PM accuracy. This second experiment helped to solidify the idea that focality hinges on two main components—cue specificity and task-appropriateness with careful consideration to the demands and processing required by the ongoing task. Both of these components are integral in establishing a PM intention as focal or nonfocal. Researchers need to be precise in how they define their PM intentions to ensure their task is actually focal, or nonfocal, and to what degree.

The findings of this second experiment also helped to determine what findings from the first experiment are fairly consistent across different ongoing tasks, and which dependent measures fluctuate with a different ongoing task. The ongoing task used in this second experiment was not extremely common in PM research, so researchers should conduct further work using more popular tasks such as lexical decision tasks and examine how that task interacts with PM cues along the focality continuum.

CHAPTER 11: GENERAL DISCUSSION

Collectively, the two proposed experiments clarify the definition of focality, a term that is pervasive in PM research, yet lacks solid grounding in what it really means for a task to be focal or nonfocal. These studies are some of the first to directly compare various nonfocal conditions defined by crossing cue specificity and task appropriateness. These comparisons help to establish the nature of focality and the differing pattern of findings when comparing PM accuracy to RT costs. Often in the literature, researchers describe PM tasks as either focal or nonfocal, which establishes the concept of focality as a dichotomy, though perhaps incorrectly so. The current studies demonstrated importance of both components of focality—cue specificity or task-appropriateness—and that the impact task-appropriateness had on PM task performance greatly depended on the ongoing task context. Pitting these two constructs against each other has rarely been carried out in prior research and has never been done as an insight into PM focality.

Together, these two studies speak to the role that levels of processing play in focality as semantic and orthographic PM and ongoing tasks are compared using the same stimuli. Orthographic processing is classified as shallow, while semantic processing is deep (Morris et al., 1977; Seamon & Virostek, 1978). The proposed studies teased apart task-appropriateness and levels of processing by having a semantic match between the PM task and the ongoing task in Experiment 1 and an orthographic match in Experiment 2. This contrast between the semantic and orthographic ongoing tasks helped to establish how much impact depth of processing has on PM performance. The complete wipeout of the task-appropriateness effects found in Experiment 1 by the switch in ongoing task for Experiment 2 demonstrate the care and attention needed to establish conditions with regards to their cue specificity and task-appropriateness in the context of an ongoing task especially for nonfocal tasks.

The designs for these two studies allows for some comparisons across experiments. The stimuli used is identical, allowing for possible differences in attention allocation to shine through in ongoing task costs under different ongoing tasks and different PM conditions. The subtle instructional differences between the conditions may offer insight into when participants adjust or establish the amount of attention they allocate towards completing the ongoing and PM tasks. In comparing the PM accuracy findings for both experiments (see Figures 7 and 22) PM detection was much higher in Experiment 2. The orthographic conditions showed a major improvement when they were task-appropriate. The semantic nature of Experiment 1 perhaps made the difficulty of these intentions harder for participants to assess. Future work should ask participants to rate their perceived difficulty of their intention to explore the possibility that this perception changed based on the ongoing task context.

Theoretical Support

The outcomes of these studies help to shape the PM field in how acutely researchers ought to define focality and how to establish focal and nonfocal conditions in future research. The results of these studies also offer grounds for comparing the theories mentioned previously (PAM, multiprocess framework, dynamic multiprocess framework, Delay Theory) and help to uncover when intention interference is greatest, when it is smallest, and whether cue specificity or task-appropriateness helps to reduce interference more.

The findings from the first experiment primarily align with the expectations set forth by the multiprocess framework and the dynamic multiprocess framework. There were no costs to the ongoing task in the TAP/Specific, focal condition while all three of the nonfocal conditions did show an RT cost compared to the control group. The PAM theory and Delay Theory would have anticipated finding a cost in the focal condition as well as in the nonfocal conditions, but this was not the case. Distinguishing between the multiprocess framework and the dynamic multiprocess framework requires further analyses that break the PM block into pre-PM cue and post-PM cue trials to see if participants are actively switching between monitoring and spontaneous retrieval strategies. Breaking the PM block into subsections of the ongoing task will help to uncover possible differences in ex-gaussian parameters after detecting the initial PM cue. Additional analyses on these two experiments will continue with more detailed examination of the strategies reported in the post-experiment questions to see if participants have insight into which strategy they are employing.

Experiment 2 painted a very different picture from Experiment 1, however. The *only* group that experienced ongoing task costs was the focal group, while all the nonfocal conditions were comparable to the control group. These exact cost findings were not expected by any of the current PM theories, which normally all anticipate that the focal group shows the least amount of costs. The PAM theory would have predicted costs for all experimental groups, yet here none of the nonfocal groups experienced a cost. The multiprocess framework might suggest post-hoc that the focal group was utilizing a monitoring strategy, while all of the nonfocal groups were utilizing spontaneous retrieval, but this pattern of outcomes does not jive with the standard predictions established by the multiprocess framework, or the dynamic multiprocess framework. Further analyses of the self-reported strategy use could also shed light on this particular pattern and reveal if participants actually report spontaneous retrieval-type strategies more often in those nonfocal conditions compared to the focal group. The pattern of cost patterns may yet be explained by the dynamic multiprocess framework, again by dividing the PM block into subsections. Participants in the focal condition might have employed a heavy monitoring strategy initially, and switched to a spontaneous retrieval mindset after garnering experience with the PM cues. Its possible that extreme switches in strategy might cloud the cost outcomes, even in the ex-gaussian parameters, as all of the PM block trials were analyzed as one distribution.

The Delay Theory predicted costs in the trial type that most closely matched that of the PM cues—the living trials and the “3 or more” trials—but the actual results showed conflicting ongoing task performance costs with faster RTs for those trial types across both experiments but worse ongoing task accuracy for the living trials as well as the “3 or more” trials. Possible reasoning for these costs are explored next.

Difficulty of the Ongoing Task

It is possible that much of the ongoing task cost outcomes in the second experiment were muddled by the overall difficulty of the ongoing task, especially compared to the ease of the animacy judgement used in Experiment 1. Animacy is something people learn and use regularly everyday (e.g. is that a stick or a snake) while determining the number of enclosed spaces in a sequence of letters is rarely, if ever used by the average person. While it is a simple task on the surface, more practice with the task may show that indeed we were seeing practice effects in Experiment 2, and that there are costs with the addition of a PM intention (possibly that even align with standard predictions from current theories).

The difficulty with the ongoing task in Experiment 1 likely stemmed from the particular stimuli used, especially for the living wordlists. Traditionally, when thinking about animacy, the normal categories of living stimuli are plants and animals. As the general PM cue used any animal words, the ongoing task had to be comprised of living things that were not animals, making these living things more abstract than our standard animacy classification.

The difficulty of the ongoing tasks differed across experiments, and likely differed in participants assumptions about the difficulty of their tasks. Participants likely initially thought their animacy decision would be easy, and gradually learned in the baseline block it was more challenging than anticipated. Conversely, participants in the second experiment likely thought their task was foreign, confusing, and difficult. With practice they probably realized it was not as difficult as anticipated. The rate of these changes in perceived task difficulty might have been different across experiments and is something that should be explored more extensively as it may have impacted the findings of the current work, especially when considered simultaneously with the perceived difficulty of the PM intention. In the second experiment, participants in the focal condition were asked to make a decision about the number of enclosed spaces in a word, and then also press Y if they saw any double letter o's throughout the experiment. This task *sounds* objectively more difficult than pressing Y when seeing the word kangaroo or rooster. The perceived difficulty may align with levels of processing but may also interact with the perceived difficulty of the ongoing task. To test this notion, other ongoing tasks—both semantic and orthographic in nature—should be used in future PM experiments looking at cue specificity and task-appropriateness.

Practical Implications

The results of the current studies help to inform what behavioral patterns to anticipate under various focality levels. These expectations may be especially useful to neurological PM researchers, as a more nuanced approach to focality may help explain disparate activation patterns. ERP findings show greater levels of activation in support of increased monitoring behavior during nonfocal tasks compared to focal tasks (Cona et al., 2013). In an fMRI study, Gordon et al. (2011) had participants complete a category matching task for the ongoing task and the PM intention was to indicate when they saw either a specific word (focal condition) or when they saw a specific syllable (nonfocal condition). Gordon et al. (2011) found a positive relationship between PM accuracy and activation in the medial temporal region, strongest in the hippocampus. They found no structural relationships related to nonfocal PM performance. The researchers took these findings as evidence supporting a spontaneous, more automatic retrieval during focal tasks (Gordon et al., 2011). McDaniel et al. (2013) found that both focal and nonfocal tasks led to activation in the Anterior Prefrontal Cortex (aPFC), though this activation was strongly connected to the precuneus on nonfocal trials or with the right middle temporal gyrus on focal trials. Later, McDaniel and colleagues (2015) found that the medial PFC was deactivated during both focal and nonfocal PM trials. They also found that demanding focal PM tasks and nonfocal tasks both activated the insula, which they took to indicate this brain region is related to actively monitoring for the PM cue (McDaniel et al., 2015). Additionally, McDaniel et al. (2015) found activation in Brodmann's Area 9 during a non-demanding focal PM task, which led the researchers to conclude this area may be involved in involuntary episodic retrieval and transient activation of the focal PM task. In Cona et al.'s (2016) meta-analysis of PM focality in neurological studies, they found an overall activation in the left aPFC during nonfocal tasks, and activation in the cerebellum and ventral parietal regions during focal tasks. This led Cona et al. (2016) to conclude that bottom-up automatic processing occurs during focal tasks and top-down processes activate during nonfocal tasks. The main outcome to remember from these neurological findings is that brain activation patterns are different depending on the focality of the PM task, but that the definition of the PM task as focal or nonfocal was based on cue specificity in some experimental designs (Cona et al., 2013), and task-appropriateness in others (Gordon et al., 2011; McDaniel et al., 2013), and both in various meta-analyses and reviews (Cona et al., 2016; McDaniel et al., 2015).

Other factors in PM research—cue salience, task importance, planning capabilities, strategies to improve PM, other populations—are still of interest, especially in how they interact with a nuanced focality. Marsh et al. (2000) found that salient cues essentially override task-appropriateness in capturing attention and leading to very high PM performance when participants receive a salient cue. Whether salience impacts all four focality conditions in the same manner is yet to be explored. In a similar fashion, how much does task importance change performance or cost in a focal setting compared to the three different nonfocal settings? Are planning techniques and memory strategies most useful only under certain focality parameters? Do certain populations that struggle with PM tasks primarily have issues with one type of nonfocal task? All of these questions remain unanswered, as focality has previously only been comprised of "focal" or "nonfocal". The nuanced approach to focality established with these two experiments could offer further insight into discrepant findings in the literature examining these other variables (e.g. why some aging studies find age-related differences on a nonfocal task and others find no differences) and help to establish the idea that not all nonfocal tasks are equally nonfocal.

Collectively, all of these factors—cue salience, task importance, planning strategies, task difficulty—should be considered when mapping out theoretical predictions for various PM intentions. Evidence from the current studies suggest predictions would differ based on whether the ongoing task is a semantic task or not, as well as the difficulty of the task. These studies highlight the need for a more thorough theory to describe PM performance in a wider variety of tasks and conditions.

In conclusion, these two experiments helped to define the construct of focality in PM research, and push the field towards a more nuanced approach in selecting PM cues and ongoing tasks. These experiments are some of the first to directly compare cue specificity and task appropriateness, two major principles in PM research, and how they impact PM performance, ongoing task accuracy, and ongoing task speed. The current experiments compared three distinct nonfocal conditions, a novel comparison that has previously not intentionally been carried out. These studies also expanded the work on RT distribution analyses in PM research, which broadens the scope and ability to compare ongoing task costs. Finally, the ability to use the exact same word lists across two different ongoing tasks and four different PM cue conditions helps to control extraneous variations often present across experiments and conditions and allows for clearer comparisons across conditions.

REFERENCES

Abney, D. H., McBride, D. M., & Petrella, S. N. (2013). Interactive effects in transfer-appropriate processing for event-based prospective memory: The roles of effort, ongoing task, and PM cue properties. *Memory & Cognition, 41*, 1032-1045. Doi: 10.3758/s13421-013-0324-7

Anderson, N. D. & Craik, F. I. M. (2000). Memory in the aging brain. In Tulving, E. & Craik, F. I. M. (Eds.) *The oxford handbook of memory* (411-425). New York, NY: Oxford University Press.

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*(2), 492-499. https://doi.org/10.3758/s13423-014-0700-8

Brewer, G. A. (2011). Analyzing response time distributions: Methodological and theoretical suggestions for prospective memory researchers. *Journal of Psychology, 219*, 117-124. Doi: 10.1027/2151-2604/a000056

Burgess, P. W., Gonen-Yaacovi, G., & Volle, E. (2011). Functional neuroimaging studies of prospective memory: What have we learnt so far? *Neuropsychologia*, *49*, 2246-2257. https://doi.org/10.1016/j.neuropsychologia.2011.02.014

Cohen, A. L. & Hicks, J. L. (2017). *Prospective Memory: Remembering to Remember, Remembering to Forget.* Springer Nature. doi:10.1007/978-3-319-68990-6

Cona, G., Bisiacchi, P. S., & Moscovitch, M. (2013). The effects of focal and nonfocal cues on the neural correlates of prospective memory: Insights from ERPs. *Cerebral Cortex, 24*, 2630-2646. https://doi.org/10.1093/cercor/bht116

Cona, G., Bisiacchi, P. S., Sartori, G., & Scarpazza, C. (2016). Effects of cue focality on the neural mechanisms of prospective memory: A meta-analysis of neuroimaging studies. *Scientific Reports*, *6*, 1-13. Doi: 10.1038/srep25983

Craik, F.I.M. and Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671-684. https://doi.org/10.1016/S0022-5371(72)80001-X

Craik, F. I. M. & Rose, N. S. (2012). Memory encoding and aging: A neurocognitive perspective. *Neuroscience and Biobehavioral Reviews, 36*, 1729-1739. https://doi.org/10.1016/j.neubiorev.2011.11.007

Craik, F. I. M. & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*(3), 268-294.

Einstein, G. O. & McDaniel, M. A. (1990). Normal aging and prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 717-726. http://dx.doi.org/10.1037/0278-7393.16.4.717

Einstein, G. O., & McDaniel, M. A. (1996). Retrieval processes in prospective memory: Theoretical approaches and some new empirical findings. In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.) *Prospective memory: Theory and applications* (115-141). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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

Einstein, G. O., McDaniel, M. A., Richardson, S. L., Guynn, M. J., & Cunfer, A. R. (1995). Aging and prospective memory: Examining the influences of self-initiated retrieval processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 996-1007. http://dx.doi.org/10.1037/0278-7393.21.4.996

Einstein, G. O., McDaniel, M. A. & Scullin, M. K. (2012). Prospective memory and aging: Understanding the variability. In M. Naveh-Benjamin & N. Ohta (Eds.) *Memory and aging: Current issues and future directions* (153-179). New York, NY: Psychology Press Taylor & Francis Group.

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

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

Ellis, J. & Milne, A. (1996). Retrieval cue specificity and the realization of delayed intentions. *The Quarterly Journal of Experimental Psychology, 49A*, 862-887. http://dx.doi.org/10.1080/713755662

Faul, F., E. Erdfelder, A.G. Lang, and A. Buchner (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 3*9*, 175-191. https://doi.org/10.3758/BF03193146

Godden, D. R. & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology, 66*, 325-331. Doi: 10.1111/j.2044-8295.1975.tb01468.x

Gollwitzer, P. M. (1999). Implementation intentions: strong effects of simple plans. *American psychologist*, *54*, 493-503. http://dx.doi.org/10.1037/0003-066X.54.7.493

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

Graf, P., Uttl, B., & Dixon, R. (2002). Prospective and retrospective memory in adulthood. In P. Graf & N. Ohta (Eds.) *Lifespan Development of Human Memory* (257-282). Cambridge, MA, US: The MIT Press.

Guynn, M. J. (2010). A two-process model of strategic monitoring in event-based prospective memory: Activation/retrieval mode and checking. *International Journal of Psychology, 38*, 245-256. http://dx.doi.org/10.1080/00207590344000178

Heathcote, A., Brown, S. & Cousineau, D. (2004). QMPE: Estimating lognormal, Wald, and Weibull RT distributions with a parameter dependent lower bound. *Behavior research methods, instruments, & computers: A journal of the Psychonomic Society, Inc*, *36*, 277-290. https://doi.org/10.3758/BF03195574

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

Heathcote, A., Popiel, S. J., & Mewhort, D. J. K. (1991). Analysis of response time distributions: An example using the stroop task. *Quantitative Methods in Psychology, 109*, 340-347. http://dx.doi.org/10.1037/0033-2909.109.2.340

Henry, J. D., MacLeod, M. S., Phillips, L. H., & Crawford, J. R. (2004). A meta-analytic review of prospective memory and aging. *Psychology and Aging*, *19*(1), 27-39. http://dx.doi.org/10.1037/0882-7974.19.1.27

Hering, A., Rendell, P. G., Rose, N. S., Schnitzspahn, K. M., & Kliegel, M. (2014). Prospective memory training in older adults and its relevance for successful aging. *Psychological Research*, *78*, 892-904. https://doi.org/10.1007/s00426-014-0566-4

Hicks, J. L., Cook, G. I., & Marsh, R. L. (2005). Detecting event-based prospective memory cues occurring within and outside the focus of attention. *The American Journal of Psychology, 118*(1), 1-11. http://www.jstor.org/stable/30039040

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

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

Ihle, A., Ghisletta, P., & Kliegel, M. (2016). Prospective memory and intraindividual variability in ongoing task response times in an adult lifespan sample: The role of cue focality. *Memory, 25*, 370-376. http://dx.doi.org/10.1080/09658211.2016.1173705

Jäger, T. & Kliegel, M. (2008). Time-based and event-based prospective memory across adulthood: Underlying mechanisms and differential costs on the ongoing task. *The Journal of General Psychology, 135*(1), 4-22. http://dx.doi.org/10.3200/GENP.135.1.4-22

Kliegel, M., Altgassen, M., Hering, A., Rose, N. S. (2011). A process-model based approach to prospective memory impairment in Parkinson's disease. *Neuropsychologia, 49*, 2166-2177. https://doi.org/10.1016/j.neuropsychologia.2011.01.024

Kliegel, M., Jäger, T., & Phillips, L. H. (2008). Adult age differences in event-based prospective memory: A meta-analysis on the role of focal versus nonfocal cues. *Psychology and Aging, 23*, 203-208. http://dx.doi.org/10.1037/0882-7974.23.1.203

Kvavilashvili, L., & Ellis, J. (1996). Varieties of intention: Some distinctions and classifications. In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications* (23-51). Mahwah, NJ: Lawrence Erlbaum Associates.

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

Loft, S. & Remington, R. W. (2013). Wait a second: Brief delays in responding reduce focality effects in event-based prospective memory. *The Quarterly Journal of Experimental Psychology*, *66*, 1432-1447. Doi: 10.1037/a0033702

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. http://dx.doi.org/10.1037/a0033702

Marsh, R. L., Hicks, J. L., & Cook, G. I. (2005). On the relationship between effort toward an ongoing task and cue detection in event-based prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(1), 68-75. http://dx.doi.org/10.1037/0278-7393.31.1.68

Marsh, R. L., Hicks, J. L., & Cook, G. I. (2006). Task interference from prospective memories covaries with contextual associations of fulfilling them. *Memory & Cognition*, *34*, 1037-1045. https://doi.org/10.3758/BF03193250

Marsh, R. L., Hicks, J. L., & Cook, G. I. (2008). On beginning to understand the role of context in prospective memory. In M. Kliegel, M. A. McDaniel, & G. O. Einstein (Eds.), *Prospective Memory: Cognitive, neuroscience, developmental, and applied perspectives* (77-100). New York, NY: Taylor & Francis Group.

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

Marsh, R. L., Hicks, J. L., & Hancock, T. W. (2000). On the interaction of ongoing cognitive activity and the nature of an event-based intention. *Applied Cognitive Psychology*, *14*, 29-41. Doi: 10.1002/acp.769

Mayes, A. R. & Roberts, N. (2002). Theories of episodic memory. In Baddeley, A., Conway, M., & Aggleton, J. *Episodic memory: New directions in research* (86-109). New York, NY: Oxford University Press.

Maylor, E. A. (1996). Age-related impairment in an event-based prospective-memory task. *Psychology and Aging*, *11*(1), 74-78. http://dx.doi.org/10.1037/0882-7974.11.1.74

McBride, D. M. & Abney, D. H. (2012). A comparison of transfer-appropriate processing and multi-process frameworks for prospective memory performance. *Experimental Psychology*, *59*, 190-198. https://doi.org/10.1027/1618-3169/a000143

McDaniel M. A. & Einstein, G. O. (1993). The importance of cue familiarity and cue distinctiveness in prospective memory. *Memory*, *1*(1), 23-41. http://dx.doi.org/10.1080/09658219308258223

McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. *Applied Cognitive Psychology*, *14*, S127-S144. Doi: 10.1002/acp.775

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

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

McDaniel, M. A., Robinson-Riegler, B., & Einstein, G. O. (1998). Prospective remembering: Perceptually driven or conceptually driven processes? *Memory & Cognition*, *26*, 121-134. https://doi.org/10.3758/BF03211375

McDaniel, M. A., Umanath, S., Einstein, G. O., & Waldum, E. R. (2015). Dual pathways to prospective remembering. *Frontiers in Human Neuroscience, 9*, 392-404. doi: 10.3389/fnhum.2015.00392

McGann, D., Ellis, J. A., & Milne, A. (2002). Conceptual and perceptual processes in prospective remembering: Differential influence of attentional resources. *Memory & Cognition, 3*, 1021-1032. https://doi.org/10.3758/BF03194320

Meeks, J. T. & Marsh, R. L. (2010). Implementation intentions about nonfocal event-based prospective memory tasks. *Psychological Research, 74*, 82-89. https://doi.org/10.1007/s00426-008-0223-x

Meier, B. & Graf, P. (2000). Transfer appropriate processing for prospective memory tests. *Applied Cognitive Psychology, 14*, 11-27. Doi: 10.1002/acp.768

Meiser, T. & Schult, J. C. (2008). On the automatic nature of the task-appropriate processing effect in event-based prospective memory. *European Journal of Cognitive Psychology, 20*, 290-311. http://dx.doi.org/10.1080/09541440701319068

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16*, 519-533. https://doi.org/10.1016/S0022-5371(77)80016-9

Mullet, H. G., Scullin, M. K., Hess, T. J., Scullin, R. B., Arnold, K. M., & Einstein, G. O. (2013). Prospective memory and aging: Evidence for preserved spontaneous retrieval with exact but not related cues. *Psychology and Aging, 28*, 910-922. http://dx.doi.org/10.1037/a0034347

Psychology Software Tools, Inc. [E-Prime 2.0]. (2012). Retrieved from http://www.pstnet.com.

Reese-Melancon, C. (2013). Age, focal processing, and monitoring in event-based prospective memory. *Journal of Adult Development, 20*, 151-157. doi: 10.1007/s10804-103-9166-6.

Rummel, J., Smeekens, B. A., & Kane, M. J. (2016). Dealing with prospective memory demands while performing an ongoing task: Shared processing, increased on-task focus, or both? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*, 1047-1062. Doi: 10.1037/xml0000359

Scullin, M. K., McDaniel, M. A., & Shelton, J. T. (2013). The Dynamic Multiprocess Framework: Evidence from prospective memory with contextual variability. *Cognitive Psychology*, *67*(1), 55-71. https://doi.org/10.1016/j.cogpsych.2013.07.001

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

Seamon, J. G., & Virostek, S. (1978). Memory performance and subject-defined depth of processing. *Memory & Cognition, 6*, 283-287. https://doi.org/10.3758/BF03197457

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

Smith, R. E. & Bayen, U. J. (2006). The source of adult age differences in event-based prospective memory: A multinomial modeling approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 623-635. ttp://dx.doi.org/10.1037/0278-7393.32.3.623

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

Smith, R. E., Hunt, R. R., & Murray, A. E. (2016). Prospective memory in context: Moving through a familiar space. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* Advance online publication. http://dx.doi.org/10.1037/xlm0000303

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*, doi: 10.1037/xlm0000400

Underwood, A. G., Guynn, M. J., & Cohen, A. L. (2015). The future orientation of past memory: The role of BA 10 in prospective and retrospective retrieval modes. *Frontiers in Human Neuroscience, 9,* 1-12. Doi: 10.3389/fnhum.2015.00668

Uttl, B. (2011). Transparent meta-analysis: Does aging spare prospective memory with focal vs. non-focal cues? *PLoS ONE, 6*(2), 1-19. https://doi.org/10.1371/journal.pone.0016618

West, R. (2011). The temporal dynamics of prospective memory: A review of the ERP and prospective memory literature. *Neuropsychologia, 49*, 2233-2245. https://doi.org/10.1016/j.neuropsychologia.2010.12.028

West. R. & Craik, F. I. M. (1999). Age-related decline in prospective memory: The roles of cue accessibility and cue sensitivity. *Psychology and Aging, 14*, 264-272. http://dx.doi.org/10.1037/0882-7974.14.2.264

West, R. & Craik, F. I. M. (2001). Influences on the efficiency of prospective memory in younger and older adults. *Psychology and Aging, 16*, 682-696. http://dx.doi.org/10.1037/0882-7974.16.4.682

Wilson, M. D. (1988). The MRC psycholinguistic database: Machine readable dictionary, Version 2. *Behavioural Research Methods, Instruments and Computers, 20*, 6-11.

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

APPENDIX A: STIMULI WORD LISTS

Living/ 3 or more spaces			Statistics	Living/ less than 3 spaces			Statistics
acrobat	doctor	nomad	Average Concreteness: 514.15	adult	knight	pupil	Average Concreteness: 521.71
admiral	donor	observer		actor	lady	scholar	
amateur	dreamer	outsider		artist	lawyer	scout	
athlete	editor	painter		aunt	lecturer	sculptor	
babe	educator	parent	Average familiarity: 481.88	author	liar	servant	Average familiarity: 506.09
baby	emperor	pastor		butcher	lily	sister	
bachelor	female	patient		chief	lunatic	spruce	
bacteria	fielder	patriot		child	maid	student	
baker	general	peasant	Average imaginability: 514.29	citizen	maker	tailor	Average imaginability: 527.53
bandit	genius	people		clown	mankind	thief	
banker	geologist	person		cousin	master	thinker	
barber	graduate	poet		crew	mechanic	tourist	
baron	grocer	police	Average number of letters: 6.47	daisy	merchant	traitor	Average number of letters: 5.63
bishop	guard	pope		duke	miner	tulip	
brother	guest	prisoner		earl	minister	tyrant	
builder	guide	prophet		elm	minor	typist	
burglar	herdsman	reaper	Average number of syllables: 2.20	enemy	mister	umpire	Average number of syllables: 1.87
busybody	hostage	rebel		father	monarch	uncle	
captain	husband	retailer		florist	mother	usher	
champion	invader	seaman		foe	musician	victim	
colonel	islander	sender		guy	orchid	vine	
composer	judge	sergeant		heir	owner	visitor	
comrade	leader	singer		hero	outlaw	voter	
creator	madman	soldier		hunter	pianist	waiter	
dad	magician	stranger		imitator	pine	widow	
dame	maiden	surgeon		janitor	plant	wife	
dancer	member	sycamore		jockey	prey	woman	
dean	neighbor	teacher		junior	priest	worker	
deputy	nephew	weaver		jury	prince	youth	

Nonliving/ 3 or more spaces			Statistics	Nonliving/ less than 3 spaces			Statistics
aerial	entrance	peso	Average Concreteness: 530.36	article	helmet	rhyme	Average Concreteness: 522.20
basement	estate	piano		aspirin	history	salary	
basket	evening	pocket		autumn	hurdle	salute	
board	exterior	poetry		basin	icicle	sash	
boat	fudge	poison	Average familiarity: 487.34	blister	inferno	satchel	Average familiarity: 487.11
boulder	garment	polo		bristle	inquiry	satin	
boundary	globe	position		brush	invoice	sauce	
brake	glove	profile		bucket	jewel	station	
bread	graphite	quarter	Average imaginability: 527.63	buckle	journal	sunset	Average imaginability: 523.84
bridge	grate	racquet		bump	level	surface	
bungalow	grave	railroad		calculus	library	surprise	
burlap	hardware	report		canal	linen	tidbit	
cage	heap	residue	Average number of letters: 6.33	carat	locker	token	Average number of letters: 6.36
camera	hexagon	revolver		casket	lotion	traction	
candle	jade	shortage		cavern	machine	tractor	
capsule	kerosene	slope		cement	manicure	trail	
carpet	leather	sonata	Average number of syllables: 2.03	charter	manor	trailer	Average number of syllables: 2.18
ceremony	lemonade	spade		chisel	mansion	uniform	
concrete	magnet	spatula		circuit	mantle	utensil	
croquet	mortgage	squeak		citation	mercury	velocity	
damage	mosque	stable		clarinet	metal	velvet	
dimple	oatmeal	tape		cocktail	moisture	vestment	
disaster	opal	toaster		costume	movie	violin	
drove	outpost	tomahawk		emulsion	pants	volume	
dungeon	overcoat	triangle		eternity	picture	whisper	
easel	package	tripod		exhaust	primary	whistle	
edition	paper	trombone		fabric	rack	window	
empire	parade	vodka		factory	racket	winter	
enamel	paradise	wage		frame	recital	zenith	

APPENDIX B: FULL EXPERIMENT 2 METHODOLOGY

Experiment 2 Methodology is covered below in full.

Experiment 2 Method

In the second experiment I manipulated the cue specificity (specific or general) and the task-appropriateness (TAP or TIP) of a PM cue during an orthographic ongoing task in order to get the full picture of focality and how it relates to PM accuracy and ongoing task costs. I obtained Institutional Review Board approval through Louisiana State University prior to data collection (see Appendix C).

Participants

Identical to Experiment 1, participants recruited for this study were Louisiana State University undergraduate students taking psychology courses. They received course credit through the Sona Systems experiment portal for their participation in this study. All participants provided informed consent before participating in the study. Again, identical to the first experiment, estimates for the sample size needed to carry out this study came from an interaction found by Abney et al. (2013), who compared a semantic PM task to an orthographic PM task while completing a semantic or orthographic ongoing task in a similar fashion to the current experiments. I conducted power analyses using G*Power software (Faul, Erdfelder, Lang, & Buchner, 2007) based on effect sizes reported by Abney et al. (2013). Abney et al. (2013) found a significant main effect of task-appropriateness, with a medium effect size of $\eta_p^2 = .08$. Using G*Power, I converted this value to $f = 0.295$ for use in the sample size estimates. Using their effect size and setting the other parameters to appropriate levels ($\alpha = 0.05$, power = 0.95, the numerator df =1, and the number of groups to 4), the a priori power calculation for an ANOVA: Fixed effects, special, main effects and interactions using G*Power lead to a sample size computation equal to 152 total participants. This is equal to 38 people per group; and with the inclusion of the control group, the grand total of participants collected for Experiment 2 was estimated at 190 people. For conservative estimates, a goal of N=200 allowed for possible errors in collection with 40 people per group. Additional power analyses were conducted based on Abney et al.'s (2013) other results, including their RT ex-Gaussian analyses. However, each subsequent power analysis resulted in smaller sample size estimates, so I retained the most conservative estimate for the proposed study.

Two hundred six participants were tested in Experiment 2. Of these participants, two experienced computer glitches that prevented the full recording of their data and were excluded from all analyses.

Materials

The current experiment was programmed and conducted using E-Prime 2.0 software (Psychology Software Tools, 2012), which controlled stimulus presentation and recorded participant responses and RTs. Words used for the ongoing task were generated by the MRC Psycholinguistic Database (Wilson, 1988) and constrained to include only nouns with 3-9 letters. Words were excluded if they were an animal or a word with double letters (e.g. bull) to avoid creating additional PM cues in certain PM conditions, resulting in 348 total words (these can be found in Appendix A). These words were used for both Experiment 1 and 2, so balancing the words based on their classifications for the ongoing task was needed. The lists were balanced by animacy and the number of enclosed spaces so that exactly half of the words used were living, and the other half were nonliving (relevant for Experiment 1); additionally, half of each living and nonliving category had 3 or more enclosed spaces, and half had less than 3 enclosed spaces (relevant for the ongoing task used in Experiment 2). Each of the word lists used in the ongoing task (living words/3 or more spaces; living words/ less than 3 spaces; nonliving words/3 or more spaces; nonliving words/ less than 3 spaces) were constructed to be comparable based on measures of concreteness (*M* = 522.20 *SD* = 6.63), familiarity (*M* = 490.61, *SD* = 10.63), imaginability (*M* = 523.32, *SD* = 6.28), the number of letters (*M* = 6.20, *SD* = 0.38), and the number of syllables (*M* = 2.07, *SD* = 0.15) calculated from the MRC Psycholinguistic Database (Wilson, 1988). All words (n = 432) were normed to get a frequency of agreement in classifying the words as living or non-living. Approximately 63 individuals took part in the norming process. Words had to meet a classification ratio of .65, meaning at least 65% of the participants agree in classifying the concept as living or nonliving. The final word lists ended with an average agreement on classification above 90% for the two living word lists, and above 97% agreement for the two nonliving lists, with each list trimmed to n = 87.

Participants saw two PM cues selected from six possible PM cues counterbalanced across participants. The PM cues were the same for all groups, only differing by the instructions they read, and each cue was presented twice. Use of the same cues across groups was necessary to avoid any differences in cue properties that could have impacted PM performance (Marsh et al., 2000; West & Craik, 2001). The PM cues used for all groups were: baboon, goose, moose, raccoon, rooster, and kangaroo. Ex-Gaussian analyses were conducted using the Quantile Maximum Probability Estimator software (QMPE; Heathcote, Brown, & Cousineau, 2004).

Procedure

Participants were tested individually. Each participant was randomly assigned to an experiment and condition between Experiments 1 and 2. Simultaneous collection of both experiments protected against any possible differences introduced by participants that complete experiments early or late in the semester and allows for cleaner comparisons across experiments. After providing their informed consent, participants read the instructions about the task.

The ongoing task required participants to make an orthographic judgment about how many enclosed spaces are in the words shown to them on the screen. They were asked to carry out the task as quickly and accurately as they can. Participants then completed a short block of practice trials (20 trials total) of the ongoing task and asked questions if needed. After the practice block of trials, participants carried out a baseline block of only the ongoing task (172 total trials). After completing the baseline block, participants read the directions for completing the PM task (excluding the control condition). The PM instructions provided were manipulated across participants as a between subject manipulation. The TAP and TIP conditions in this second experiment were reversed from Experiment 1 in order to match the orthographic processing used in the ongoing task. Now, the TAP conditions required participants to either press “y” when they see two consecutive double letter o’s (specific condition), or to press the “y” key when they see any two consecutive letters, or doubles (general condition). The TIP/Specific condition required participants to press “y” when they see either of the following words: baboon or goose (other participants received two other specific cues from moose, rooster, raccoon, kangaroo used as a counterbalancing measure). The TIP/General condition asked participants to press “y” when they see any animal word, excluding humans. After the participants indicated that they understood both the PM instructions and the ongoing task instructions they filled out demographic information (age and sex) that served as a distraction/ delay task prior to starting the PM trials. Once participants began the ongoing task (180 trials total; 4 of which were PM trials), the 2 PM cues appeared embedded evenly spaced across the second half of the ongoing task trials (roughly every 28 trials) to allow for a longer delay between encoding the intention and the first PM cue. Once participants finished with the task, they were asked six questions about possible monitoring, as a manipulation check to insure they understood the PM instructions. The questions used can be found in Table 7. Participants were then debriefed and given credit through Sona Systems for their participation.

Design

This experiment manipulated two independent variables (cue specificity and task-appropriateness) with two levels each (specific and general, TAP and TIP, respectively), all between-subjects, and compared these groups to a control condition with no PM intention. Table 5 provides a schematic of the conditions included in this experiment, just reversed for TAP and TIP demarcations. Additional variables of interest built into the design as within-subjects manipulations included the trial type ("less than 3" and "3 or more") and block (baseline block and PM block), which became import when examining ongoing task accuracy, and ongoing task RTs.

APPENDIX C: IRB APPROVAL

ACTION ON EXEMPTION APPROVAL REQUEST

Institutional Review Board
Dr. Dennis Landin, Chair
130 David Boyd Hall
Baton Rouge, LA 70803
P: 225.578.8692
F: 225.578.5983
irb@lsu.edu
lsu.edu/research

TO: Bethany Lyon
Psychology

FROM: Dennis Landin
Chair, Institutional Review Board

DATE: September 6, 2017

RE: **IRB#** E10588

TITLE: Focality in Prospective Memory: A Thorough Examination of Cue Specificity and Task-Appropriateness

New Protocol/Modification/Continuation: New Protocol

Review Date: 9/6/2017

Approved ___X___ **Disapproved** __________

Approval Date: 9/6/2017 **Approval Expiration Date:** 9/5/2020

Exemption Category/Paragraph: 2a

Signed Consent Waived?: No

Re-review frequency: (three years unless otherwise stated)

LSU Proposal Number (if applicable):

Protocol Matches Scope of Work in Grant proposal: (if applicable)

By: Dennis Landin, Chairman ________________

PRINCIPAL INVESTIGATOR: PLEASE READ THE FOLLOWING –
Continuing approval is CONDITIONAL on:

1. Adherence to the approved protocol, familiarity with, and adherence to the ethical standards of the Belmont Report, and LSU's Assurance of Compliance with DHHS regulations for the protection of human subjects*
2. Prior approval of a change in protocol, including revision of the consent documents or an increase in the number of subjects over that approved.
3. Obtaining renewed approval (or submittal of a termination report), prior to the approval expiration date, upon request by the IRB office (irrespective of when the project actually begins); notification of project termination.
4. Retention of documentation of informed consent and study records for at least 3 years after the study ends.
5. Continuing attention to the physical and psychological well-being and informed consent of the individual participants, including notification of new information that might affect consent.
6. A prompt report to the IRB of any adverse event affecting a participant potentially arising from the study.
7. Notification of the IRB of a serious compliance failure.
8. **SPECIAL NOTE: When emailing more than one recipient, make sure you use bcc. Approvals will automatically be closed by the IRB on the expiration date unless the PI requests a continuation.**

* *All investigators and support staff have access to copies of the Belmont Report, LSU's Assurance with DHHS, DHHS (45 CFR 46) and FDA regulations governing use of human subjects, and other relevant documents in print in this office or on our World Wide Web site at http://www.lsu.edu/irb*

VITA

Bethany Alice Lyon grew up in northern Illinois and received her bachelor's degrees in Psychology and Biology from Augustana College. She became very interested in memory and especially how memory changes with age. She decided to pursue these interests in the Psychology Department at Louisiana State University. Upon completion of her doctorate, she will continue to learn and teach others about memory and psychological research.