THE DYNAMICS OF GLOBAL STATES IN EXECUTIVE CONTROL

by

JASON HUBBARD

A DISSERTATION

Presented to the Department of Psychology
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2017

DISSERTATION APPROVAL PAGE

Student: Jason Hubbard

Title: The Dynamics of Global States in Executive Control

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Psychology by:

Ulrich Mayr	Chairperson
David Unsworth	Core Member
Elliot Berkman	Core Member
Santiago Jaramillo	Institutional Representative

and

Scott L. Pratt	Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded June 2017

DISSERTATION ABSTRACT

Jason Hubbard

Doctor of Philosophy

Department of Psychology

June 2017

Title: The Dynamics of Global States in Executive Control

In the present work, we examine how the cognitive system responds to complex environments. It has been proposed that executive control, which is responsible for orchestrating high-level behavior in such environments, operates according to different broad processing modes, one geared towards stability and focus ("maintenance"), and the other that's open to environmental influence ("updating"). Aging work has proposed that this latter mode is over-represented in older age, leading to deficits in many, but not all cognitive domains. Across three studies, we sought to identify the dynamics of the updating state in particular, and how those dynamics are shifted in older age. In Chapter II, we used a paradigm designed specifically to enforce maintenance and updating states with an age-comparative sample, and found that older adults show increased behavioral costs (reaction times) and distractibility (distractor fixations) consistent with being "chronic updaters". In Chapter III we probed the updating state by examining spontaneous fixations towards irrelevant cues, allowing us to identify how it occurs both in response to the task context, and independently from it. We found that older adults were more sensitive to global changes in the task context (single versus mixed-task blocks), but also showed a stronger tendency to update independently from the task. Younger adults, by contrast, were more prone to update in response to transient task

events. In Chapter IV, we lay the groundwork to address these questions with neuroimaging, using machine learning to extract information regarding the task context (task set, targets, distractors, response-selection) in a task-switching paradigm on a trial-by-trial and moment-by-moment level. This opens the door for more directly measuring neural signatures of updating and gives a more high-fidelity measure to examine the dynamics of how and when it occurs. Together, this work provides some insight into the dynamics and age-differences involved in global processing states, which heretofore have been under-investigated in the literature. Additionally, we provide important analytic and methodological advancements for extending this work in the future.

This dissertation includes previously unpublished co-authored material.

CURRICULUM VITAE

NAME OF AUTHOR: Jason Hubbard

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR
San Francisco State University, San Francisco, CA
University of California, Berkeley, CA

DEGREES AWARDED:

Doctor of Philosophy, Psychology, 2017, University of Oregon
Master of Science, Psychology, 2012, University of Oregon
Master of Arts, Psychology, 2011, San Francisco State University
Bachelor of Arts, Cognitive Science, 2005, University of California, Berkeley

AREAS OF SPECIAL INTEREST:

Cognitive Neuroscience
Decision Making
Executive Control

PROFESSIONAL EXPERIENCE:

Teaching Assistant, Department of Psychology, University of Oregon, 2011-2017

GRANTS, AWARDS, AND HONORS:

Winner, 6th Annual Graduate Research Forum, University of Oregon, 2015

Henry V. Howe Scholarship, University of Oregon, 2013

fMRI Training Grant, University of Michigan, 2012

PUBLICATIONS:

Hubbard, J., Kuhns, D., Schäfer, T. A., & Mayr, U. (2016). Is Conflict Adaptation Due to Active Regulation or Passive Carry-Over? Evidence from Eye Movements.

Journal of Experimental Psychology: Learning, memory, and Cognition, 43(3), 385-393.

Hubbard, J., Harbaugh, W.T., Srivastava, S., Degras, D., & Mayr, U. (2016) A General Benevolence Dimension that Links Neural, Psychological, Economic, and Life-span Data on Altruistic Tendencies. *Journal of Experimental Psychology: General, 145*(10), 1351.

Hubbard, J., Molapour, T., & Morsella, E. (2016). The Subjective Consequences of Experiencing Random Events. *International Journal of Psychological Studies*, *8*(2), 120.

Kikumoto, A., Hubbard, J., & Mayr, U. (2016). Dynamics of Task-set Carry-Over: Evidence from Eye-Movement Analyses. *Psychonomic Bulletin and Review*, *23*(3), 899.

Mayr, U., Kuhns, D., & Hubbard, J. (2014). Long-Term Memory and the Control of Attentional Control. *Cognitive Psychology*, *72*, 1-26.

Hubbard, J., Rigby, T., Godwin, C., Gazzaley, A., & Morsella, E. (2013). Representations in working memory yield interference effects found with externally-triggered representations. *Acta Psychologica*, *142:1,* 127-135.

Hubbard, J., Gazzaley, A., & Morsella, E. (2011). Traditional response interference effects from anticipated action outcomes: A response-effect compatibility paradigm. *Acta Psychologica*, *138*, 106-110.

Morsella, E., Montemayor, C., Hubbard, J., & Zarolia, P. (2010). Conceptual knowledge: Grounded in sensorimotor states or a disembodied deus ex machina? *Behavioral and Brain Sciences*, *33:6,* 455-456.

Morsella, E., & Hubbard, J. (2010). Controlled-Reflective Processes Arise from Integrative Action-Goal Selection in the Ventral Pathway. *European Journal of Personality, 24:5,* 404-422.

ACKNOWLEDGMENTS

I would like to express my gratitude to Ulrich Mayr for his guidance over the years, and for continually presenting me with new challenges. Additionally, I would like to thank Elliot Berkman, Nash Unsworth, and Santiago Jaramillo for their thoughtful input on this manuscript. I am indebted everyone in the Cognitive Dynamics Lab who has provided support, both direct and indirect, with special thanks to Atsushi Kikumoto for imparting his EEG expertise and for his incisive feedback.

Dedicated to my wife Kellie whose love, strength, and support has brought me here today, and to my son Miles who keeps me focused on the future.

TABLE OF CONTENTS

LIST OF FIGURES

Figure | Page

LIST OF TABLES

Table Page

CHAPTER I

INTRODUCTION

A central challenge of the cognitive system is to guide behavior in a coherent fashion across varied and unpredictable environments. We know from previous cognitive and neuroscience work that the brain is modular in nature, having roughly compartmentalized processes responsible for handling different types of information (Gazzaniga, Ivry, & Mangun, 2002), and thus the challenge is to tie together these disparate systems to produce fluent and adaptive behavior. While there is considerable variability in its exact definition (Banich, 2009), the construct of executive control (also called cognitive control) is considered the process responsible for such orchestration. The challenge faced by executive control as a cognitive process, and by the researchers who study it, is that it should operate in various contexts and influence behavior across many lower-level systems. If we consider executive control as that which produces "optimal" behavior in a given environment, there is still the difficulty in defining universally-optimal behavior. In short, such behavior does not exist, as there are always cases where behavior that is adaptive in one context can be harmful in another. One prominent example is the ability to maintain focus in the face of distracting information in the environment. This is certainly adaptive for situations commonly encountered in school or the workplace, as well as many cognitive tasks (e.g., Stroop, Flanker; Stroop, 1935; Eriksen & Eriksen, 1974). However, this focus would actually be harmful in other situations, for instance while navigating a crowded space in an unfamiliar environment, or avoiding an unexpected obstacle when driving a vehicle. This ability to balance more stable or focused, versus more flexible behavior in light of changing contexts is something that has been studied extensively (c.f., Driesbach & Goschke, 2004; Aston-Jones & Cohen, 2005; Durstewitz & Seamans, 2008). It also underscores a potential solution for a cognitive system that must orchestrate complex and adaptive behavior in changing contexts—rather than coordinating all lower-level processing in a way that is tailored to each unique context, perhaps the system can simply switch between more global processing "modes" that influence how information is processed in the lower-level systems. This is a way that executive control may influence behavior in a way that's adaptive for changing environments, while reducing the complexity involved in exerting that influence.

This idea of switching between broadly stable or broadly flexible modes carries some weight in the neuroscience literature, where it has been formulated in computational models (O'Reilly, 2006; Aston-Jones & Cohen, 2005; Durstewitz & Seamans, 2008). In one case, O'Reilly (2006) proposes a model of working memory (WM), where dopamine-based interactions between the prefrontal cortex (PFC) and the basal ganglia either keep WM in a protected, "maintenance" state where its contents are resistant to change, or an "updating" state, where it is opened up to information from the environment, allowing the WM contents to be updated. Aston-Jones and Cohen (2005) propose a model based on the noradrenaline system, which has long-range projections across the cortex and thus can exert broad influence. In their model, the brain alternates between a stable, "exploitation" mode where current representations are robustly maintained, and a flexible "exploration" mode where those representations can change based on environmental influences. Similarly, Durstewitz and Seamans (2008) propose a model based on broad influences in the dopamine system, which shifts the brain between a stable ("D1"), and a flexible ("D2") state, again corresponding to the extent to which representations are influenced by the external environment.

From a somewhat independent trajectory, the aging literature has converted on a similar idea, that perhaps many of the cognitive deficits observed in older adults could actually be attributed to a bias towards a more flexible (updating) mode. This was prompted largely from the finding that, despite the well-accepted notion that older adults show deficits in executive control, they show relatively spared performance on specific measures, such as switch costs in task switching (Kray & Lindenberger, 2000; Mayr, 2001). At the same time, they show profound deficits in the same tasks when comparing performance between longer task-switching versus single-task blocks (so-called "global costs"; Mayr, 2001; Kray & Lindenberger, 2000). Mayr (2001) proposed that a parsimonious explanation for this effect is older adults' reliance on costly set-updating strategy. Further work examining this question more closely using eye tracking found that older adults were less prone to switch to a more stable state when explicitly prompted to, and that this was primarily due to a tendency to fixate on irrelevant cues in the environment (Spieler, Mayr, & Lagrone, 2006). This is consistent with other aging work in the memory literature that has found that older adults adopt costly strategies that rely on external information, even when they are capable of performing well when the information is taken away (Rogers, Hertzog, & Fisk, 2000). Lindenberger and Mayr (2014) summarize these and other findings to suggest that with older age, comes a tendency towards "chronic updating" where older adults are much more often in this global updating state. The reason this has been overlooked is that cognitive tasks can vary on the extent of their reliance on the external environment. Many of the common tasks used require focused attention, and thus older adults show large deficits, but as in the case of task switching, adopting a more flexible, updating state can be less harmful or even beneficial.

From a functional and a neurobiological level, the adoption of these global control states seems to be a plausible mechanism by which executive control operates. The aging literature also lends some support that older adults are "stuck" in the updating state. What has yet to be established are the *dynamics* of these states, how they ebb and flow in complex task environments. One straightforward question concerns the timescale of these states. Some of the work suggests that the maintenance state may be more of a default, with rapid updates to quickly refresh from the environment (O'Reilly, 2006; Aston-Jones & Cohen, 2005). But the aging literature suggests that updating may occur over much large timescales. Thus one can examine whether an updating state appears to be a transient phenomenon, or extends over longer periods of time. Another unexamined question is the extent that updating is exogenously- or endogenously-driven, in other words, triggered by the task context, or operating in a more random fashion (or both). Lastly, the age differences as summarized in Lindenberger and Mayr (2014) actually do not directly examine the question whether older adults are "chronic updaters", thus a direct test of this hypothesis is needed.

In the present project, we aimed to answer the above questions, which carry considerable methodological challenges. First, as mentioned above, being in a maintenance or updating state can have different behavioral consequences depending on the context, and thus they are difficult to infer unless there is some crisp distinction between them. In Chapter II, we build off of previous work developing a paradigm to specifically distinguish between maintenance and updating states (Mayr, Kuhns, & Hubbard, 2014). This is accomplished by having single-task blocks that promote a maintenance mode, punctuated by unpredictable, forced updates. By examining performance time-locked to these forced updates, we examine the consequences of updating, and test specifically how older adults respond differently to this context. Additionally, we use eye tracking in order to probe the allocation of attention throughout the trial in addition to the ultimate response, again time-locked to the forced updates. At each step we directly address the hypothesis of whether older adults are more prone to update. In Chapter III, we have similar forced updates, but we also specifically examine endogenously-driven updates that occur in intervening trials. We're able to do this by using eye tracking in a paradigm where participants can fixate on peripheral, irrelevant cues. Since these occur in an unambiguous task context, we use these cue fixations as an indicator of being in an updating state, and we specifically examine how these updates may be triggered by local task factors (e.g., conflict, errors), the global task context (single versus mixed-task blocks), or completely independent from these (i.e., endogenously-driven). At each level we also probe the extent to which they're sensitive to age differences.

While these data are convincing, they carry the limitations inherent in all behavioral work, where we must infer the information processing steps, while only being able to observe the ultimate outputs (RTs, eye movements). Ideally, the most powerful method would be to directly observe how different features of the task environment are processed with high temporal resolution. Accordingly, we developed a task-switching paradigm similar to those used in chapters II and III and in previous eye tracking work

(Mayr, Kuhns, & Rieter, 2013; Kikumoto, Hubbard, & Mayr, 2015), but adapted for electroencephalography (EEG) to measure brain activity on a millisecond timescale. Then, using machine learning techniques, we examine the extent to which we can extract information regarding each task element at each timepoint, including those which cannot be observed from behavior, such as the abstract task set. Further, we do this on a trial-by-trial level, circumventing the limitations imposed by other averaging techniques, like event-related potentials. These data provide the initial foundation to build on, in order to further address the questions explored in chapters II and III. Next, we can extend the paradigm to see how these abstract representational codes are modified by more stable or flexible task contexts (e.g., single versus mixed blocks, or high versus low switch rate blocks), while harnessing the high-fidelity measures developed in Chapter IV. This study was co-designed with Atsushi Kikumoto, who also provided technical background and details regarding the EEG data collection and preprocessing, and the main text (but not the extended methods) was co-written with Ulrich Mayr.

CHAPTER II

EXAMINING "CHRONIC UPDATING" IN OLDER AGE

Introduction

Executive control is considered the construct responsible for the orchestration of coherent, goal-directed behavior, which operates across varied contexts and time scales. Consequently, disruptions in executive control can lead to profound impairments in day-to-day functioning. This has led many to suggest that age-related cognitive decline is related to a breakdown of executive control specifically, since older adults exhibit a wide array of impairments (Kausler, 1991; Salthouse, 1991; Royall et al., 2004). This is also unsurprising given that executive control is associated with function of the prefrontal cortex (PFC), which undergoes large structural and functional changes in older age (Braver & West, 2008; Raz, 2000). Historically, a number of tasks have been used to index the strength or quality of executive control (e.g., Stroop, Flanker, Wisconsin Card Sorting, Tower of Hanoi, task switching; Verhaeghen, 2011; Salthouse, 2005; Banich, 2004), and in general many show reduced performance in older compared to younger adults. However, there are cases where older adults show relatively spared performance in executive control measures (Verhaeghen, 2011; Mayr, Kliegl, & Krampe, 1996; Mayr, 2001; Kray & Lindenberger, 2000). One surprising, yet consistent finding is that older adults show little difference in switch costs in task switching (Kray & Lindenberger, 2000; Mayr, 2001). This begs the question whether the apparent executive control deficits are the result of generalized degradation, or more specific alterations in component processes. One possibility that has been raised is that observed deficits in executive control arise from biases in more "global" processing states, which may be adaptive under certain conditions, but detrimental in many laboratory tasks (Lindenberger & Mayr, 2014). This runs counter to the perspective that cognitive processes are simply declining or slowing with older age (e.g., Salthouse, 1996).

These age-related biases in broad processing states have been most clearly articulated in the task switching literature, where typically age differences are prevalent in the "global" task context, but absent with regard to the "local" effects. In task switching, one can index the cognitive system's response to local shifts by comparing performance between task-switch and task-repeat trials within the same blocks (i.e., the classic switch cost). To index more global shifts in control, one can instead compare responses to task-repeat trials embedded within task-switching blocks to trials from separate single-task blocks (Kray & Lindenberger, 2000; Lindenberger & Mayr, 2014; Mayr, 2001). This "global cost" indexes the system's response to the same local demand (repeating the same task that was just performed), but with the global context of being in a switching environment or not. The common finding is that there are relatively small differences between age groups in local processing (switch costs), but robust differences in global costs (Kray & Lindenberger, 2000; Mayr, 2001). According to Mayr (2001), the most parsimonious explanation for these patterns of results is that older adults show a greater reliance on set-updating processes in general. In other words, older adults are more inclined to update the task set based on information from the environment, instead of relying on endogenously-generated representations of task sets. Using this strategy during task-switching blocks will not harm performance as much, since the task itself requires one to update whenever a switch occurs. However, during single-task blocks, the more optimal strategy is to rely completely on endogenous task-set representations, since the demands are unambiguous. Older adults do not adopt such a strategy in the single task blocks, and consequently show greater global costs.

Spieler, Mayr, and Lagrone (2006) explored this question more closely, using a task switching paradigm in which the global demands shifted over the course of a block. In this paradigm, a block would start with cued task switching, but one of the tasks would "fade out" halfway through the block, leaving only one relevant task for the rest of the block. The currently-relevant tasks were indicated by peripheral cues on every trial, with the fade-out task clearly crossed out with a thick red line. Thus, the end of these blocks were equivalent to a single-task block, and responses on these trials were compared with pure single-task blocks. They found that this fade-out cost in particular was increased in older age, suggesting that older adults had difficulty shifting from task-switching to a single-task context within the same block. The younger adults, on the other hand, rapidly adjusted to the new context and showed equivalent performance to the control blocks. Spieler et al. (2006) were also able to examine these effects more closely using eye tracking. If older adults are more prone to update from the environment, then they would be more likely than the younger adults to attend to the task cues, particularly when they were no longer relevant (i.e., in the fade-out phase). When examining fixations to the peripheral task cues they found exactly that—the behavioral fade-out costs could be attributed to the cue fixations in particular. Critically, in a separate experiment in which the task cues were taken away during the fade-out phase, they found that older adults performed similarly to the younger adults, indicating that the fade-out costs were limited to situations where redundant information is present on the screen (Spieler et al., 2006).

Lindenberger and Mayr (2014) summarize these and other findings and suggest that this dependence on the external environment may be a general pattern of aging that extends beyond the task-switching context. They suggest that it is often missed because different tasks vary on the extent that they rely on support from the environment versus internal task-set representations. Similar to Mayr (2001), they propose that with increasing age people adopt a costly "updating" policy more frequently. Both the concepts and the terminology come from previous computational work in working memory (O'Reilly, 2006). Here, "updating" refers to sampling from the environment and updating internal representations. This is in contrast to a "maintenance" mode where internal representations (e.g., task sets) are robustly maintained and resistant to what is occurring in the external environment. Lindengerger and Mayr (2014) suggest that older adults will sample from the environment even when the task context is unambiguous and they are actually capable of performing the task without external support (Spieler et al., 2006; Rogers, Hertzog, & Fisk, 2000). In some cases, such a policy is beneficial, and may counteract the cognitive system's compromised ability to maintain abstract task sets (Braver et al., 2001; Gazzaley, 2013), but in cases where the use of endogenous representations would lead to more efficient processing, such a policy can be detrimental. Importantly, adopting this updating mode of processing leads to costs only under specific circumstances, and not a generalized slowing of performance. This can explain why older adults can show spared performance in some executive control measures (i.e., local switch costs) but not others. While there is some indication across different studies that this "chronic" updating is a real phenomenon of aging, it has yet to be tested explicitly in a single study. What is needed is a paradigm that can clearly distinguish when the system is in an updating versus a maintenance state. Mayr, Kuhns, & Hubbard (2014) developed such a paradigm, which promoted a maintenance mode in most trials, with unpredictable, forced updates. This basic setup allowed them to test competing theories regarding the mechanisms involved in task switching. They found that task switch effects could be better explained based on previous experiences with tasks stored in long term memory (LTM), rather than passive carry-over of control settings between adjacent trials. This and other work (Mayr & Bryck, 2005) found specific task-switching effects in the absence of immediate task switches. The present investigation uses the some of the data from Mayr et al. (2014) and adds an age-comparative sample (Experiment 1), but focuses on the costs associated with updating, rather than the specific task-switch effects reported previously. Specifically, we sought to provide more direct evidence that older adults are chronic updaters. Additionally, we collected an independent sample of younger and older adults who completed the same task with eye tracking (Experiment 2) to more precisely characterize how an updating state influences attentional allocation in both age groups.

The paradigm used in Mayr et al., (2014) and in the present investigation involves the selection of spatial locations based either on endogenous attention (a centrally-located symbolic cue) or on exogenously-driven attention (a red, sudden-onset stimulus). By design, the exogenous task is easier to perform, as it has a more automatic pull on attention, while the endogenous task requires processing of the abstract cue. Participants perform single-task blocks of either the endogenous (hereafter, “endo”) or exogenous (hereafter, “exo”) task, and on some trials, conflict is added by presenting the stimulus for the irrelevant task. Embedded within these blocks but with a low probability (p = .25) are interruptions in which the stimuli are replaced by a math equation, requiring participants to indicate whether it's true of false. The design is such that during the regular trials (between interruptions), the task is unambiguous and thus promotes a protected, maintenance state. The interruptions, on the other hand, result in a forced update, where the task set must be abandoned in order to solve the math equation. The trial immediately following the interruption requires further updating and retrieval of the task-relevant representations from LTM. Mayr et al. (2014) found that during these immediate post-interruption trials (i.e., in the updating state) the system is particularly susceptible to competing task sets that were not distracting just 2 trials earlier (i.e., the irrelevant stimuli on conflict trials). Mayr et al., (2014) also found that, as predicted by instance theories of LTM (Hintzman, 1986; Logan, 1988), having memory traces of the competing task sets were critical for these observed interference effects, and that performing the tasks under conditions of conflict increased the strength of these traces. This was supported by the use of 2 control groups, who performed only the endo or exo task throughout the experiment. For these participants, there were still "conflict" trials, but they had no history of the irrelevant tasks themselves. Having this control can thus distinguish effects driven by the distracting stimuli themselves from having a history with the competing task set. Lastly, we included for the older adults a group that performed both tasks like the Experimental group, but without any interruptions. This allowed us to obtain a cleaner comparison of being in a more global updating context or not.

As this paradigm involves the selection of spatial locations on the screen, in Experiment 2 we use eye tracking in order to examine how attention is directed throughout each trial, indexing *how* participants are distracted in the updating versus maintenance states. This approach has also been successfully used in previous work to distinguish between different models of executive control in task-switching (Kikumoto, Hubbard, & Mayr, 2015; Mayr, Kuhns, & Rieter, 2013). One question we can answer with eye tracking which cannot be addressed with RTs alone is the extent to which the age differences are driven by early distraction from irrelevant items on-screen, versus later and more deliberative effects, such as double-checking.

Across both experiments, the findings generally support the hypothesis that older adults are more prone to update, showing both larger costs following forced updates, as well as a persisting cost that extends over multiple trials after that. Most effects cannot be described in terms of general slowing in older adults. Further, the eye tracking results reveal that there are both early and late attentional influences that drive these behavioral costs, which are much larger in older adults. However, some effects were more subtle than anticipated, suggesting that the "chronic updating" either does not occur as frequently as the other findings suggest, or that older adults have other strategies to compensate for this tendency.

Methods - Experiment 1

Participants

Participants were either undergraduates from the University of Oregon (*n*= 40) who participated for course credit, or seniors (age 65-80, *M*=70.6, *SD*=5.0) from the surrounding community in Eugene, OR (*n*=64). Data from the young group have been previously reported in Mayr et al., (2014), Experiment 1.

Participants were assigned to one of 4 groups, which determined which of the tasks they were asked to perform throughout the experiment. The Experimental group (*n*

= 41) performed alternating blocks of the endo and exo task with the math equation interruptions, while two control groups either performed the endo or exo task throughout the experiment (*n* = 40, task counter-balanced across participants). Another control group of senior participants (*n* = 23) alternated between the endo and exo tasks, but without the math equation interruptions.

Paradigm

The stimuli for the basic paradigm is depicted in Figure 1. As mentioned above, the paradigm involves two tasks, designed to pit endogenously-controlled and exogenous attention against each other. Both tasks required participants to attend to a letter within a circular frame (diameter = 13 mm = 1.5°) presented along a larger circular array (radius = 70 mm = 8°). Six of the circular frames remained on screen, and were always white on a black background. Additionally, on some trials a single, red frame appeared between any two of the white ones. During the response-stimulus interval (RSI, 1000 ms) these frames were filled with the "&" symbol (see trial sequence in Figure 2). With the stimulus onset one frame would contain either an L or R, and all other circles contained either P's or T's (note that Figures 1-3 show the stimuli for Experiment 2, which used P and K). Each trial, the task was to find the appropriate circle and press either the left or right arrow key, depending on whether it contained an L or R, respectively. In the center of the screen was a smaller array of circles (diameter of each = 4 mm = .5°; diameter of array = 14 mm = 1.6°), corresponding to each of six large white circles. This small array served as a cue for the endo task. During the RSI these circles were filled with red, then at stimulus onset all but a single circle turned white. The remaining red circle indicated which of the large circles would be the target (i.e., contain an L or R). For the exo task, the target was always the single red, sudden-onset frame that appeared between two of the white circles at stimulus onset. On these trials, this circle contained either an L or R. On conflict trials, the stimulus for the irrelevant task was presented along with the relevant one—so on endo conflict trials, the sudden-onset also appeared (but filled with a P or T), and on exo conflict trials, the central cue pointed to one of the white circles. On no-conflict exo trials, all central circles remained white. The block structure is depicted in Figure 3. Participants performed single-task blocks of 80 trials of either the endo or exo task (called the "center" and "surround" task, respectively). Performance in these single-task blocks was interrupted occasionally by math equation trials (Figure 3). On these interruption trials, a math equation (e.g., "7 * 8 24 = 32") was presented instead of the regular displays and participants had to respond with a correct/incorrect judgment. The probability of correct equations was p = .5. Incorrect equations were off by ±1 or 2. Problems were constrained to produce solutions in the positive range. Participants used the arrow keys to indicate whether the equation was correct or incorrect (left key = incorrect, right key = correct). Immediately after responding the next endogenous or exogenous-task stimulus display appeared. For each trial, the probability of a number task was p = .25, with the constraint that two interruption trials could not occur consecutively. In case of either primary-task or interruption-task errors a short error tone occurred. Participants were cued about the relevant task at the beginning of each block. Participants were randomly assigned to 4 groups. In the Experimental group, they alternated between single-task blocks of the endo and exo task. Two Control groups performed either the endo or the exo task throughout the entire experiment. These participants were still exposed to conflict trials (e.g., showing a sudden-onset on an endo trial), but had no history with the irrelevant task. This allowed us to cleanly distinguish between effects associated with the task stimuli themselves from those related to the LTM traces of each task. Lastly, for the older adults only, we had a group of No-Interruption Controls, who alternated between both tasks, but never encountered math trials. This gave us a clear index of being in an overall updating context. After 2 practice blocks, participants performed 8 blocks of 80 trials, for a total of 640 trials.

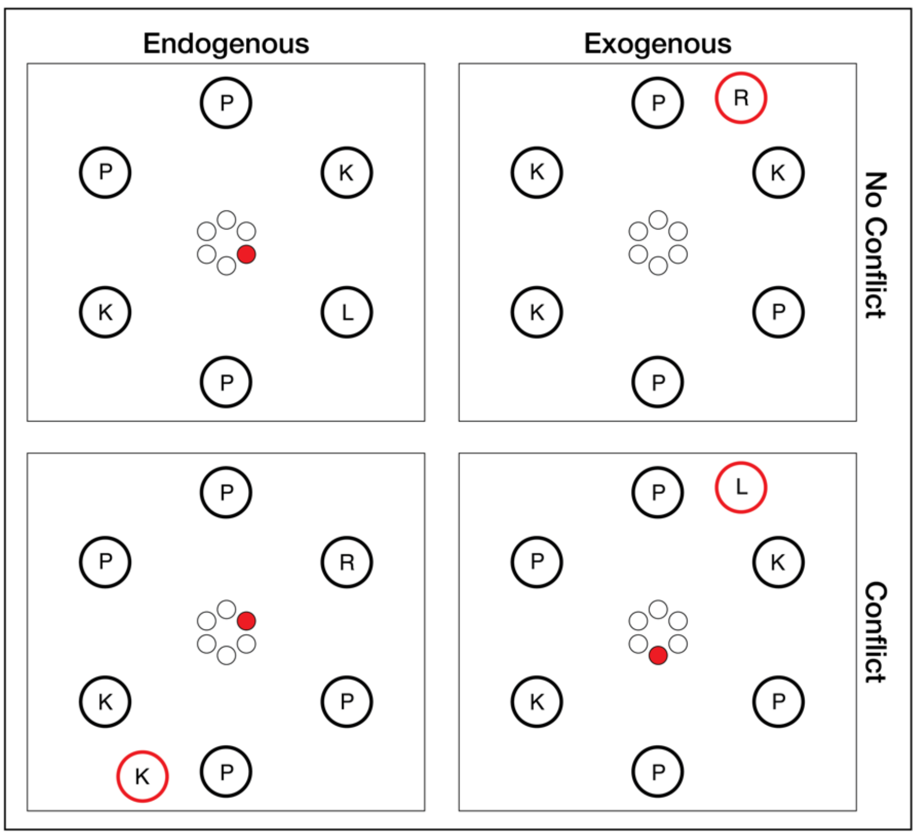

Figure 1. Stimuli used in experiments 1 and 2. Two tasks (Endogenous and Exogenous) involve attending to either a central cue, or a peripheral, red sudden-onset. The task is always to attend to the appropriate circle and press the left key if it contains an L or right key if it contains an R. On conflict trials, the irrelevant stimulus is also presented.

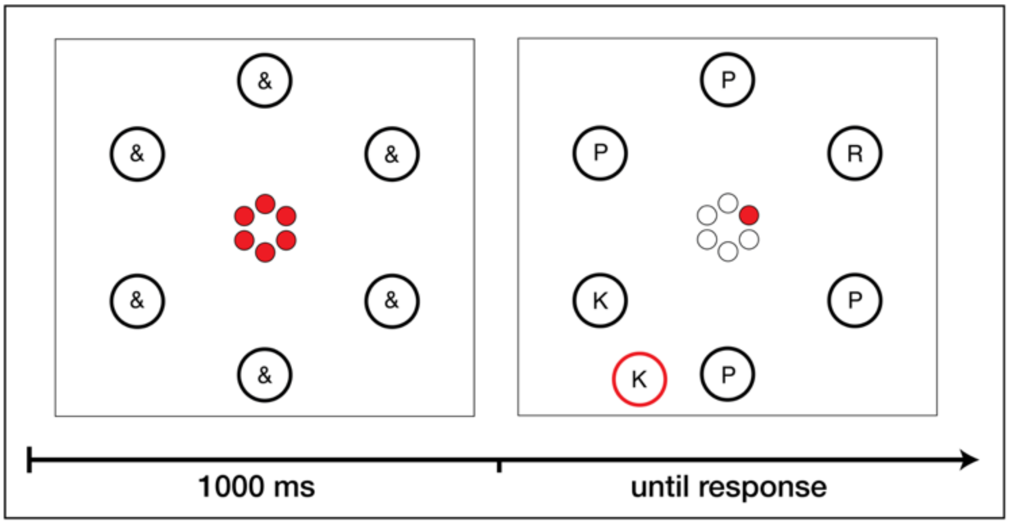

Figure 2. Trial sequence. Example trial showing the response-stimulus interval and stimulus onset. Stimuli remained on-screen until a response was made. A brief auditory tone (100 ms) was emitted when errors occurred.

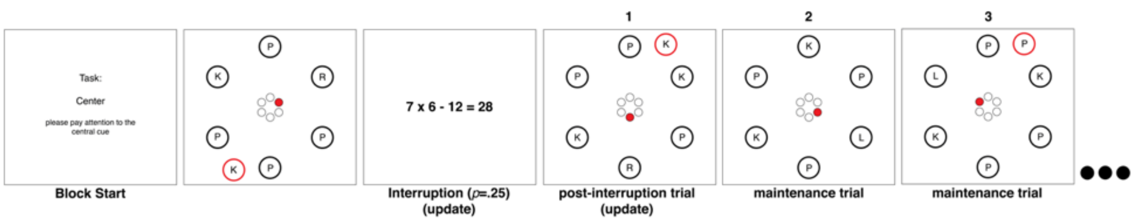

Figure 3. Sequence of trials and interruptions. Each block consisted of a single task, which was instructed at the beginning. Interruptions (math equations trials) appeared unpredictably and remained on screen until a response was made (correct/incorrect). Note that the numbers that indicate the post-interruption trial number, used in the subsequent analyses. After the interruption, the main task proceeded as before.

Methods - Experiment 2

Participants

As in Experiment 1, participants were either undergraduate students at the University of Oregon who participated for course credit (n = 23), or seniors (n =22; age 65-80, M = 70.4, SD = 4.0) recruited from the surrounding community in Eugene, OR.

Stimuli

The timing and basic stimulus setup of Experiment 2 was the same as the Experimental group in Experiment 1 (see Figure 1), but with small differences in stimulus dimensions and colors. Participants were seated approximately 65 cm from the screen. The stimulus frames (diameter of each circle = 21 mm = 1.9 degrees) were presented around a virtual circle with a radius of 71 mm (5.4 degrees). The sudden-onset frame was presented between these positions along a slightly larger virtual circle (radius = 86 mm = 6.6 degrees). As in Experiment 1, the six white stimulus frames contained the "&" symbol during the RSI (1000 ms, Figure 2), but task-irrelevant circles contained the letters P or K. The cue circles (diameter of each circle= 7 mm = 0.6 degrees) were arranged in the same manner as in Experiment 1, but instead of always being red during the RSI, could be presented in one of 6 possible colors (green, blue, yellow, orange, magenta, or purple). The color of the central circles varied randomly from trial to trial, but did not repeat across 2 consecutive trials. Likewise, the color of the sudden-onset circle varied across trials, but was always a different color from the central cue. These alternative colors were included to rule out the possibility that the color red, or any low-level priming due to the color, could be responsible for the effects observed in Experiment 1. The color did not lead to any significant differences and was disregarded in the analysis. The parameters of the math trials were the same as in Experiment 1, but the equation was always presented at a random location on the screen instead of the center, in order to eliminate any bias towards the central cue on immediate post-interruption trials.

Eye Tracking

Eye movements were measured using the SR Research desk-mounted Eyelink 1000, controlled by the Eyelink Toolbox in MATLAB (Cornelissen, Peters, & Palmer, 2002) at a sampling rate of 1000 Hz. Fixations were recorded when neither a blink nor a saccade was present, and saccades were defined for each pair of successive data samples for which the velocity of eyes exceeded 30°/s or the acceleration surpassed 8,000°/s2. Calibration for eye position registration occurred after the end of the practice blocks and repeated every 2 blocks.

Procedure

Like Experiment 1, participants performed single-task blocks of either the endogenous or exogenous task. Additionally, in order to obtain a more pure measure of the maintenance state, Experiment 2 used within-subjects manipulation where half of all blocks did not contain any math equation interruptions, just as the no-interruption controls in Experiment 1. Participants alternated between pairs of blocks with and without interruptions, and were informed at the beginning of each block whether it would contain math equations. The probability of an interruption was the same as in Experiment 1 (p = .25). Task blocks occurred in an ABBA sequence, with one AB pair containing interruptions, and the second pair without interruptions. These pairs always included one endo and one exo block (order counter-balanced across subjects). The order of no-interruption blocks was counter-balanced across participants. After 2 initial practice blocks, participants performed 12 blocks of 80 trials each.

Results - Experiment 1

Analysis

A total of 4 senior participants were excluded from analysis for failing to complete the experiment, leaving a total of 60 seniors for analysis (n = 20 for each group). The younger sample, as reported previously, included n = 20 in the Experimental, and n = 20 in the control group. For all analyses, we excluded error trials, math equation trials, trials after errors (whether they occurred after math equations or the standard tasks), and trials with RTs slower than 4000 ms. To statistically test for the behavioral effects, we used linear mixed effects modeling, using sum/deviation contrast coding (-.5 vs. .5) for the factors Conflict, Task, Experimental group, and Age group (young= -.5, old = .5). As in Mayr et al., (2014), we focus the analysis on the difference between updating trials that immediately follow interruptions (math equations), and maintenance trials that occur in the rest of the experiment (trial 2+ following the last interruption, see Figure 3). In the previous work, it was evident that younger adults enter the maintenance state by the second trial following an interruption, but in an additional step, we examine whether this is true of the older adults by comparing trials 2 and beyond following the interruption. For simplicity, trials exceeding 8 trials from the last interruption were coded as 8. We believed this provided reasonable interpretability of the results, and guarded against outliers (i.e., really long runs without interruptions) driving the effects. Trials that preceded the first interruption of the block were excluded from the analysis. Thus we can define the maintenance state as the point where RTs level off following the interruption, allowing us to index how many trials it takes to re-enter this state. Unless specified otherwise, all within-subject main effects were also included as random effects for each subject. Analyses we carried out using the lme4 package in R (Bates, Mächler, Bolker, & Walker, 2015). For each result, we report the unstandardized estimate and the corresponding t value. As p values cannot be reliably deduced from these models (Baayen, Davidson, & Bates, 2008, Footnote 1) we report only t values, but use the general rule of thumb that a t exceeding 2.0 is statistically significant.

Predictions

The paradigm is designed such that throughout the single-task blocks, the system remains in a relatively protected (i.e., maintenance) state where task-relevant representations are robustly maintained in working memory and are resistant to interference. Interruption trials then necessitate an update of working memory in order to perform the task, and trials immediately following those interruptions again require updating and retrieving the task-relevant representations from LTM (Mayr et al., 2014). Importantly, any competing task-related representations that have left traces in LTM can enter working memory during this vulnerable state. Mayr et al., (2014) found that experience with the task under conditions of conflict increased the strength of these LTM traces (Logan, 1988), and consequently, the updating effects observed. The control groups were included in order to test this LTM trace account explicitly—if the observed updating effects are due to some aspect of the stimuli themselves, then we would expect the same results regardless of whether participants had experience with the irrelevant task. If, on the other hand, it has to do with the LTM traces left through experience with the other task, then the updating effect should be absent in the single-task control groups. The primary hypothesis, that older adults have a greater tendency to update, should have two main consequences in the present paradigm. First, we would expect a greater cost on immediate post-interruption trials (i.e., the updating effect) for older adults, which should be larger in the Experimental group, and particularly large on conflict trials. Secondly, we expect that older adults will have a more difficult time entering a maintenance state—by examining the trials beyond the immediate post-interruption trial, we would expect a persisting cost for the older adults, while they younger adults re-enter a maintenance state by trial 2 post-interruption. This, again, is expected to be amplified by conflict, which can also be observed by comparing maintenance trials (i.e., trial 2+ following the last interruption) between the Experimental and the no-interruption control groups.

Updating effect

Figure 4 depicts the mean RT across each experimental condition (Task, Conflict) and group (Age, Experimental/Control group), aligned to the last interruption (math) trial. Trials that are 8 or greater from the last interruption is coded as 8. In this figure and the following analyses, we used the data trimming procedures explained above. Note that on immediate post-interruption trials there is a substantial cost, particularly in the exo task (but also the endo task in the older group). As a first step, we examined this updating effect—the increase in RT on immediate post-interruption trials—as a function of experimental condition and age group using linear mixed-effects modeling. We used a single Update factor coding for the immediate post-interruption, and since the control groups only performed a single task, we run a separate model for each task to compare across the Experimental and Control groups. To make sure we are comparing updating trials to true baseline performance, we also exclude trial 2 post-interruption from this analysis, since the overall patterns suggest that older adults do not return to baseline until around trial 3 (Figure 4). This is explicitly tested in analysis of the maintenance trials below. All results from these two models are presented in Table 1, with those of theoretical interest in bold. Of particular interest is the Age x Experimental x Update interaction, which would indicate an increased updating cost for the older adults (relative to the control participants), as well as the Age x Experimental x Update x Conflict interaction indicating an even stronger cost on post-interruption, conflict trials. As evident in Table 1, the Age x Experimental x Update interaction was significant for the endo task (b = 132, t = 2.28) but not for the exo task (b = 81, t = .89). However, the Age x Experimental x Update x Conflict interaction was significant for both tasks (endo, b = 109, t = 2.01; exo, b = 193, t = 3.41), suggesting that conflict is particularly harmful to performance on the post-interruption trials for the older adults.

Focusing on the exo task in Figure 4, it is clear that the older adults in the control condition show an updating effect, while this is not true of the younger adults. While this was not necessarily expected, it was expected that the conflict effect should only be present for the Experimental condition, which is true in both age groups, but larger for the older adults. As a follow-up, we also examined whether the Age x Experimental x Update interaction is evident in each conflict condition. Accordingly, we ran 4 additional models testing for this interaction in each level of Conflict and Task. As seen in Table A1 in the Appendix, the interaction is evident selectively in the endo, conflict trials (b = 184, t = 2.24). Together, the results indicate that older adults are indeed more perturbed by the interruption, and that this is particularly exaggerated in conditions of conflict, a pattern that is consistent with an increased tendency to update following the enforced update of the math trial.

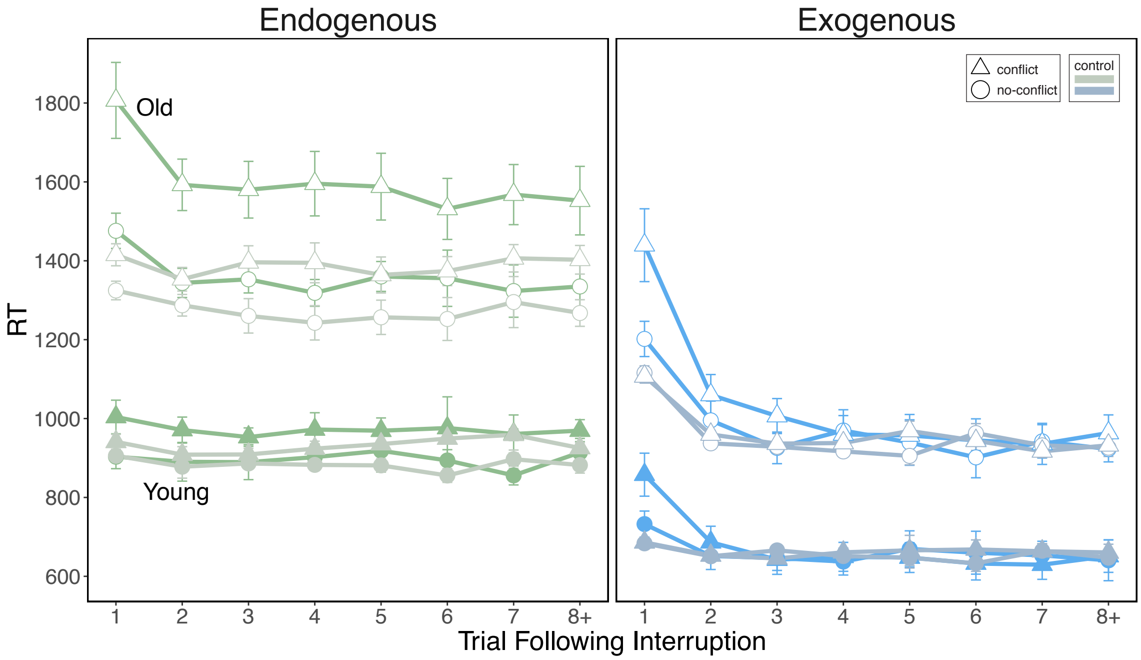

Figure 4. RT Results from Experiment 1.

Return to Maintenance State

As a next step, we want to examine differences in the trials following the immediate post-interruption trial. If older adults are prone to be stuck in an updating state, then we would expect an effect of the interruption that extends beyond the immediate post-interruption. Figure 4 suggests that this may be the case, particularly in the exo task, since RTs return almost immediately to baseline for younger adults, but seem to decline gradually in the older adults. We test for this effect by looking at the steepness of the decline in RTs as the trial moves away from the last interruption, and the extent to which it is steeper for older adults. Accordingly, we excluded the immediate post-interruption trial, and focused on the linear effect of trials 2-8 post-interruption (with 8 also including trials 9 and above) using two more models (one for each task). The post-interruption trial number was first centered and used in as a predictor along with Experimental group, Conflict and Age, while controlling for quadratic effects across the post-interruption trials. Thus a significant (negative) linear effect would reflect a gradual decline in RTs in trials beyond the immediate post-interruption. Somewhat surprisingly, there was a significant main Linear effect for the exo task (b = -8, t = 4.32) indicating a gradual decline across all groups (Table A2). But this was in the presence of the expected Linear x Age interaction (b= -11, t = 3.11), indicating a steeper decline for older adults. There was also an Experimental x Linear interaction in the exo task (b = -12, t = 3.45), indicating a more gradual decline for the experimental group relative to controls; the Age x Experimental x Linear interaction was in the expected direction, but non-significant (b = -12, t = 1.71). Lastly, there was a Linear x Conflict interaction (b = -7, t = 2.69) indicating a more gradual decline for conflict trials, but this did not interact with Age. We did not observe any linear effects in the endo task, which is unsurprising given the pattern in Figure 4. Overall, the results suggest that there is indeed a more gradual decline in RTs for the older adults in the exo task, particularly in the conflict trials, and the experimental group. We also repeated the analysis, running a model for each task and age group separately. In the exo task, the older adults showed a main Linear effect (b = -13, t = 4.80), an Experimental x Linear interaction (b = -18, t = 3.33), and a Conflict x Linear interaction (b = -11, t = 2.61). *For* the exo task in the young adults, the only significant effect was the Experimental x Conflict x Linear interaction (b_{young} = -15, t = 2.82). For the endo task, the only significant effect across both age groups was an Experimental x Conflict x Linear effect for the older adults (b = -21, t = 1.97). These results are consistent with the overall model, and show that the older adults show a more gradual decline to the maintenance state, but that the younger adults do show a similar effect in the exo, conflict trials in the Experimental group, which as reported in Mayr et al., (2014) have the most potent updating effects.

Table 1. Overall Model for Experiment 1

	estimate	*t*
Endo Task		
Age	539	8.39
Experimental	125	1.94
Conflict	136	9.86
Update	71	4.89
Age x Experimental	185	1.44
Age x Conflict	147	5.34
Experimental x Conflict	104	3.78
Age x Update	103	3.56
Experimental x Update	71	2.46
Conflict x Update	30	2.28
Age x Experimental x Conflict	134	2.42
Age x Experimental x Update	**132**	**2.28**
Age x Conflict x Update	44	1.70
Experimental x Conflict x Update	84	3.25
Age x Experimental x Conflict x Update	**110**	**2.13**
Exo Task		
Age	413	7.61
Experimental	74	1.37
Conflict	62	7.02
Update	213	8.47
Age x Experimental	63	0.58
Age x Conflict	36	2.04
Experimental x Conflict	112	6.33
Age x Update	238	4.73
Experimental x Update	142	2.81
Conflict x Update	91	6.44
Age x Experimental x Conflict	148	4.20
Age x Experimental x Update	**76**	**0.76**
Age x Conflict x Update	28	0.99
Experimental x Conflict x Update	189	6.69
Age x Experimental x Conflict x Update	**193**	**3.41**

No-interruption Controls

Figure 5 repeats the mean RTs for the older experimental group presented along the performance by the no-interruption control group (which contained only older adults).

With the other control groups, we wanted to capture the effect of updating that arose as a consequence of the memory traces associated with each task, whereas in this group we attempt to capture the cost of being in an updating versus no-updating context. Since there were no interruptions in this control groups, we compared performance on maintenance trials in the experimental group to the controls. The critical questions here are whether there's an overall cost to being in this updating context (i.e., a main effect of Experimental) and whether this interacts with Conflict. Since these control participants completed both tasks, we include a single model including both tasks and an additional Task factor. Results are presented in Table A3, which and show a larger Conflict effect in the Experimental group (b = 47, t = 2.16), which was further increased in the endo task (b = 102, t = 3.32). As a follow up we use two additional models (one for each task) to examine to what extent the effects are present in both tasks. This analysis indicated that both effects were present in the endo task, which is unsurprising given the small conflict effect in maintenance trials for the exo task (Figure 5).

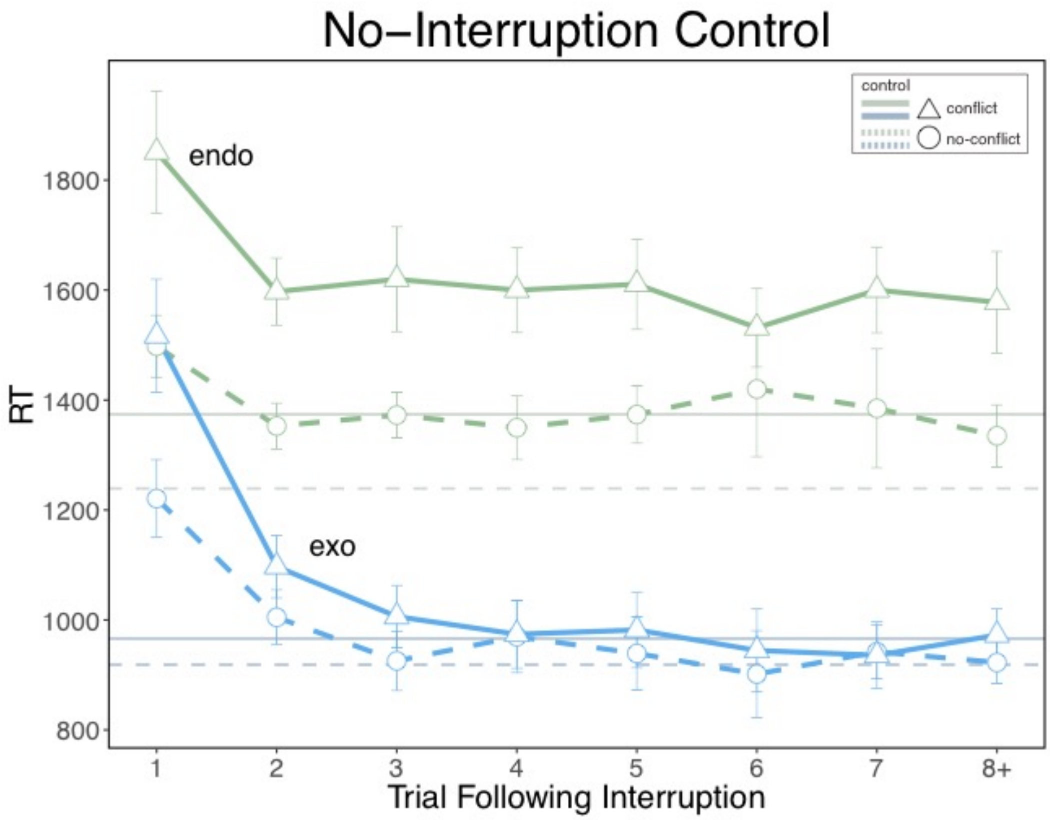

Figure 5. No-interruption control group in Experiment 1 for both the endo (green) and exo (blue) tasks. Control participants are depicted as horizontal lines, with the older adults in the Experimental group repeated from Figure 4.

Controlling for General Slowing

The models above suggest that age differences in this paradigm are particularly sensitive to the updating context, however it can be difficult to rule out the alternative hypothesis that the observed effects are due to generalized slowing which happen to occur on the slowest, post-interruption trials. This is a general concern in the aging literature, and many have found that when controlling for generalized slowing effects, the age differences aren't as broad as they seem on the surface (c.f., Mayr et al., 1996; Verhaeghen & Cerella, 2002; Verhaeghen, 2011). One way to account for such effects is to examine performance in older and younger adults as a function of how much they deviate from some baseline condition in the experiment (Verhaeghen & Cerella, 2002). In Experiment 1, we have the control group which only has experience with one of the tasks. When comparing them to the Experimental group, we can isolate the cost associated with updating when competing LTM traces of multiple tasks are present (Mayr et al., 2014), which is why the Age x Experimental x Update interaction was of particular interest in the overall analysis. To address the present question, we can look at performance in a similar manner as Figure 4, but for the Experimental group calculated the percent change from the corresponding control group. Since generalized slowing in the older adults should be present in both groups, by baselining to this performance, we emphasize the differences that are specific to the Experimental group. By calculating as a percent difference, we also normalize both age groups with respect to the degree of increase in RT as well—in other words, a 100 ms increase in the younger group may be more meaningful than the same increase in the older group. Thus we first calculated the group mean performance for the control group for each experimental factor, and for trials 1-5 post-interruption. For the Experimental group, we then calculated the average RT for each condition and post-interruption trial (1-5), and calculated the percent change from the corresponding control group. The average deviations from the control groups are depicted in Figure 6. As evident, there are large effects in the immediate post-interruption trials for the older adults above that of the younger adults. This suggests that there is indeed an extra cost for the older adults in these trials, above and beyond general slowing. To test this, we submitted the percent change values to 2 repeated-measures ANOVAs. The first examined the updating effect, contrasting the immediate post-interruption trials to the rest, with the within-subjects factors of Task, Conflict, and Update, and Age as a between-subjects factor. In this analysis, there was a main effect for the Update factor ($F(1,39) = 39.92$, $p < .001$) as well as an Update x Task interaction ($F(1,39) = 25.92$, $p < .001$), and an Update x Task x Age interaction ($F(1,39) = 6.69$, $p = .01$), indicating a larger effect for the exo task, which was even larger for the older adults. We ran a second model examining the linear effect of the post-interruption trials (excluding the immediate post-interruption) to see if older adults were also slower to return to maintenance when controlling for general slowing. This model showed a main linear effect for the post-interruption trial, $F(1,39) = 19.73$, $p < .001$, but not the expected interaction with Age ($p = .13$). There was also an interaction with Task ($F(1,39) = 6.67$, $p = .01$), and Conflict ($F(1,39) = 19.24$, $p < .001$) indicative of the steeper slope in the exo task and on conflict trials for both age groups. These factors did not further interact with age (all *p's* > .35). Together the results suggest that there is indeed a difference in the older group that can be attributed to the updating state *per se* and not just general slowing, but that the age difference in returning to the maintenance state may be a more subtle effect that is at least partly driven by slowing in older adults.

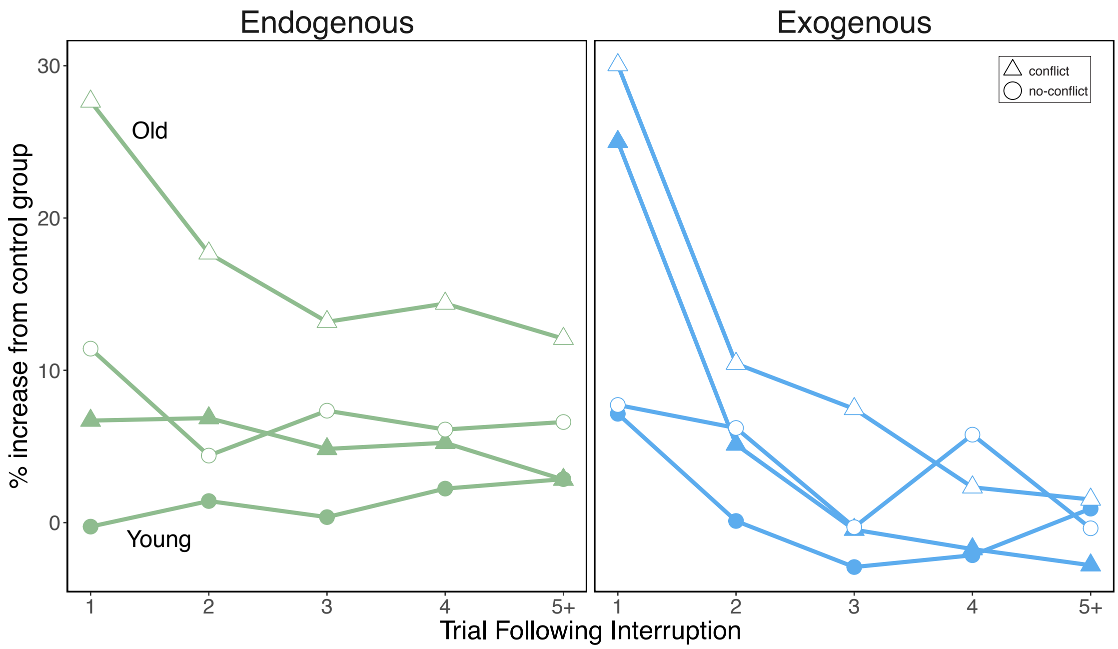

Figure 6. Visualizing performance in the Experimental group as a percentage change from the Control group in Experiment 1.

Results - Experiment 2

Out of the 45 participants, 2 younger and 4 older participants were excluded due to failing to complete the experiment, leaving a total of 21 younger and 18 older for analysis. Mean RTs for Experiment 2 are depicted in Figure 7. We used the same data trimming procedures as in Experiment 1, excluding error trials, math equation trials, trials after errors (whether they occurred after math equations or the standard tasks), and trials with RTs slower than 4000 ms. Trials were again analyzed based on their relation to the interruption, from 1 to 4+ trials post-interruption (and trials 1-8+ for the linear effect), using the same multilevel modeling techniques as in Experiment 1. In order to gain a cleaner index of the updating cost in this experiment, half of the blocks did not contain any interruptions, allowing for the same comparisons as in the no-interruption controls in

Experiment 1, but in a within-subjects manner. We will first focus on the behavioral effects in the experimental blocks (which included interruptions) as a replication of the results in Experiment 1.

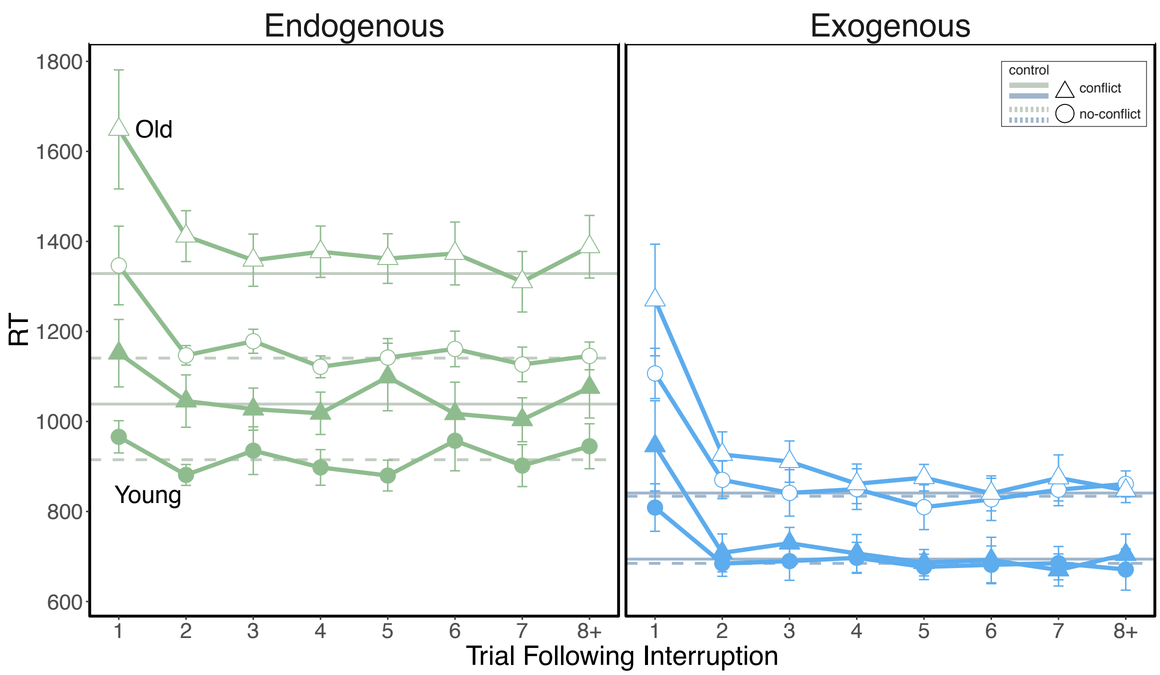

Figure 7. Mean RTs for Experiment 2. The no-interruption control blocks are depicted as horizontal lines, with solid lines indicating conflict trials.

Updating Effect

Similar to Experiment 1, we begin with an overall model specifically focusing on the Age x Updating effect. We restrict this initial analysis to the Experimental blocks only, and compare the no-interruption control blocks in a separate model. Again we focus on RT on immediate post-interruption trials, but since we do not have the same single-task control, examine the effect across both tasks. Thus the model contains the factors Task, Conflict, and Age. Again we exclude trial 2 post-interruption as it does not appear to reflect true baseline performance in the older adults. The results of this model are presented in Table A4. In this analysis, there was a significant Age x Update interaction (b = 153, t = 3.74) indicating an overall increased updating effect for the older adults. This did not interact with Task or Conflict, however. Running separate models within each task revealed that the Age x Update effect was present in both tasks (b_{Exo} = 159, t = 2.87; b_{Endo} = 162, t = 3.21). Running a separate model within each level of Conflict (Table 7) revealed that the Age x Update effect was also present in conflict and no-conflict trials as well (b_{NoC} = 138, t = 3.61; b_{Con}= 177, t = 3.54). These results replicate those of Experiment 1 in that the updating effect is larger for the older adults, but a notable difference is the presence of an updating effect for the younger adults in the endo task. However, this effect was significantly smaller in the endo task for the younger adults (b = -101, t = 5.54).

Return to Maintenance

As in Experiment 1, there does appear to be a more gradual decline towards baseline performance for the older adults, particularly in the exo task. We repeated the same analysis examining the overall decrease of RT across post-interruption trials, excluding immediate post-interruption trials, and looking at the linear effect of trials 2-8+ post-interruption, while controlling for the quadratic effect. This revealed a main Linear effect across all conditions and age groups (see Table A5; b = -3.6, t = 1.97) as well as an interaction with Age (b = -7, t = 1.97), and with Conflict (b = -6, t = 2.65), but no Age x Linear x Conflict interaction. Thus there does seem to be a more gradual decline in RT across the post-interruption trials which is steeper in older adults, and is also steeper in conflict trials. Repeating this analysis separately within each age group showed a main

Linear effect (*b* = -7, *t* = 3.19) and a Linear x Conflict interaction (*b* = -8, *t* = 2.66) in the older adults only. Younger adults did show a Linear x Task interaction (*b* = -7, 2.25), however. Together the results converge well with those of Experiment 1, which show a slower decline to the maintenance state in the older adults, and in conflict trials. A somewhat different pattern is that we see these effects across both tasks, whereas it was only present for the exo task in Experiment 1.

Table 2. Overall Model for Experiment 2

	estimate	*t*
Both Tasks		
Age	297	5.69
Task	293	43.57
Conflict	153	13.46
Update	212	10.35
Age x Task	128	9.52
Age x Conflict	70	3.08
Task x Conflict	111	8.22
Age x Update	**153**	**3.74**
Task x Update	-98	-7.29
Conflict x Update	111	8.25
Age x Task x Conflict	90	3.32
Age x Task x Update	23	0.85
Age x Conflict x Update	37	1.36
Task x Conflict x Update	-72	-2.68
Age x Task x Conflict x Update	-25	-0.46
Endo Task		
Age	358	5.53
Conflict	207	11.45
Update	163	6.46
Age x Conflict	112	3.09
Age x Update	**162**	**3.21**
Conflict x Update	75	3.72
Age x Conflict x Update	18	0.44
Exo Task		
Age	239	4.51
Conflict	92	9.22
Update	257	9.31
Age x Conflict	5	0.26
Age x Update	**159**	**2.87**
Conflict x Update	139	8.51
Age x Conflict x Update	12	0.36

No-Interruption Control Blocks

As in Experiment 1, we wanted to compare performance on the no-interruption control blocks (horizontal lines in Figure 7) to the Experimental blocks as an index of the cost of being in an updating context. Accordingly, used a mixed-effects model comparing RTs in maintenance trails in the Experimental blocks (excluding trials 1 and 2 post-interruption) to all trials in the control blocks, with Age, Task, and Conflict as additional predictors. This analysis revealed a main effect of the updating context (b = 11.81, t = 2.25), and an interaction with Conflict (b = 17.40, t = 2.70) indicating slower RTs and larger conflict effects in the experimental blocks. However, this did not interact with Age, suggesting that both groups were affected similarly by being in the updating context.

Controlling for General Slowing

We conducted a similar procedure as in Experiment 1 in order to address the notion that the age differences are attributed to general slowing effects. However, in this case we do not have the same control group that had experience with only one task. Instead, we compared performance on the experimental blocks to the no-interruption control blocks. This is a slightly different comparison compared to Experiment 1, and captures the cost associate with being in an updating versus non-updating context. Nevertheless, any age differences that manifest while controlling for these blocks will be difficult to attribute to general slowing effects. Thus for each participant, task condition, and post-interruption trial (1-5) we computed the percent deviation from the control block performance. The average deviations are depicted in Figure 8. Note that they show a similar pattern to that in Experiment 1, where older adults show larger costs on immediate post-interruption trials, above-and-beyond general slowing effects. Again we submitted these values to 2 repeated-measures ANOVAs. The first model examined the updating effect, and found a main effect of updating ($F(1,37) = 121$, $p < .001$), and an Age x Update interaction ($F(1,37) = 7.07$, $p = .01$). This model also revealed a larger updating effect in the exo task ($F(1,37) = 24.9$, $p < .001$), and in conflict trials ($F(1,37) = 20.5$, $p < .001$), but no other age interactions. In the ANOVA examining the linear effect of trials 2-5, we found only a marginally significant Linear x Task effect ($F(1,35) = 3.59$, $p = .07$), but no other age interactions.

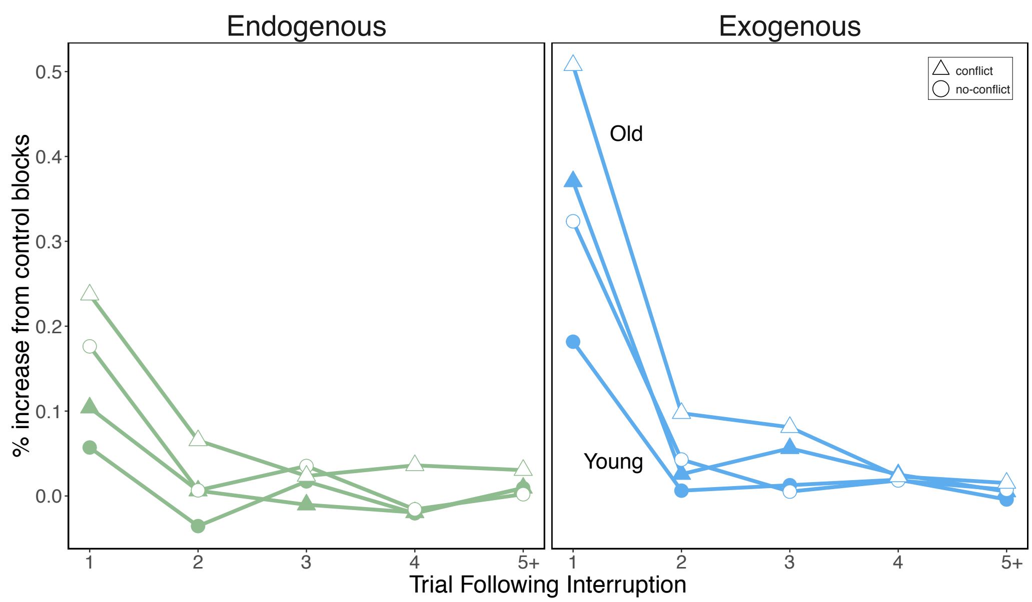

Figure 8. Visualizing performance in the experimental blocks as a percentage change from the control blocks in Experiment 2.

Eye Movements

Next we look at the eye movements in Experiment 2 to shed light on how the updating costs described above occur. One question that is difficult to address from the

RTs alone is the extent to which the costs are due to early distraction by irrelevant items on the screen or later double-checking of those items. If, for instance, updating is really a result of a more deliberative process, where the participants are simply being conscientious about returning to the task, then we would expect differences in the later part of the trial. If updating is associated more with involuntary distraction, then we would expect differences primarily in the early part of the trial. We began by establishing elliptical regions of interest (ROIs) around the sudden-onset exo stimulus (if present), the endo stimulus (if present), and a circular ROI around the central endo cue. For each item, any fixation that fell within the ROI after stimulus onset was coded as a "hit" (1) and all other fixations were coded as misses (0). The elliptical regions ensure that fixations directed towards adjacent items on the screen are not coded as hits, but those that under- or overshoot the item are still identified. Eye tracking data were preprocessed using the itrackR package in R (http://github.com/jashubbard/itrackR). Using the fixations that were coded as hits or misses for each item, we then calculated the probability of a fixation each item across time for the first 500 ms following stimulus onset. For simplicity, this interval was broken into 25ms bins. For each time bin, if any fixations during that time were coded as a hit (1) then that bin was coded as a 1. If all fixations that occurred within that time bin missed the item of interest, then that bin was coded as 0. If no fixations occurred during a given time bin, then the bin was coded as missing data. Thus each trial was represented as a binary vector indicating whether an item was fixated and at what time. Averaging across trials gives a representation of the probability of fixating an item across time. Similar to the behavioral analyses, we examine the attentional dynamics for each task and conflict condition, relative to the last interruption. For simplicity we depict only up to 4 trials post-interruption. These time-courses are depicted in Figures 9-10. As a reference for baseline performance, fixations towards each item on the no-interruption control blocks are depicted in gray and repeated across each panel. Thus the extent that fixations deviate from this baseline can be thought of as the cost associated with the updating context.

Figure 9 depicts the fixations over time for the exo task, separately for each age group, conflict condition, and post-interruption trial. Separate lines indicate the different items (the exo target versus central cue). Fixations towards the endo stimulus are not shown because they were very few trials where this item was fixated on the conflict trials. Examining the immediate post-interruption trials (the first row), clear age differences emerge. First, on conflict trials fixations towards the irrelevant central cue are increased for older adults across the entire time interval, suggesting both early distraction (before 200ms) but also later double-checking throughout the interval. The younger adults, by contrast, show only a slight pull towards the cue compared to baseline blocks around 200-500ms after stimulus onset. Fixations towards the task-relevant stimulus are also greatly reduced in the older adults in these trials, starting from very early in the trial (~150ms for young, ~200ms for old). Younger adults also show a reduction in fixations across the time interval, but to a smaller degree compared to the older adults. For the younger adults, by trial 3 post-interruption the target fixations approach that of the control blocks, particularly within the first ~250 ms of the trial, and fixations to the cue are essentially absent. For older adults, there is a greater deviation from the control blocks, even in the early part of the trial across all but trial 4+, and there is some indication of distraction by the central cue even by trial 4 post-interruption.

Turning to the no-conflict trials we can see that, despite the fact that the central cue is blank on these trials, older adults are nonetheless dwelling on this item in the immediate post-interruption trials, and this effect persists for the next couple of trials. As in the conflict trials, target fixations also show a greater deviation from the control blocks in the older adults on the immediate post-interruption trials in the early portion of the trial (< 250 ms) and persisting throughout the interval. For the older adults, this early deviation is present even in trials 2-3 post-interruption, while in younger adults this difference is almost absent by trial 2. As in the conflict trials, even in trial 4 post-interruption, older adults show a reduction in target fixations throughout the interval, while younger adults essentially return to baseline performance. Overall, examining the fixations over time reveals that the costs observed updating effects in the RTs arise from the eyes dwelling on the central cue early in the trial, but also a reduced tendency to fixate the target later in the trial. In older adults, this seems to occur even when the central cue is not indicating the conflicting task. They also reveal a delayed tendency to fixate the exo target in the early part of the trial, extending beyond the immediate post-interruptions, particularly for the older adults.

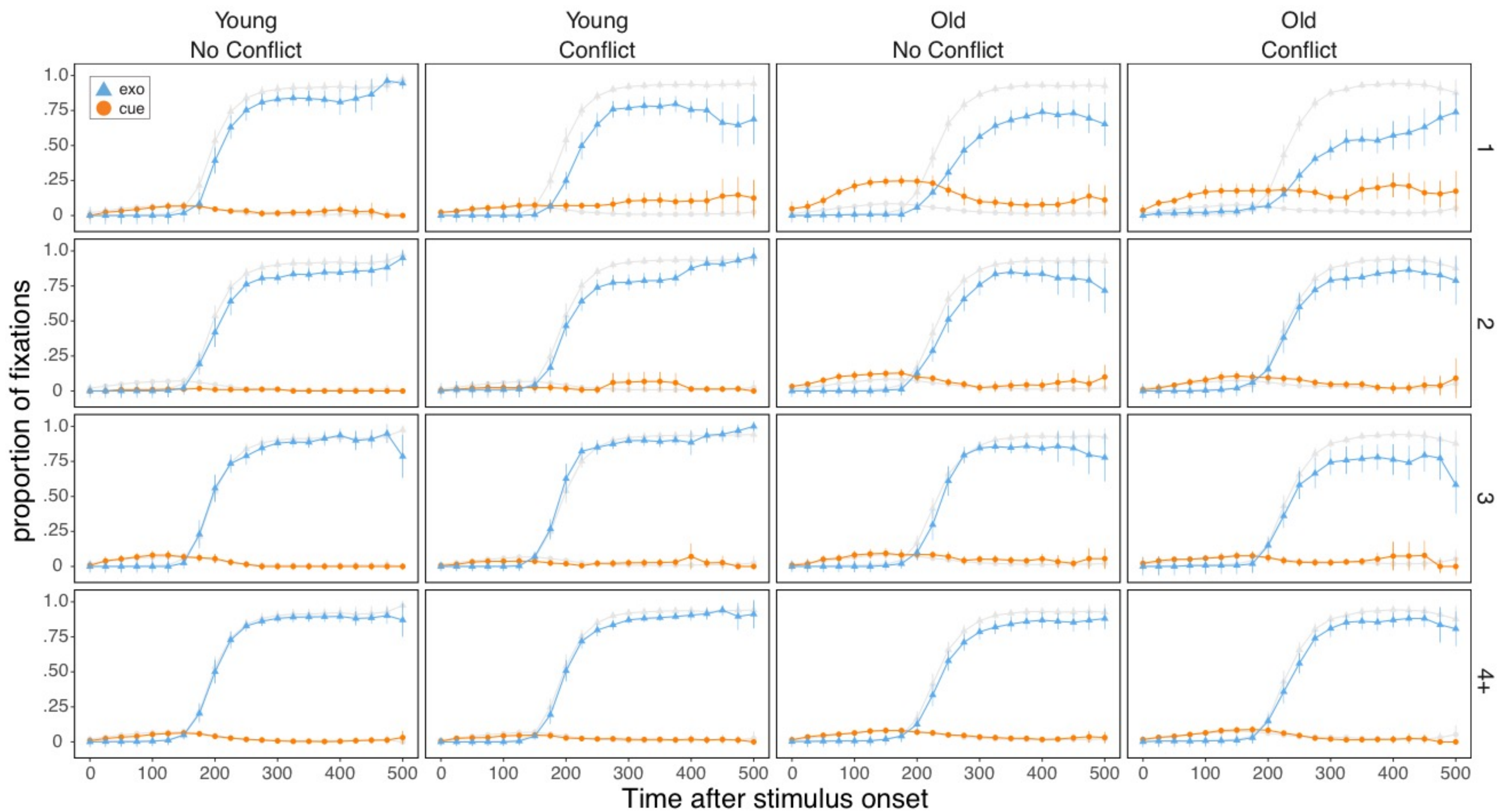

Figure 9. Fixation time-courses for exo trials, relative to stimulus onset. Rows indicate the post-interruption trial (1-4+), while the age groups and conflict conditions are in columns. The time-course for the control blocks is depicted in gray and repeated across all 4 panels. Vertical lines indicate within-subject confidence intervals calculated within each condition and age group (Cousineau, 2005). Note that the endo stimulus was present on conflict trials, but is not displayed here because there were very few fixations.

Endo Task Fixations

Figure 11 depicts the corresponding time-courses for the endo trials. In the conflict trials, we show fixations towards all three possible items (central cue, endo target, and exo distractor) since there are a considerable number of fixations towards all items in these trials. This is by design, given the distractibility of the sudden-onset. Focusing on the conflict trials, we see that there is a dramatic difference in the immediate post-interruption trials, where there is a much larger tendency to fixate the exo distractor in the older adults (blue line). In fact, after 400ms post-stimulus, older adults are equally likely to be fixating the exo distractor and the central cue. These dynamics are very different from the control blocks, where older adults show a large bias towards the central cue. In the younger adults, there is an increased tendency to fixate the distractor (and decreased tendency to fixate the cue), but by 350 ms post-stimulus, the eyes are more likely to go towards the cue. Further, by trial 2 post-interruption the fixations look very similar to the control blocks, except for a small tendency to fixate the distractor early in the trial (200-300ms), and this is essentially eliminated by trial 3. In the older adults, by contrast, there is a sizeable disruption on trial 2, and an increase in distractor fixations that persists even in trial 3. This increased tendency to fixate on the distractor is also sustained over a longer period of the trial compared to the younger adults, particularly in trials 1-2. The no-conflict trials are shown for completeness, but were not expected to reveal as dramatic of age differences compared to the other conditions. Since the exo distractor was not present in these trials, only fixations towards the central cue and endo stimulus are shown. We still see the greatest deviation from baseline performance in the older adults in the immediate post-interruption trials, but this is completely resolved by trial 2. Notably, for the older adults, we see both early and late deviations, again suggesting both early and late attentional components driving the RT effects.

Together, the pattern suggests for the endo trials, the age differences are manifested in both early and later parts of the trial, in contrast with the exo task where arguably most of the meaningful differences were present in the early part of the trial. In line with the exo task, though, we see a much greater disruption immediately following the interruption trial in the older adults.

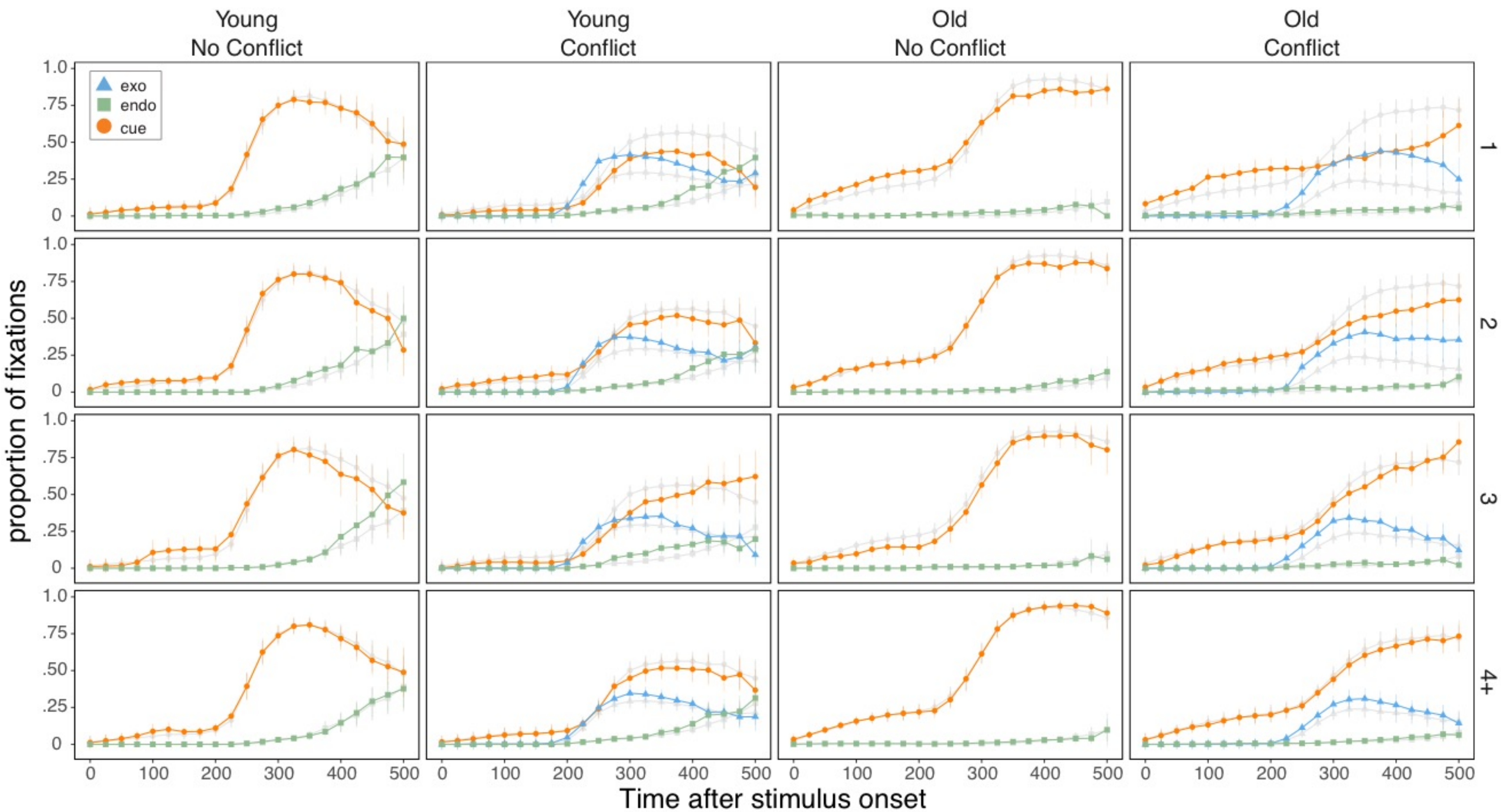

Figure 10. Fixation time-courses for endo trials, relative to stimulus onset. Rows indicate the post-interruption trial (1-4+), while the age groups and conflict conditions are in columns. The time-course for the control blocks is depicted in gray and repeated across all 4 panels. Vertical lines indicate within-subject confidence intervals calculated within each condition and age group (Cousineau, 2005).

Discussion

In the present work, we examined age differences in a paradigm that can clearly distinguish between maintenance and updating modes of processing (Mayr et al., 2014). Using this paradigm allowed us to extend previous work suggesting that age differences in executive control can be described by a bias towards "chronic updating" in older adults. Across two different experiments, the results are generally consistent with this pattern, but the age differences were not quite as dramatic as expected. In the behavioral version of the task, we found that older adults exhibited greater interference immediately following interruptions. We also found that they took more trials to re-enter a protective, maintenance state than young adults, but unexpectedly, the younger adults also seemed to take more than a single trial to enter that state as well. It is important to note that these effects were seen in a context with no task-switching demands (beyond switching from the math task) since only one task was completed during each block; despite the fact that the task was unambiguous, we observed costs that extended beyond the immediate post-interruption trial. Using control groups who only completed a single task throughout the experiment allowed us to distinguish effects driven by the stimuli themselves, versus those driven by having LTM traces of each task (Mayr et al., 2014). In the previous work, and in the age-comparative sample, we saw that having experience with both tasks leads to the strongest updating effect, particularly under conditions of conflict.

One unexpected finding was the presence of an updating effect in endo task for the older, but not the younger group in Experiment 1. Note that for the younger sample, as reported in Mayr et al. (2014), the updating effect was only present for the exo task. As explained in Mayr et al. (2014) for the younger adults, the dominance of the exo task leads to them lay relatively weak LTM traces while performing that task, since the endo distractor is easily ignored on maintenance trials, leading to little experienced conflict. Then, while performing the endo task, attention is drawn towards the distractor almost entirely by bottom-up exogenous attention (Theeuwes et al., 1998) rather than the LTM traces associated with the endo task. This operates similarly whether in the maintenance or updating state (Mayr et al., 2014). The endo task, on the other hand, has strong LTM traces due to high experienced conflict when the exo stimulus was present. Then, when performing the exo task, these LTM traces have the greatest influence in the updating (post-interruption) state, leading to a strong updating effect (see Mayr et al., 2014 for more detail). This account can also explain why older adults may show an updating effect in the endo task—if they are less prone to enter a maintenance state, then they should experience more conflict on maintenance trials, even while performing the exo task. During the endo task, when they enter the updating state, attention is pulled both by exogenous attention *and* the strong LTM traces associated with the exo task. Importantly, this explanation is supported by comparison with the control groups; if the effects observed were a result of older adults being more distracted by the exo stimulus in general, then we would expect an equally large updating effect for both them and the experimental group, but instead we see a larger updating effect for those who experienced both tasks.

For the older adults, the use of the no-interruption control group allowed us to have a baseline measure for the maintenance trials in the experimental group to more cleanly index the persisting cost associated with the forced update. This comparison revealed that older adults in the Experimental group show an increased cost in the maintenance trials overall, and a larger conflict effect in both tasks as a result of being in this context. Thus, in addition to the greater tendency to be in an updating state, being in a context where forced updates are expected leads to greater interference throughout the experiment.

One important alternative hypothesis to address regarding the age effects is the notion that the age differences are just a consequence of general slowing in older adults. To address this, in both experiments we compared performance in the experimental groups/blocks as a function of the corresponding control group/blocks. These analyses still revealed robust age-related updating effects, but little evidence regarding the gradual return to baseline performance. This suggests that the more gradual return to the maintenance state in older adults may be driven in part by general slowing in this group, but that the older adults are still uniquely sensitive to the updating trials that cannot be explained in this way.

Experiment 2 served as a replication of Experiment 1, and added eye tracking to provide additional insight into the allocation of attention throughout the trial. Additionally, we used a within-subjects design that compared performance on updating and non-updating blocks as an attempt to gain a cleaner baseline measure of the maintenance state. The behavioral results largely replicated the pattern seen in Experiment 1, with respect to the larger updating effect in older adults, and a more gradual decline towards the maintenance state. One notable exception was the unexpected updating effect in the endo task for the younger adults. This does go against the LTM trace account as described above and in Mayr et al., 2014, but notably this updating effect was much smaller when compared to the exo task for this group. In line with Experiment

1, comparison with the no-interruption control blocks revealed an apparent persisting cost and increased conflict effect associated with being in an updating context but surprisingly, these effects did not interact with age. This was not expected, given the previous evidence suggesting that older adults are uniquely sensitive to such global costs (Mayr & Lindenberger, 2014), at least in the domain of task switching. It is possible that this manipulation was not successful inducing a change in global state as intended, and instead participants simply adopted a similar global strategy regardless of which block they were in. While the fixation time-courses suggest that attentional allocation might be more optimal on these control blocks, for the most part, participants reach this optimum performance on all but the first couple trials following the interruption, again suggesting that they're operating similarly in both contexts. Clearly more work will be needed to confirm the effectiveness of this within-subject design. It may be that a between-subjects design is necessary in order to make meaningful comparisons regarding these updating costs.

Examining the patterns of fixations allowed us to directly address the question of whether the observed age-related updating effects are a consequence of early distraction by competing items on-screen, or later double-checking later in the trial. Accordingly, we examined how attention was directed towards the different items on the screen for each trial type, relative to the last interruption trial, and how the dynamics diverted from baseline performance on control blocks. Consistent with the behavioral results, we found that in the updating (immediate post-interruption) trials that older adults were disproportionately affected, and that they took more trials to approach baseline performance. For the updating trials in the exo task, the effects seemed to be driven both by early and late components, with a larger tendency to fixate the endo cue in the early part of the trial, but also an increased tendency to fixate it throughout the trial. Younger adults, by contrast, showed a slower tendency to move the eyes towards the sudden onset in the early part of the trial, a tendency that quickly went away in the subsequent trials. While the tendency to fixate the central cue was reduced in the subsequent trials, older adults continued to be delayed in fixating the sudden-onset compared to younger adults. Interestingly, in the no-conflict trials, older adults showed a very similar pattern, despite the central cue being blank on those trials, while younger adults were relatively unperturbed. On endo, conflict trials, older adults exhibited very strong distractibility by fixating the exo distractor, which persisted throughout the trial and was also present in the subsequent trials. Younger adults, by contrast, showed modest distractibility in the updating trial, which was quickly resolved by trial 2. Together, the fixation patterns suggest both early and late components of distractibility contribute to the observed updating effects, and give some indication that the age differences may be more sensitive to the later components.

Taken together, the results give some indication that older adults have an increased tendency to adopt an updating state. However, the relative weakness of the effects in the trials extending beyond the immediate post-interruption suggest that the story may be more nuanced. With a really strong tendency to update, one might expect that older adults would never return to baseline performance (as determined by the control groups/blocks), but we see that they clearly do, but just a bit more gradually than the younger adults. Similarly, the eye tracking results show dramatic age difference in the immediate post-interruption with persisting, but subtle effects in the subsequent trials.

While we believe that the paradigm is successful in promoting updating and maintenance states, one weakness is that we do not have a clear indicator of when updating occurs in all but a few trials following the interruptions. It may be that older adults are in an updating state much more frequently, but as we can only capture this through RTs or fixations, if we do not have the appropriate control conditions for comparison, they are difficult to see. In Experiment 2, there's some indication that the manipulation was not successful, and we do not have a control group as in Experiment 1 that only experienced one task. On the other hand, making the between-subjects comparisons in Experiment 1 carries its own difficulties as well. To complement these findings, one would need a paradigm where endogenously-driven (not experimentally-driven) attempts to update can be easily captured. The ideal scenario would be an index of updating that could be captured using neuroimaging, which does not rely on overt behavior. While there are theories regarding specific neural structures involved (O'Reilly, 2006), this is very difficult to capture at the slow time resolution of fMRI, and with the poor spatial resolution of EEG. Again, it will require careful experimental design to find a reliable measure.

Bridge

In this study, we examined updating as enforced by the task itself, namely, the interruption trials. This, coupled with the fact that participants performed single-task blocks, provided a relatively sharp distinction between the updating and maintenance states, as discussed above and reported in Mayr et al. (2014). Next, we want to examine updating as it occurs in a more natural context, and thus we need an overt measure to capture this updating. In Chapter III, we use a paradigm that has enforced updates, but between them we can index updating though fixations towards irrelevant, peripheral cues. This allows us to examine how updating may occur both in response to, and independently from the task context.

CHAPTER III

TRACKING THE DYNAMICS OF EXOGENOUSLY AND ENDOGENOUSLY-DRIVEN UPDATING

Introduction

In a given day, our behavior may require close attention to our surroundings, or it may require strict focus in the face of competing distractions. For instance, while navigating a crowded square in an unfamiliar foreign country, one may need to pay attention to the landmarks and crowds of people around them in order to find their destination. Conversely, while working on a time-sensitive project at work, one may need to filter out the conversations of surrounding co-workers and focus on the task at hand. Notably, the appropriateness of being either open to one's environment or filtering it out to focus on a goal is context-specific; there is no "right" way of operating across all scenarios. Experimentally we can observe similar effects in paradigms such as the Stroop or Eriksen flanker tasks (Stroop, 1935; Eriksen & Eriksen, 1974), where one must focus attention on a particular feature or spatial location and act according to instructed goals, often suppressing the urge to perform a more automatic task (e.g., reading the word in the Stroop task). This ability to suppress more prepotent actions in the service of an instructed goal is a central aspect of executive control, and given the higher-order nature of this construct, deficits in executive control can lead to wide-ranging dysfunction. In older age, we see an apparent decline in tasks that require focused attention, as well as an array of other functions (Kausler, 1991; Salthouse, 1991; but see Verhaeghen, 2011), leading some to posit that with older age comes a specific breakdown of executive control. Task switching has also been used as a hallmark measure of executive control, but unlike the Stroop and flanker, requires more flexibility in behavior, frequently in response to rapidly-changing and unpredictable cues. By contrasting performance on task-switch versus task-repeat trials, the switch cost provides an index of the high-level processes responsible for re-configuring the current task set (Rogers & Monsell, 1995). Surprisingly, in spite the robustness of switch costs and the clear age-related decrements in executive control as measured through Stroop-like tasks, some have found relatively spared performance in switch costs in older age (Mayr, Kliegl, & Krampe, 1996; Mayr, 2001; Kray & Lindenberger, 2000). Upon closer inspection, the story is not so contradictory, and substantial age differences *do* emerge when comparing performance between longer periods where task switches are expected to occur or not (Mayr, 2001; Spieler, Mayr, & Lagrone, 2006). So-called "global costs" compare performance between single-task blocks and the no-switch trials in mixed blocks. This indexes the cost associated with being in a task switching context, even when the current trial itself does not involve a switch. Importantly, these dissociations underscore a more general way of interpreting performance on executive control tasks—on one hand, it involves responses to local changes that occur in the moment (for instance, whether the task just switched or repeated), but it is also dictated by the expectations of the global task context (e.g., being in a single versus mixed-task block). This framing converges well with other work suggesting that the brain switches between different broad processing "modes" which influence how readily information from the external environment is processed (O'Reilly, 2006; Durstewitz & Seamans, 2008; Aston-Jones & Cohen, 2005). There are different theories regarding the exact mechanism, but the commonality is that there are two broad modes, one characterized by a focused state where the currently-relevant representations are well-protected from the influences of the outside environment, and another flexible state that is biased towards taking in information from the environment and updating those representations (hereafter called the “maintenance” and “updating” modes, respectively; O’Reilly, 2006). Importantly, as mentioned above, general stability or flexibility can be adaptive for different types of scenarios, as purely focused behavior would be too rigid, and purely flexible behavior would be too distractible. Recently, it has been proposed that the deficits observed in older age are a product of older adults being chronically “stuck” in this latter updating state (Lindenberger & Mayr, 2014). This bias towards updating essentially makes older adults more reliant on the external environment—a strategy that can be adaptive in some circumstances (particularly when one’s ability to maintain internal representations is compromised; Gazzaley, 2013), but can be detrimental particularly in tasks that require selective attention. This also explains why others have found that older adults are prone to focus on information present in the environment even when it’s completely irrelevant (Spieler et al., 2006; Rogers, Hertzog, & Fisk, 2000). Across studies, there is evidence suggesting that older adults are indeed “chronic updaters”, but there has been little work examining this question directly. Further, the few cases where it has been examined have focused on a relatively broad level—for instance the transitions going from a mixed to a single-task block (e.g., Spieler et al., 2006). To our knowledge, no one has explicitly examined the properties of a task context that drive transitions between the maintenance and updating states at a fine-grained level (but see Mayr, Kuhns, & Hubbard, 2014), and the extent to which such transitions are endogenously- or exogenously-driven. Accordingly, in the present work, we sought to identify the antecedents to the updating state in a complex task-switching scenario, and how such antecedents may differ across age groups. There are several possibilities regarding what drives these transitions, and age differences may be manifested at any one of them. First, there are local properties of the task that would be expected to induce a switch towards updating within a given trial, or in a subsequent one. This includes conflict (at the attentional selection or response selection phases), errors (both current-trial errors and trials after errors; Dutilh, Forstmann, Vandekerckhov, & Wagenmakers, 2013; Ratcliff, Thapar, Gomez, & McKoon, 2004; Starns & Ratcliff, 2010) or task difficulty. As suggested in the aging literature, there are also properties of the global context that may lead to increased adoption of one mode or another (e.g., single-task versus task-switch blocks). From the previous work, we would expect that these global influences will be particularly sensitive to age differences in updating (Mayr, 2001; Spieler et al., 2006; Kray & Lindenberger, 2000). There is also the possibility that these transitions are (at least partly) endogenously-driven, occurring independent of the task context. One way this may be manifested is a passive carry-over phenomenon, such that cue fixations on one trial will be highly predictive of cue fixations on the next, independent of any task factors. Previous work has shown that such carry-over models of high-level control can explain different effects in the domain of task switching (Mayr, Kuhns, & Reiter, 2013; Kikumoto, Hubbard, & Mayr, 2015) and conflict adaptation (Hubbard, Kuhns, & Mayr, 2016).

In order to examine these questions, we need a reliable indicator of being in an updating state, and importantly, we want to measure cases where such updates are not necessarily required by the task itself. Otherwise we are effectively measuring participants' adherence to the task rules. With this in mind, we used a modification of the paradigm used in Spieler et al. (2006) where task switches were infrequent and thus most of the time the current task was unambiguous. This minimized any external task demands requiring an updating state. As updating is defined as an increased reliance on the external environment, we then used eye tracking in order to track the extent to which participants focused on information presented on-screen. We did this with the use of distant task cues that were always present, but only informative at very particular, well-defined points. Any fixations to these cues that occurred on different trials were deliberate and, given the unambiguous task context, unnecessary. For these reasons, we regarded these redundant cue fixations as an indicator of being in an updating state. We then examined the local, global, and endogenous factors that were predictive of these fixations, both within and across age groups. Experiment 1 used a similar cueing procedure used in Spieler et al. (2006) combined with an attentional capture paradigm used previously to examine updating effects in task switching (Mayr, et al., 2014). This was performed by a group of undergraduate students, and examined both task-level (local) and block-level (global) effects that influenced updating. This paradigm primarily involves conflict in directing attention to different spatial locations, in some cases broadening the attentional window. This in itself could have led to task-specific effects influencing cue fixations. Therefore, in Experiment 2 we used an identical cueing procedure, but coupled with a task that involved focusing on the center of the screen, with a response-selection demand as seen in previous interference tasks (e.g., Stroop flanker). Therefore, cue fixations could not be explained in terms of the task demands. We used a variation of the Stroop paradigm which requires a manual, rather than verbal response to color words presented in different colored "inks" (for reviews, see Lu & Proctor, 1995 and MacLeod, 1991). Experiment 2 served as a replication, as well as the addition of an age-comparative sample to specifically examine the hypothesis that older adults update more frequently and identify the factors that give rise to this increased updating. To preview, we found that participants exhibited a sizeable number of unnecessary cue fixations in both experiments (indicative of updating), and that some local task factors (e.g., conflict) triggered updating in younger adults, while carrying little influence in older adults. Conversely, in line with previous work (Lindenberger & Mayr, 2014; Kray & Lindenberger, 2000), older adults were largely driven by the global task context, as well as more endogenously-driven "inertia" from updating in previous trials. Importantly, unnecessary cue fixations were associated with worse performance on subsequent trials, suggesting that they were not merely a strategy to increase accuracy on the task. Overall, results were indeed consistent with the notion that older adults are "chronic updaters" and also provides a more complete picture regarding the dynamics of how and why shifts in broad control states occur.

Methods

Subjects

In Experiment 1, $n = 31$ students from the University of Oregon participated for course credit. Experiment 2 included a total of $n = 34$ University of Oregon students and an age-comparative sample of $n = 31$ seniors from the surrounding community in Eugene, OR. Seniors were compensated $15/hour for their time.

Overall Procedure

Both experiments involved the same basic structure, but differed in the particular tasks used. Each was based on a variable-runs task-switching paradigm (e.g., Altman & Gray, 2002), where the currently-relevant task was indicated by peripheral cues at the upper corners of the screen. Figure 1 depicts the basic cueing procedure used across both experiments.

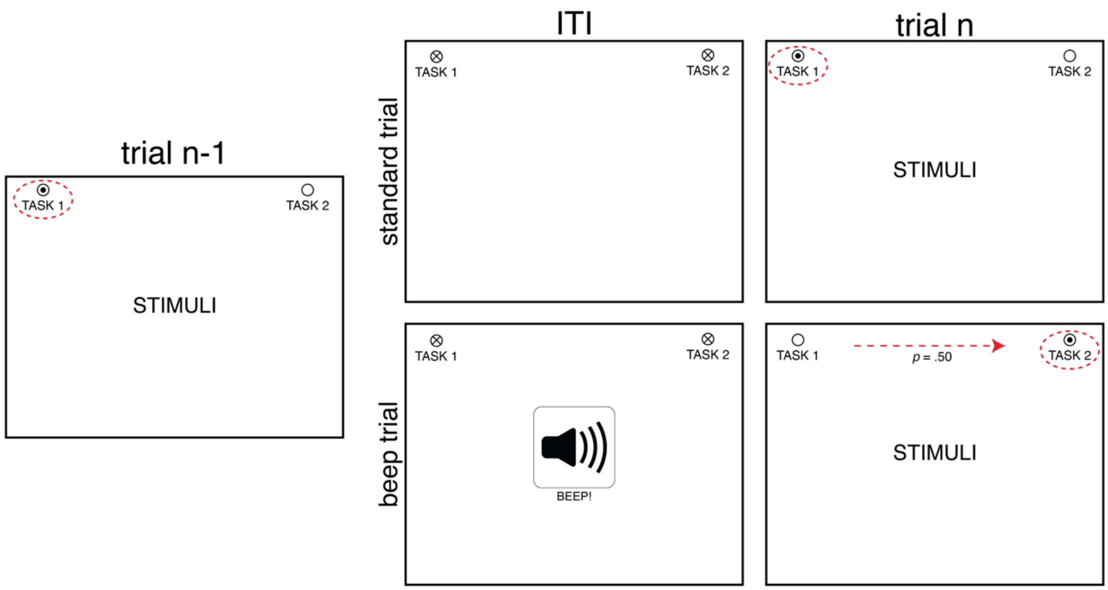

Figure 1. Basic cueing procedure used across both experiments. The currently-relevant task is indicated by the small dot that appears in either the left or right circle. The current task is highlighted here for illustration. Note cues are not to scale.

Each cue consisted of a word (*CENTER*/*SURROUND* for Experiment 1, *COLOR*/*WORD* for Experiment 2) underneath a small circle. A small dot appeared in the center of the circle corresponding to the currently-relevant task. Task switches were signaled by an auditory cue (500 hz beep) that indicated that a 50% probability of a task switch. Due to their size and location, participants had to fixate the cues on these "beep trials” in order to know the current task. Each block began with a beep trial to indicate the starting task, and beeps occurred in either 1 out of 8 (p = 12.5%) or 1 out of 12 (p = 8.33%) trials. Task switches could only occur on beep trials, and thus the currently-relevant task should be unambiguous on intervening trials. A central focus of the analysis is on the occurrence of unnecessary cue fixations on these intervening trials. Each experiment began with a short practice block involving task switches, followed by two single-task blocks (which contained beeps, but participants were informed that the task would never change), then mixed blocks (14 blocks of 80 trials in Experiment 1, 10 blocks of 80 trials in Experiment 2) followed by two more single-task blocks.

Eye Tracking

Eye movements were measured using the SR Research desk-mounted Eyelink 1000, controlled by the Eyelink Toolbox in MATLAB (Cornelissen, Peters, & Palmer, 2002) at a sampling rate of 1000 Hz. Fixations were recorded when neither a blink nor a saccade was present, and saccades were defined for each pair of successive data samples for which the velocity of eyes exceeded 30°/s or the acceleration surpassed 8,000 $°/s^2$. Calibration for eye position registration occurred after the end of practice blocks and repeated every 2 blocks.

Paradigm – Experiment 1

In Experiment 1, participants completed a modified version of a paradigm reported previously (Mayr et al., 2014) to examine the influences of long term memory traces and updating on task-switch effects. The paradigm consists of two tasks, where participants made speeded responses to a letter (L or R) that appeared either within one of six stimulus circles indicated by the central cue (endogenous attention task) or within a peripheral, sudden-onset stimulus (exogenous attention task; see Figure 2). The endogenous (hereafter, “endo”) task required processing the symbolic central cue in order to direct attention to the appropriate location, while the exogenous (hereafter, “exo”) task required more bottom-up attention towards the sudden onset (Theeuwes, Kramer, Hahn, & Irwin, 1998), which contained the relevant stimulus. Conflict was generated on 50% of trails by presenting not only the currently relevant cue/stimulus (e.g., the central cue), but also the currently irrelevant cue/stimulus (e.g., the sudden onset). On each trial, the target stimulus (as indicated by the central cue, or the sudden-onset) contained an L or R, indicating the participant should press the left or right arrow key, respectively. The currently-relevant task was signaled via peripheral cues as discussed above and would only switch on half of the rare “beep trials”. For participants, the endo and exo tasks were named the “center” and “surround” task, respectively.

Stimuli – Experiment 1

Figure 2 presents the basic stimulus setup for Experiment 1. Participants were seated 65 cm from the computer display. Six circular stimulus frames (diameter of each circle = 21 mm = 1.9 degrees) were presented along a virtual circle (radius = 75 mm = 6 degrees) around the screen's center. These circles where always presented in white on a black background. Within each circle the "&" symbol or the letters L or R could be presented in white, size 14 Helvetica font. An additional, sudden-onset circle of the same size could appear between two of the regular circle positions (90 mm = 7.25 degrees from center). This circle was always presented in red and could also contain the letters L or R in white, size 14 Helvetica font. At the center of the screen there were six smaller cue circles (diameter of each circle= 7 mm = 0.6 degrees). These were arranged in a way that mirrored the larger set of 6 stimulus circles (radius of the central cue circle= 12 mm = 1.1 degrees), such that for each position in the larger stimulus circle, there was a corresponding, smaller cue circle. The smaller cue circles could be presented either in white or red.

The stimulus display also contained two task cues in the upper left and the upper right corners (27 degrees from screen center). The task cues consisted of a verbal task label (*CENTER* or *SURROUND,* size 14 Helvetica font) along with a small circle (diameter = 9 mm = 0.8 degrees). During inter-trial-interval (ITI), both of these circles were filled with small dots. With stimulus onset, the dot remained in the center of the relevant-task circle, while the irrelevant-task circle was empty. This ensured that cues were uninformative during the ITI, and that the relevant cue would not be a sudden-onset stimulus, which would draw attention in an automatic manner (Theeuwes, et al., 1998). Because the dot was small, it was necessary to fixate the task cues in order to identify the current task on beep trials. Since task switches could only occur on beep trials, it should be unnecessary to fixate cues on other trials.

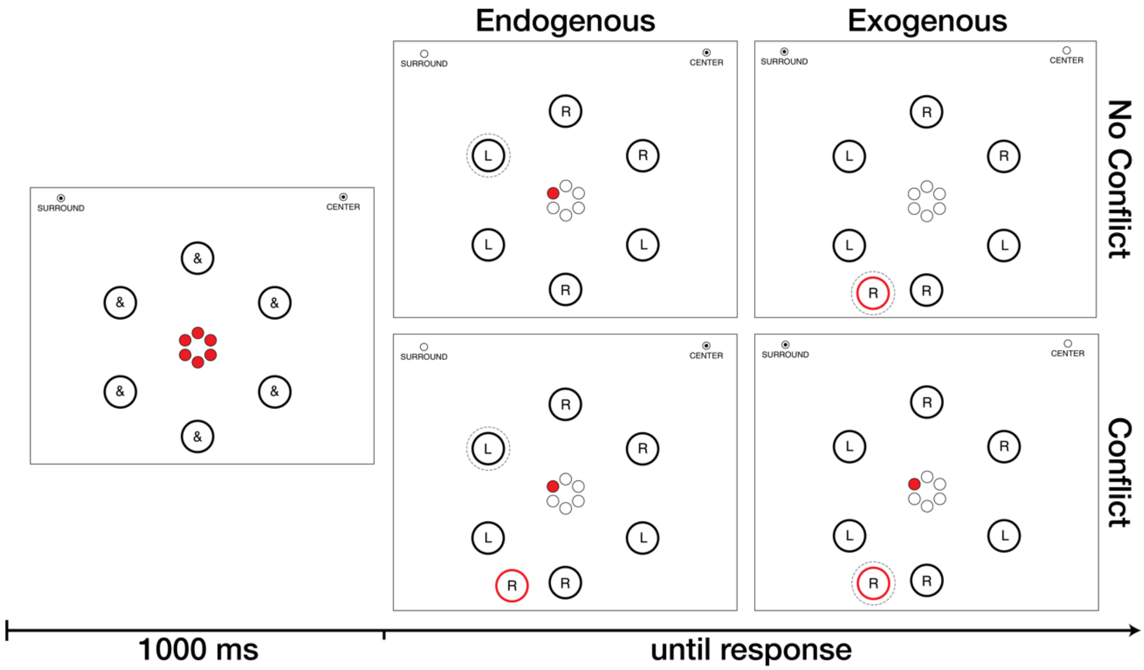

Figure 2. Stimuli and timing used in Experiment 1. Participants either focused on the central cue (Endogenous task) or red sudden-onset (Exogenous task). For illustration, the current target is highlighted here by a dotted line.

Procedure – Experiment 1

Each trial began with a 1000 ms ITI in which all of the large peripheral circles contained the "&" symbol and the central cue circles were all filled in red (Figure 2). In the endo task, at stimulus onset all central cue circles turned white except for a single one that remained red. This red circle indicated which of the corresponding peripheral circles contained the correct response-relevant stimulus (L or R). In the exo task, the response-relevant stimulus was embedded in the single red sudden-onset circle that appeared between the peripheral circles. Conflict was manipulated by presenting stimuli for the irrelevant task on 50% of trials. Thus, for the endo task, on conflict trials a red sudden-onset stimulus appeared (Figure 2). For the exo task, on conflict trials a central cue was presented (i.e., one of the central cue circles remained red) while on the no-conflict trials, all cue circles turned white. Additionally, on each trial all non-target circles also contained L's and R's in order to maximize the need to fixate the appropriate stimulus. Stimuli remained on screen until a response was made. Error feedback was given in the form of a red flash of the screen for 200 ms, while on correct trials, the circles were filled with the "&" symbol for this period.

Participants performed one 40-trial practice block followed by 14 test blocks with 80 trials each. Depending on the between-subject condition there was either a .085 or a .125 probability of a "beep trial". On beep trials, a 500 Hz tone at the beginning of the ITI indicated there was a p=.5 probability of a change in task, thus prompting participants to inspect the task cues in the upper corner of the screen. For "non-beep trials", participants were instructed that the task always remained the same as on the previous trial and cue-inspections were not necessary. Twelve participants were exposed to a .085 rate of beep trials, and 16 participants to a .125 rate. This factor interacted in no systematic way with the effects of interest and therefore will be ignored for the analyses.

Paradigm – Experiment 2

Experiment 2 was a variant of the Stroop paradigm with a manual response (MacLeod, 1991; Lu & Proctor, 1995). On each trial a color word was presented in either a matching color (e.g., *RED* in red "ink") or a mismatching color (*RED* in blue ink; see Figure 3). This word was flanked on the left and right by two squares, each presented in a particular color. These squares served as response cues—participants were required to press either the left or right arrow keys to indicate the correct response color. As in the typical Stroop paradigm, conflict/incongruent trials where defined as those where the word did not match the ink color. In these trials, one square always matched the color corresponding to the word, and the other square corresponded to the ink color. On no-conflict trials, where the word matched the ink color, one square corresponded to that color, and the other was a random color chosen from the remaining set of colors. In order to reduce the possibility of direct stimulus repetitions across subsequent trials (Mayr, Awh, Laurey, 2003), this paradigm used a set of 8 colors (red, green, blue, yellow, orange, purple, pink, and brown). Conflict was present on 50% of trials. As in Experiment 1, the currently-relevant task was indicated by peripheral cues (*COLOR* or *WORD*) using the variable-runs task-switching procedure discussed above.

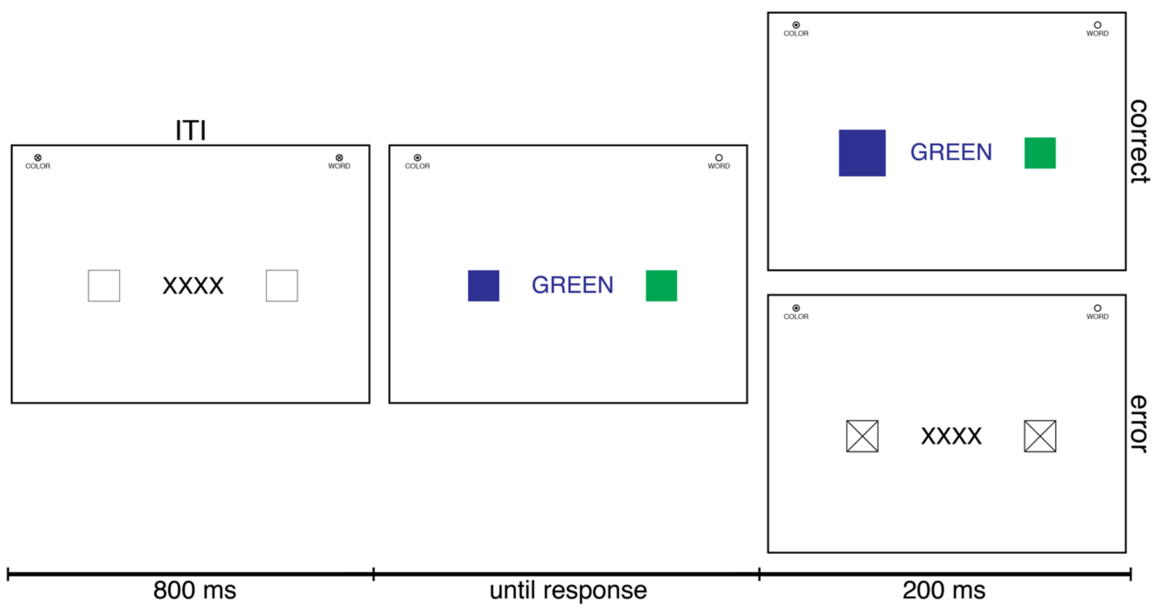

Figure 3. Stimulus Sequence for Experiment 2.

Procedure – Experiment 2

Each trial began with an 800 ms ITI in which the two response squares were filled with white, and four X's were displayed in place of the central word. During this time, both peripheral cue circles were also filled small X's. At stimulus onset, one of the 8 color words appeared in the center of the screen (size 40 Helvetica font), the response squares (width = 3.14 degrees, 6.28 degrees from center) were filled with the corresponding colors, and a dot appeared in the center of the currently-relevant peripheral cue. Stimuli remained on screen until a response (left or right arrow key) was made. For correct responses, the chosen square was enlarged by 20% for 200ms. For errors, the word changed to four white X's, and the response squares turned white and large black X's in the center of each for 200ms. Participants first completed a short 20-trial practice block (which included beeps and switches), followed by two 80-trial single-task blocks (one for the color task, one for word, counter-balanced), then 10 mixed blocks (80 trials each), and two more single-task blocks. For Experiment 2, beeps occurred at a rate of 12.5% (1 out of 8 trials). On these trials, the beep occurred in the last 200ms of the ITI, just before stimulus onset. In addition to undergraduate students at the University of Oregon, Experiment 2 also has an age-comparative sample of seniors (age 65-80) recruited from the surrounding community in Eugene, OR.

Results

Out of the 31 participants in Experiment 1, three were excluded due to corrupted eye tracking data, leaving a total of 28 for the analysis. Of the 34 younger adults who participated in Experiment 2, two were excluded due to corrupted eye tracking data, two were excluded for failing to finish the experiment, and two were excluded for failing to follow task instructions (i.e., fixated the cues on less than 25% of beep trials). From the 31 seniors in Experiment 2, two were excluded for having error rates in excess of 40%, and one was excluded for failing to finish the experiment. This left a total of 28 for analysis in each age group for Experiment 2.

Basic RT Effects

In Experiment 1 the exo task is much easier to perform than the endo task, by design (Mayr et al., 2014). The mean RTs for each task, conflict condition, and block type (single vs. mixed) are depicted in Figure 4. These were analyzed using linear mixed-effects modelling using the lme4 package in R (Bates, Mächler, Bolker, & Walker, 2015), with the factors Task (1=endo, 0=exo) and Conflict (1=conflict), and a separate model for single and mixed blocks. In these and subsequent analyses, we excluded error trials, trials after errors, beep trials, trials after beeps, the first trial of each block, and very slow responses (RTs > 4000ms). As expected, the endo task had much slower RTs in both block types ($b_{\text{Single}} = 202$, $t = 6.67$; $b_{\text{Mixed}} = 176$, $t = 13.07$), and showed a greater conflict effect ($b_{\text{Single}} = 116$, $t = 8.75$; $b_{\text{Mixed}} = 91$, $t = 8.80$). We also observed a conflict effect for the exo task selectively in the mixed block ($b = 31$, $t = 2.11$).

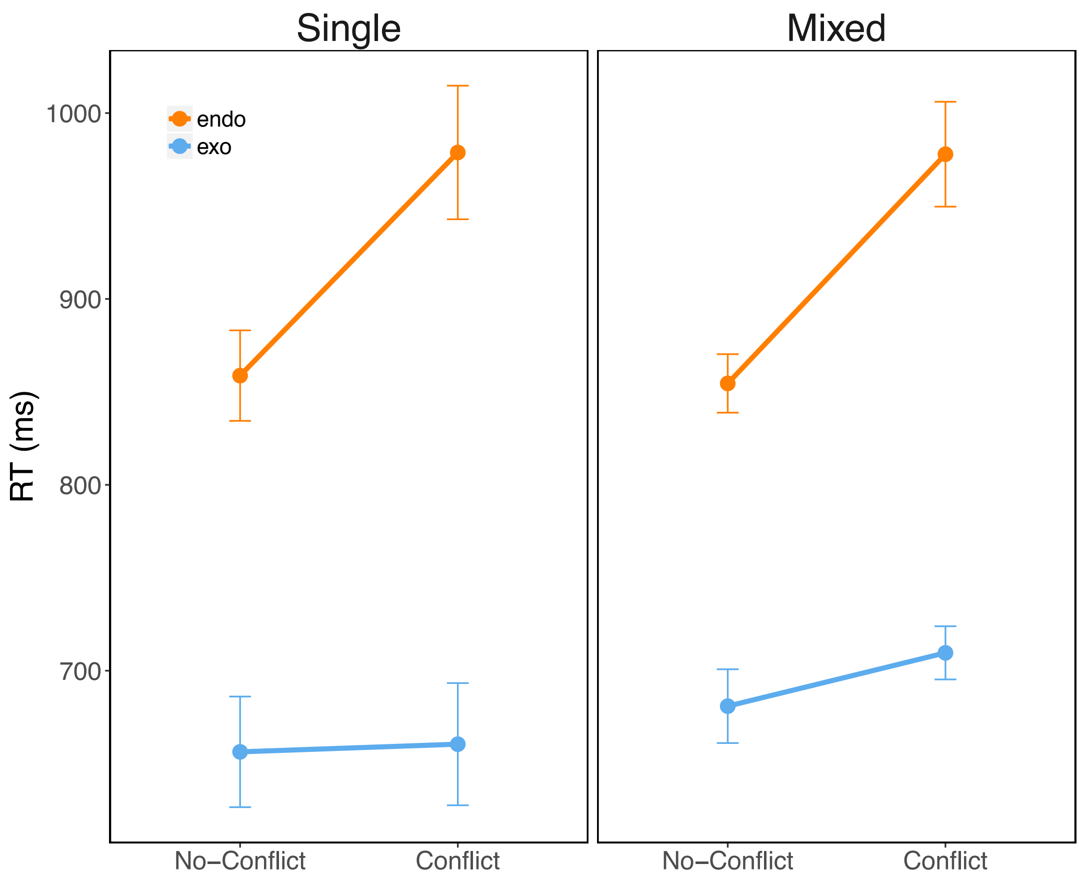

Figure 4. Mean RTs for Experiment 1. Error bars indicate within-subject standard error (Cousineu, 2005).

Figure 5 depicts the RTs for Experiment 2 for each task, conflict, and block condition, as well as each age group. These were analyzed using linear mixed-effects models in the same manner as in Experiment 1. While there is an expected overall increase in RT in the older adults, there was also an unexpected difference by task, such that the younger adults responded more quickly for the color task (consistent with previous findings in similar manual Stroop tasks, Lu & Proctor, 1995) while older adults showed the reverse effect, responding more quickly for the word task (Age x Task interaction, $b_{Single} = -339$, $t = -4.83$; $b_{Mixed} = -201$, $t = -4.17$). Previous work demonstrating a reverse Stroop effect in older adults is lacking, thus more work is needed to establish whether this differential sensitivity to the tasks can be replicated. As we are focusing predominantly on the cue fixation effects, this difference has no consequence for most of the analyses performed here, except when examining the influence of task difficulty, since the more difficult task was different for each age group (word for young, color for old). However, we found that task difficulty nor task itself had an effect on cue fixations in Experiment 2, and thus the remaining analyses we disregarded task.

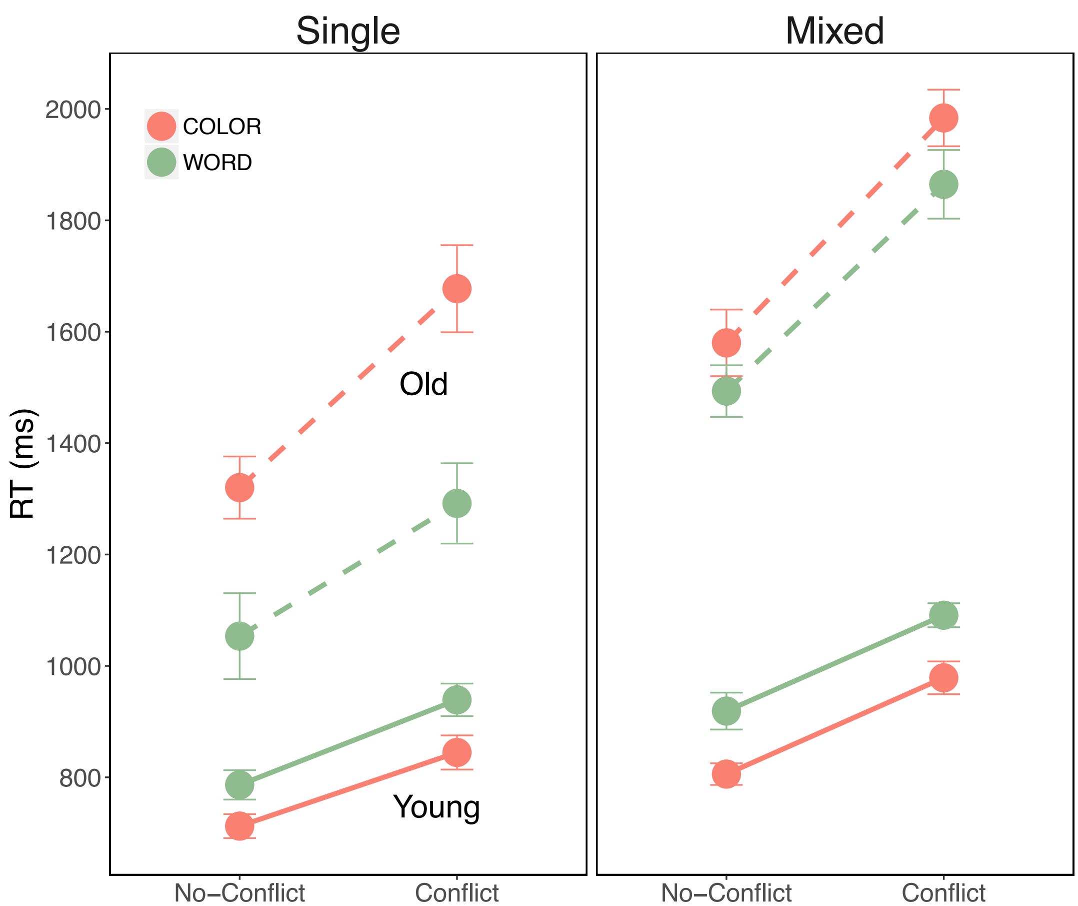

Figure 5. Mean RTs for Experiment 2. Error bars indicate within-subject standard errors computed within each age group (Cousineau, 2005).

Global Effects

As a first step, we examined the amount of unnecessary cue fixations as a result of the global task context. Previous work has shown that more global aspects of the task (for instance, comparing single-task and task-switch blocks) can be dissociable from more local task demands (switch costs), and that the former is more sensitive to age differences (Mayr et al., 1996; Mayr, 2001; Kray & Lindenberger, 2000). This previous work compared performance between a single-task and a task-switch contexts and found that older adults showed greater global costs. This is presumably a consequence of an over-reliance on set-updating, which is particularly costly for single-task blocks (c.f., Mayr, 2001). Note that in the mixed blocks, younger adults are also expected to adopt more of an updating mode in order to deal with the task switches, but in a manner that is closely tailored to the probability of switching (c.f. Mayr, Kuhns, & Rieter, 2013). Furthermore, younger adults should revert to a stable, maintenance state in the single-task blocks while older adults will be less prone to do so (Spieler et al., 2006). Thus we would expect more cue fixations in older adults overall, plus a greater increase of fixations in the mixed compared to single-task blocks. Accordingly, for both experiments we used repeated-measures ANOVA to compare the proportion of unnecessary cue fixations between the mixed and single-task blocks, and in Experiment 2 also tested for an interaction with age group. Note that Experiment 1 had a between-subjects factor that manipulated the occurrence of the beep trials (either 1 out of 8 or 1 out of 12 trials), but we found that this factor had no significant influence on cue fixations across tasks or block types ($p > .37$), and thus ignored it in subsequent analyses. Figure 6 shows the mean proportion of cue fixations by task and block type in each experiment. As evident in Figure 6, there were indeed more cue fixations in the mixed blocks in Experiment 1, $F(1,27) = 4.84$, $p < .05$, *ges* = .07. In Experiment 2, both age groups show increased cue fixations on mixed relative to single-task blocks $F(1,54) = 20.9$, $p < .001$, *ges* = .10, and as predicted, older adults showed a much greater increase in the mixed blocks $F(1,54) = 8.84$, $p < .001$, ges = .044. Notably, these are cue fixations that do not occur on the beep trials, so it cannot simply be explained by participants following the instructions, and it cannot be explained by the presence of the beeps themselves, as the single task blocks included them as well (but were completely irrelevant). Thus the mixed blocks were successful in inducing an updating state in both groups and across both experiments, but as expected, older adults were disproportionately sensitive to this change.

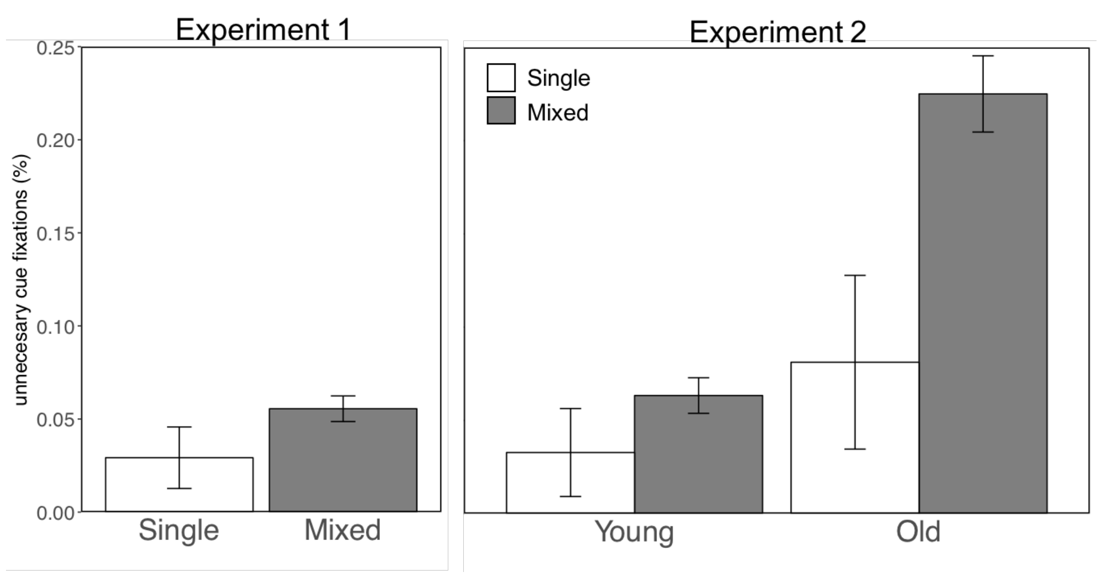

Figure 6. Mean cue fixations by block type for Experiment 1 (left) as well as age group for Experiment 2 (right).

Another way we can approach the question of whether older adults are more sensitive to the global effects is by examining (on a qualitative level) the number of cue fixations in mixed blocks as a function of the number of cue fixations in single-task blocks. This has been used in the aging literature in the past to reveal the extent to which older adults show deficits in particular cognitive processes, above and beyond simple slowing (i.e., state-trace analysis; Mayr et al. 1996; Verhaeghen, 2011). For each individual in Experiment 2, we calculated the proportion of cue fixations in the single-task and mixed blocks, and plotted them in relation to each other in Figure 7 (i.e., single-task blocks on the x axis and mixed blocks on the y axis). If there were some common process that affected the amount of cue fixations on single and mixed blocks equally, then we would expect most data points (individuals) to cluster around the diagonal. Similarly, if this process were similar across age groups, then we would see both age groups cluster around the same area of the plot. If, however, there is something unique about the mixed condition, then we would expect the points to cluster above the diagonal, and if this process were different across age groups, the younger and older adults would cluster in different places. As is evident in Figure 7, younger adults show a fairly tight clustering around the diagonal within a relatively restricted range on both block types, while older adults show much more variability, and a greater range in the mixed block. This suggests that, for older adults, there is something unique about the mixed condition that cannot be explained by some common process that affects cue fixations in general. Conversely, for younger adults there appears to be a more consistent process that affects the increase in cue fixations in the mixed blocks.

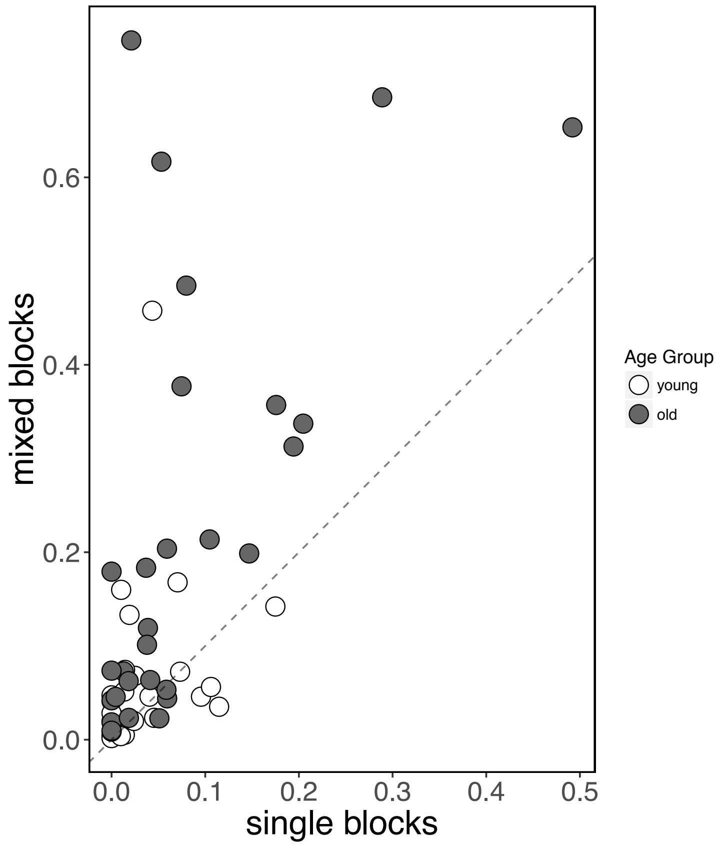

Figure 7. Cue fixations for mixed blocks as a function of single task blocks for all participants in Experiment 2.

Local Effects

Clearly being in a task-switching context has some influence on the tendency to update overall. Next we take a closer look within the mixed blocks and examine the aspects of the local task environment that influence updating in the moment. Possible factors include the task conditions themselves (conflict, task difficulty), one's behavior in the current trial (committing errors), as well as adjustments following particular trial types (trials after conflict, trials after errors). While there was a clear expectation of age differences in the global effects, here we did not have a strong *a priori* expectation regarding specific local effects that may differ across age groups. In order to examine these effects at a finer level, we examined each effect using a logistic mixed effects model with current-trial cue fixation (coded 0 or 1) as the dependent variable, and the effect of interest (e.g., Conflict, Error) coded as a binary predictor. All models included a random intercept for each subject and a random slope for the effect of interest. We excluded beep trials, trials after beeps, and trials with RTs greater than 4000 ms and, except where errors were the effect of interest, error trials and trials following errors. We also excluded trials that followed a cue-fixation trial; as described later, we found that cue fixations tended to happen in streaks, and thus we wanted to capture the initial transition towards an updating state, not the carryover from the previous trial. The same analysis routine was applied to both Experiments 1 and 2, with two exceptions: 1) Due to the dramatic task differences in tasks in Experiment 1 (by design), all models also controlled for task using a main effect (coded exo=1, endo=0) and 2) As Experiment 2 included an age-comparative sample, age group (coded 0 for young, 1 for seniors) was included in each models as well as the interaction with the effect of interest.

The estimates and *p* values for all models of local task effects are depicted in Table 1. For Experiment 2, we include the coefficient for the Age x Effect interaction as well. Note there was also a very robust main effect of age, with older adults exhibiting more cue fixations than young. Since all models had the same dependent variable, the estimate was similar across them (1.46-1.76; all *p*'s <.001) and thus was not included in the table. Similarly, the main effect for Task in Experiment 1 was highly significant and similar across models (.49-.52; all *p*'s <.001). In both experiments, we see there is an effect of current-trial conflict, such that people are more likely to fixate the cues on conflict trials (although it was marginally significant in Experiment 1; Table 1). For trials *following* conflict trials, we see a reduced tendency to fixate the cues, but only for Experiment 2. Furthermore, this interacted with Age, indicating this tendency was greatly reduced in older adults. For trials following errors, there was an increased tendency to fixate the cues in both experiments (Table 1), which again interacted with age in Experiment 2, showing a reduced tendency in older adults. For Experiment 1, we see a decreased tendency to fixate the cues on the harder task (the endo task). However, there are two plausible interpretations of this effect. The fact that the exo task was easier could have allowed the control system to be "lazy", leading to a loss of the task set, and a need to update more often. Alternatively, cue fixations could have been prompted by the spatial distribution of attention in the exo task, which necessitated a broadening of the attentional window, compared to focusing on the central cue in the endo task. Part of the motivation for Experiment 2 was to present all task-relevant stimuli at the center, so that we could examine the effect of task difficulty in particular. We can see in Table 1 that there is no hint of a difficulty effect in Experiment 2. This suggests that the broadening of the attentional window is more likely to have prompted more fixations in Experiment 1. Note that for Experiment 2, the Difficulty factor was coded according to age group (i.e., color task for young adults, word task for older adults). As a follow-up we also ran a model using a Task factor (without recoding), but there again was no effect (all p's > .30). Together the results suggest that for young adults, current-trial conflict can induce an updating state. In Experiment 2, it appears that younger adults are much less likely to update following these trials, suggesting some type of conflict-triggered adjustment. After committing errors, younger adults are also more likely to update, suggesting a loss of task set leading to the error, and a subsequent cue fixation to regain it. In both these cases, older adults are less likely to make these adjustments (since the coefficients are in the opposite direction in each case). This suggests that older adults do not respond as readily to these factors.

Table 1. Summaries of Mixed Effects Models for Local Task Factors

	Experiment 1		**Experiment 2**			
	Effect	*p*	Effect	*p*	Effect x Age	*p*
Conflict	0.32	0.06	**0.31**	**0.01**	-0.25	0.1
Lag-Conflict	-0.02	0.84	**-0.45**	**0.02**	**0.26**	**0.05**
Error	-0.06	0.85	-0.28	0.19	0.38	0.12
Lag-Error	**1.53**	**<.001**	**2.17**	**<.001**	**-1.25**	**<.001**
Difficulty	**-0.52**	**<.01**	0.04	0.81	0.02	0.9

Coefficients and p values are for the effects of interest. Each model for Experiment 1 included a predictor for Task (1= exo task, 0= endo task) not depicted here, with a coefficient that ranged between .49 and .52. Likewise, all models for Experiment 2 included a main effect of Age, with a coefficient between 1.46 and 1.76.

Endogenous Effects

In addition to the results above suggesting that unnecessary updating is induced by particular events in the task context (e.g., conflict, errors), we also wanted to examine the extent to which endogenous factors that operate independently of task context may influence updating. For instance, it is possible that a control state on one trial which leads to a cue fixation has a certain amount of "inertia" that increases the likelihood of fixating the cue on the next trial. This notion that executive control states can carry over across successive trials has been demonstrated in previous work in the domain of task switching (Mayr, et al. 2013; Kikumoto et al., 2016) and conflict adaptation (Hubbard et al., 2016). Thus we examined the extent to which fixating the cue on one trial (indicative of an updating state) influenced cue fixations on the next trial, regardless of the particular task context. First we examine this on a purely descriptive level, by seeing how often cue fixations occur relative to the last beep trial, where such fixations were necessary. If there is some kind of carry-over, then we would expect that a cue fixation prompted on a beep trial may increase the likelihood of fixating the cues (i.e., updating) unnecessarily on subsequent trials. Figure 7 shows the mean proportion of unnecessary cue fixations for successive trials after the last beep trial, separately for single and mixed blocks and (in Experiment 2) age groups. These exclude error trials, trials after errors, and slow responses (RTs > 4000 ms). In this figure, there appears to be a high tendency to fixate the cues on immediate post-beep trials, which declines gradually in subsequent trials. Note that this effect appears to be more dramatic for the older adults in Experiment 2, who also show a greater difference between mixed and single blocks (as established in the global analyses above).

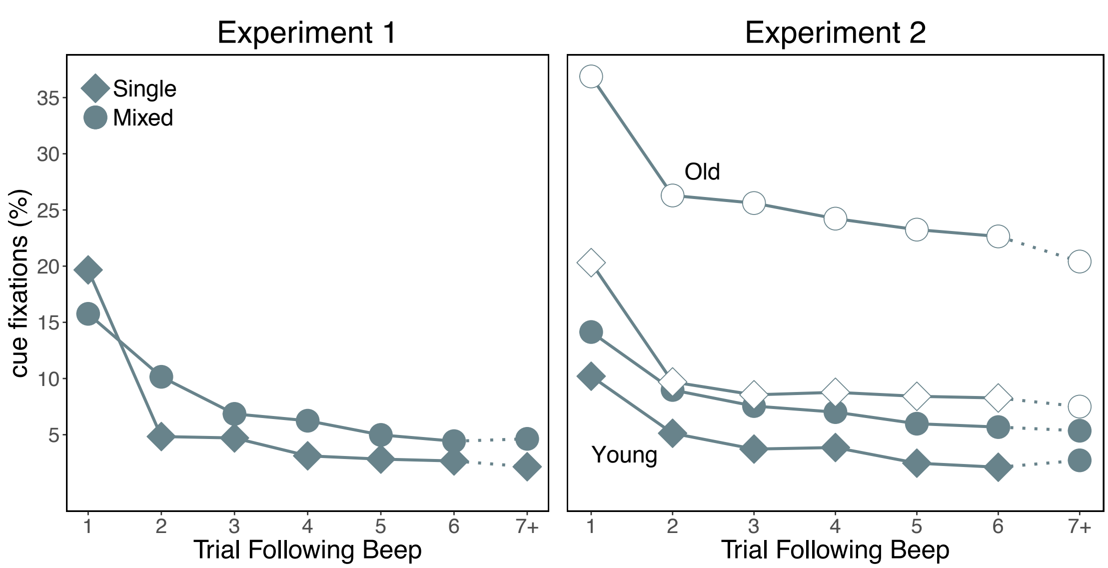

Figure 8. Proportion of cue fixations following beep trials. Unfilled points denote older participants in Experiment 2.

As a next step, we want to determine whether this carry-over effect also translates to unnecessary updating as well, and whether this tendency may be increased in older age. This is difficult to determine from the plots alone, as it is clear that the beep trials do seem to have an influence on subsequent trials. Similar to the local effects, we used multilevel logistic regression to examine current-trial unnecessary cue fixations as a function of whether the cue was fixated on the previous trial. To capture the carryover effect from unnecessary cue fixations in particular, we controlled for whether the current trial followed a beep trial (where the cue fixations *were* necessary), as well as the number of trials since the last beep. We begin with the mixed blocks only. This analysis showed a strong effect of previous-trial cue fixations influencing current-trial fixations in both experiments ($b_{Exp1} = 1.67$, $z = 7.69$, $p < .001$; $b_{Exp2} = .69$, $z = 2.47$, $p < .05$), an effect that also interacted with age in Experiment 2 ($b = .78$, $z = 2.22$, $p < .05$). This was on top of the increased (although non-significant) tendency to fixate the cue on immediate post-beep trials (*p's* > .20), and the tendency for cue fixations to gradually decrease on trials further away from the beep ($b_{Exp1} = -.02$, $z = 2.30$, $p < .05$; $b_{Exp2} = -.04$, $z = 3.35$, $p < .001$). These latter two effects did not interact with age in Experiment 2, however (all *p*'s > .20). Figure 9 illustrates the age interaction from Experiment 2 by showing the model predictions from the analysis above. We can see that, following trials with no cue fixations, there is little age difference in the tendency to fixate the cue on the current trial, while after cue-fixation trials, we see a much greater tendency in older adults. In order to see whether this inertial effect interacts with the global context, we repeated the above analyses, but including both block types, and an additional predictor coding for the mixed block (1=mixed, 0=single). In these models we found similar effects as before, but with a negative interaction between the block type and the carry-over effect (b_{Exp1} = -1.33, z = 7.76, $p < .001$; b_{Exp2} = -.80, z = 5.04, $p < .001$), but no further interaction with Age in Experiment 2 (p = .45). This indicates that with the mixed task context comes a decreased carry-over effect from updating in the previous trial. Given the relatively few fixations in these blocks (Figure 9) and the simpler task context, it is perhaps unsurprising that the inertial effect would carry more influence in these blocks. It is surprising, however, that this does not differ with age, given the overall age differences by global context (Figure 6). Thus across both experiments and age groups, we see that cue fixations (i.e., updating) tend to happen in streaks, where updating on one trial increases the likelihood of updating on the next, an effect that is increased in older adults. This occurs both in response to necessary updating imposed by the experiment, but also in terms of unnecessary updates. This suggests that being in an updating state does not always obey the boundaries of the trial structure imposed by the experimenter.

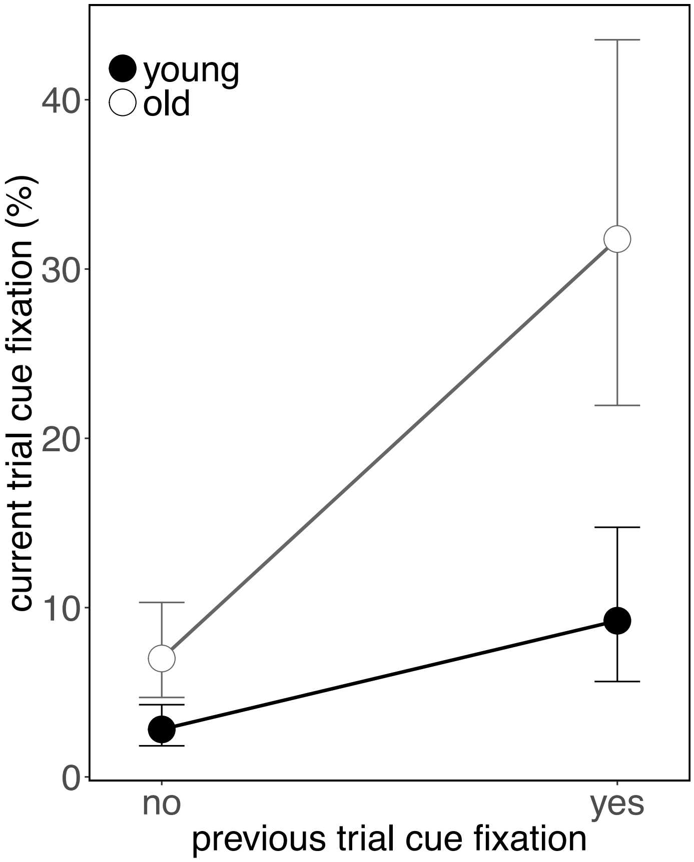

Figure 9. Model predictions examining the carry-over effect in Experiment 2. This analysis was performed on the mixed blocks only. Error bars indicate confidence intervals from the logistic mixed-effects model.

Behavioral Consequences of Cue Fixations

While the findings above suggest that certain global, local, and endogenous factors influence the tendency to update, we must also consider the possibility that these cue fixations (despite being unnecessary) may be a deliberative and strategic part of doing the task. Perhaps fixating the cues is just a conscientious method of remaining on-task, and this tendency is increased on older adults. In fact, there is a consensus in the aging literature suggesting that older adults work to avoid errors to a much stronger degree than younger adults (e.g., Starns & Ratcliff, 2010), as well as an increase in conscientiousness across the lifespan. If this is indeed the case, then cue fixations should be tied to improved performance in some way. Thus fixating the cue on trial *n* may lead to faster responding and better accuracy on trial *n*+1, and perhaps better accuracy on trial *n* as well. If, however, cue fixations are an indicator of updating inappropriately, then we would expect unchanged or even *worse* performance on the next trial. Accordingly, we examined the behavioral consequences of fixating the cues both within the current trial and in the next trial using linear mixed effects models (again, within the mixed blocks only). First we examined the effect on response time (RT) on trials immediately following cue-fixation trials. Importantly we exclude beep trials as well as trials where the cue was fixated, since these trials are almost guaranteed to have longer RTs, and we control for Task and Conflict, as they show clear RT differences as well (Figures 4 and 5). If fixating the cue leads to a refreshing of the task set that leads to more effective processing on the next trial, then we would expect faster responses following cue-fixation trials. However, across both experiments we see exactly the opposite, where fixating the cue on the previous trial leads to slower responses on the following trial ($b_{Exp1} = 53$, $t = 3.51$; $b_{Exp2} = 87$, $t = 3.93$). This result is much more in line with a model of passive carry-over of the control state between trials (Mayr et al., 2013; Kikumoto et al., 2016; Hubbard, et al., 2016). This finding aligns well with the results of the cue fixations, which appear to happen in streaks, but importantly this analysis excludes any trials where a cue fixation occurs, suggesting that performance still suffers following a string of cue-fixation trials. This, however, seems to be a general effect that does not interact with age ($t = 1.14$), so there is no evidence that it provides any additional costs or benefits to older adults. Again, this goes against the explanation that increased cue fixations in older adults arise from those participants being more conscientious. Additionally, there is no evidence that cue fixations influence error rates in the current (as reported above) and subsequent trials (all p's > .21).

Confirmatory Cue Fixations

In the analyses above, we index an updating state through unnecessary cue fixations. What we have not examined thus far is how attention is directed during this state. We have shown in previous work that updating is not necessarily characterized by undirected distraction, but that attention is biased towards specific task-relevant features (Mayr, Kuhns, & Hubbard, 2014). In the present context, we can look not only at cue fixations, but *which* cue is fixated. Since the relevant task was unambiguous non-beep trials, and the location of each task cue (*COLOR/WORD or CENTER/SURROUND*) was fixed on each block, we can distinguish cases where participants fixate the task-relevant cue, from cases where they fixate the irrelevant cue. If the updating state is associated with general distraction, or a complete loss in the task set, then we would not expect any systematic bias towards either cue. Accordingly, among all the cue-fixation trials, for each participant we calculated the proportion of those that went towards the relevant versus irrelevant cue, and examined the difference using logistic mixed-effects models with a binary predictor coding for task-relevant cues (1=relevant, 0=irrelevant). Note that this is specifically addressing the question, *given a cue fixation, which cue is fixated?* This is different from asking how many fixations occur towards one item or another, compared to all other trials. Given the task differences in cue fixations in Experiment, we also included the main effect and interaction with Task (coded 1=exo, 0=endo). This yielded a significant Task x cue-relevance interaction (*b* = 1.80, *z* = 2.14, *p* = .032). As evident in the model predictions in Figure 10, we see a bias towards the relevant cue, but only in the exo task. The endo task shows a numerical, but non-significant bias. This model was repeated for Experiment 2, which included an interaction with Task as well as Age. This model yielded a highly significant bias towards the relevant cue (*b* = .52, *t* = 8.49), but no interactions with Task or Age (all *t*'s < 1.50). Figure 11 depicts the model predictions separately for each age group, but for simplicity does not separate by task. We see that for both age groups, there is a considerable bias towards the task-relevant cue, to a similar degree as the exo task in Experiment 1.

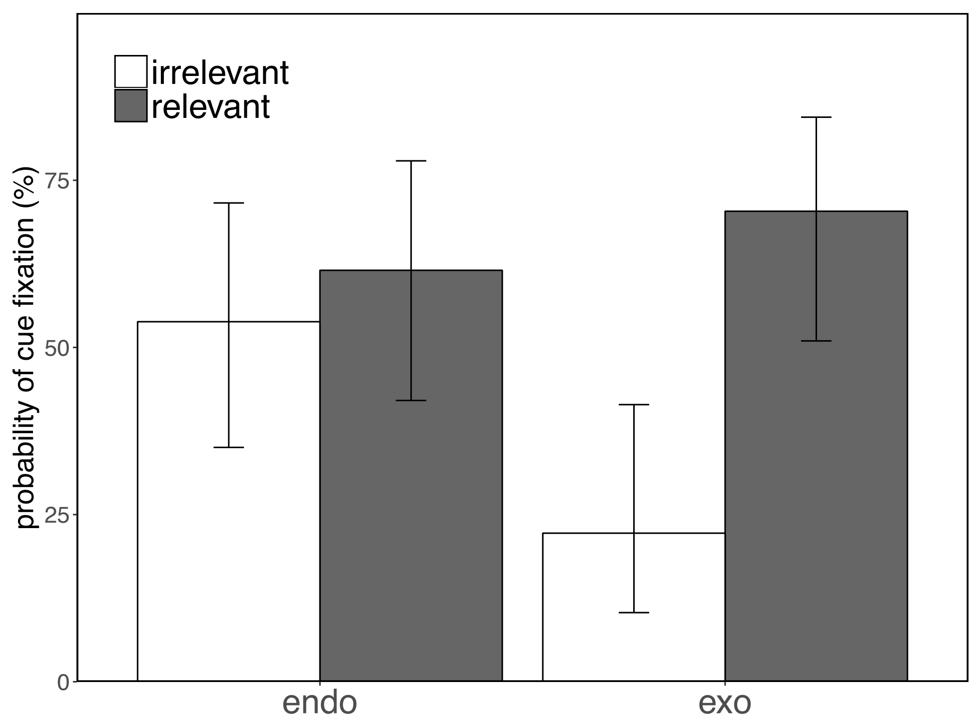

Figure 10. Proportions of relevant versus irrelevant cue fixations for Experiment 1. Values are model predictions from a logistic mixed-effects model. Error bars indicate 95% confidence intervals.

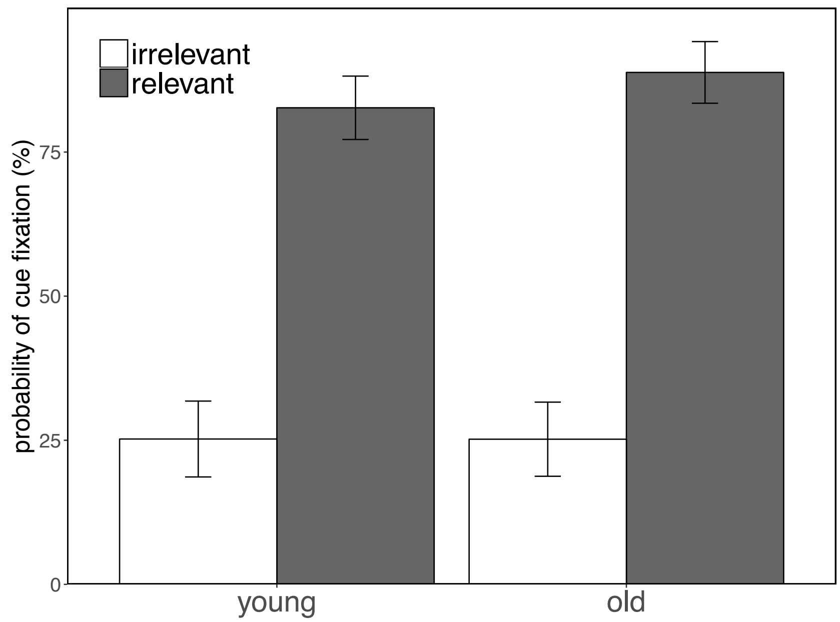

Figure 11. Proportion of relevant versus irrelevant cue fixations in Experiment 2. Values are model predictions from a logistic mixed-effects model including an interaction with Task and Age, but plotted disregarding task. Error bars denote 95% confidence intervals.

Discussion

In the present investigation, we sought to characterize the factors that influence shifts in global control states in a complex task-switching scenario. Previous work has posited the existence of at least two broad processing states—one stable maintenance state, characterized by robust maintenance of task-relevant information and another flexible updating state that brings in new information in the environment. Recent work has also suggested that many age differences observed across many tasks may be attributed to a chronic updating state—in other words, a greater reliance on information in the environment, even when that information is redundant (Rogers et al., 2000). This idea has been put forward in terms of the broad strokes, but with relatively little empirical work directly testing the assertion. The main drive of the present work was to establish the antecedents that lead to shifts towards the updating state, and to characterize where age differences emerge. In the process, we also illuminated some of the dynamics of these shifts.

Specifically, we used fixations towards the peripheral cues in an unambiguous task context as an index of an updating state, and focused on three potential levels that may drive them: 1) global shifts in task context (i.e., mixed versus single-task blocks) 2) local task factors that change on a trial-by-trial basis (e.g., conflict, errors) and 3) endogenous factors, which may operate independently of the specific context (i.e., passive carry-over). Broadly, we found in both experiments that updating in younger adults was driven largely by the local task effects, particularly following conflict trials and error trials. This suggests a more transient process that is triggered by rapid shifts in the task. This was in addition to endogenous factors that drove cue fixations, namely, passive carryover from one trial to the next. Thus updating on one trial increased the likelihood of updating on the next, a pattern that is in line with previous work in task switching and conflict adaptation (Kikumoto et al., 2016; Hubbard et al., 2016). Older adults showed a somewhat different pattern, being much less likely to update in response to the local task demands, but an increased tendency to update based on the global context (i.e., the difference between single and mixed-task blocks). Older adults also showed an increased tendency towards passive carry-over across trials. The global effects are indeed consistent with the chronic updating account of aging, with older adults being more environment-bound (Spieler, et al., 2006; Lindenberger & Mayr, 2014). The result that older adults have an increased tendency towards passive carryover is more difficult to interpret. This pattern could emerge simply as a consequence of older adults being in the updating state more often, not necessarily a result of the "inertia" of the control state carrying over across successive trials. The latter interpretation would suggest that not only are older adults more likely to update, but that regardless of what state they are in, those states are “sticky” and resistant to rapid change. It is also possible that similar effects do indeed drive younger and older adults towards an updating state, but that the threshold is much lower for older adults. Thus, the reduced local effects are a consequence of the older adults being triggered into the updating state more easily, and thus less of a sharp contrast between trial types (e.g., conflict and no-conflict). This would also lead to longer streaks of cue fixations in the older adults. Yet another possibility is a general tendency towards broader temporal construal in older adults. It is possible that the older adults are more concerned with the long-term aspects of the task, which in this case would coincide with mixed versus single-task blocks, and are simply less attentive to the faster moment-to-moment changes at the level of individual trials. Ideally, these possibilities could be explored in a paradigm in a more continuous task that does not have such abrupt changes between trials (c.f., Esterman, Noonan, Rosenberg, & DeGutis, 2012) which would allow one to distinguish shifts in the updating state that are not necessarily bound by the task structure. This would also allow one to impose gradual long and short-term shifts in the task context to specifically test the temporal construal explanation.

Are cue fixations really an index of updating?

Here we treat unnecessary cue fixations as a signal that executive control system is in an updating state, one that is biased towards refreshing information from the environment. Others may interpret these fixations as simple double-checking, a deliberate act seeking to confirm information that the participant already has access to. In fact, analysis of the cue fixations indicates that participants in both age groups are more likely to fixate the relevant task cue. This suggests that the cue fixations really are unnecessary. Previous work in aging has also shown us that older adults are in general more conscientious (Roberts, Walton, & Viechtbauer, 2006), and have an exaggerated tendency to avoid errors in laboratory tasks (Starns & Ratcliff, 2010; Ratcliff et al., 2004). Thus cue fixations could just be strategic attempts to avoid errors. Importantly, in both experiments we examined the consequences of updating, to rule out the notion that it was a sort of speed-accuracy tradeoff that led to better performance. If anything, the results were more consistent with a poorly-optimized control state that led to poorer performance (in terms of RTs), which carried over into the next trial (Kikumoto et al., 2016; Hubbard et al., 2016; Mayr et al., 2013). These effects speak against a simple speed-accuracy tradeoff. However, given that the cue fixations were not predictive of subsequent errors indicates that there was little cost to the older adults for making these eye movements, beyond a slowdown in response time. This leaves open the possibility that the cue fixations did not help performance significantly, but gave the older participants the sense that they were avoiding errors. Since error rates were not affected one way or another, this is difficult to address in the present investigation. Some may also suggest that cue fixations are more an indicator or monitoring or vigilance, rather than the specific attempt to update. The fact that the eye movements indicate knowledge of the task are in line with this explanation, as is the relatively small cost associated with the cue fixations. Notably, this pattern also converges with literature on reading behavior, which shows that in many ways the eyes follow predictable patterns that operate independently of the intentions of the reader (e.g., Engbert, Nuthmann, Richter, & Kliegl, 2005). This suggests that there is some underlying process responsible for planning eye movements and sampling from the environment that is not necessarily under volitional control. The findings here could reflect a similar process, which samples information from the task cues even though the participant herself may know exactly what they need to do. If this is indeed the case, the question is whether this kind of sampling occurs as a result of a global updating state, or operates completely independently. It is possible that this more automatic sampling may reflect an attempt to update, but without actually updating the internal representations, something more akin to monitoring. It is possible that the cue fixations observed here may be a mixture of this automatic process and more deliberate updates, making it difficult the disentangle the influence of each.

Whether the cue fixations are driven by voluntary updating or a more automatic process is still and outstanding question. At the broadest level, updating can be seen as simply an attempt to gather information from the environment—whether this is driven by a high-level intention to confirm information or a low-level process out of voluntary control is still yet to be determined. In either case, it does not change the pattern that older adults do it more—the question then focuses on how well older adults can overcome that tendency when asked. So far, the evidence suggests that as long as information is in the environment, older adults have difficulty avoiding it, but can nonetheless perform well when that information is absent (e.g., Spieler et al., 2006). This suggests a more automatic process, but does not necessarily rule out strategic effects either. Again, the pattern could be attributed to monitoring or vigilance instead of updating per se. These questions cannot be answered in the present investigation since cues were present on all trials, and we cannot compare performance in the presence or absence of cues. A thorough examination of this question will take careful task design, and would most likely require some measure of updating from neuroimaging. This would allow one to identify updating in a manner independent of overt behavior, and thus speaks against intentional, strategic factors. To our knowledge, however, such an index has yet been discovered using human neuroimaging.

Possible neural mechanisms

Previous computational and experimental work established a strong case that these broad processing states do exist, with most accounts attributing them to the dopamine or noradrenergic systems. Computational modeling of working memory first introduced the idea that the cognitive system switches between a protected state, which can be updated with the opening of a gate that allows in information from the environment (O'Reilly, 2006). These states are largely driven by dopamine-mediated interactions between the prefrontal cortex and the basal ganglia. Coming from work in animal models, other theoretical work has suggested that the relative dominance of different classes of dopamine receptors can lead to "D1" or "D2" states, that are characterized by stability and flexibility, respectively (Durstewitz & Seamans, 2008). They suggest that this will drive differences both within and across individuals. Others have suggested that tonic activity in the locus coeruleus norepinephrine (LC-NE) system can drive the brain towards a flexible "exploration" state, while phasic bursts within that system can lead to a focused "exploitation" state through an adaptive gain mechanism (Aston-Jones & Cohen, 2005). This account conveniently has the added benefit that the

LC-NE system can be indexed via fluctuations in pupil diameter (Aston-Jones & Cohen, 2005; Nieuwenhuis & Jepma, 2011). For instance, others have found that changes in tonic pupil diameter can predict exploratory choices in a n-armed bandit task (Jepma & Nieuwenhuis, 2011; Gilzenrat et al., 2010). We examined pupil diameter in the present task and found a general pattern of increased pupil diameter in trials approaching cue fixation trials, but these effects were weak and difficult to interpret due to the rapid nature of the tasks, which were not optimized for pupillometry (which is has a sluggish response of ~ 1 second). We do not have strong opinions about which particular neural mechanism gives rise to these states, and indeed it has been suggested that the dopaminergic and noradrenergic systems may activate each other's receptors (Sara, 2009). The age-related bias towards updating, and the evidence that dopamine may play a role converges well with evidence that the dopamine system is systematically depleted with older age (Bäckman et al., 2006). Thus with increasing age the downregulation of dopamine may tip overall brain dynamics towards an updating state. Again, to provide strong support for this account, one would need a reliable neural index of updating to distinguish it from more strategic factors.

Bridge

In Chapters II and III, we examined how updating occurs through examination of RT costs in relation to task events (e.g., following interruptions), and fixations towards relevant versus irrelevant items on-screen (e.g., peripheral cues). These, however, are indirect measures of what we actually want to capture, which is the actual representation of the task context. Thus in Chapter IV, we developed a paradigm similar to that in Chapter II and in previous work in our lab, but tailored to electroencephalography (EEG), a measure of electrical activity in the brain. By using these measurements combined with machine-learning techniques, we attempt to extract abstract information about the task context as it emerges in time. As a first step, this needs to be done in a very controlled context, without the additional manipulations associated with updating (e.g., interruptions). By first establishing that the information can be extracted, in future work we can extend it to more directly address questions regarding the updating and maintenance states.

CHAPTER IV

DECODING THE DYAMICS OF ABSTRACT ATTENTIONAL SETTINGS IN TASK SWITCHING

This work was co-designed with Atsushi Kikumoto, who also provided technical background regarding EEG data collection and analysis and will be a co-author on the paper. The main text below (but not the extended methods) is co-authored with Ulrich Mayr.

Since Wilhelm Wundt, it has been an important, though somewhat elusive goal to trace the cascade of information processing in humans. We show that all relevant representations within a task-switching situation—of stimuli, responses, and even abstract attentional settings—can be extracted with high temporal resolution from scalp electrophysiological signals (EEG). Importantly, with this method, we can characterize the temporal dynamics of abstract attentional settings, and probe their relevance for fluent performance.

Existing methods to infer mental processing components, such as chronometric analyses of response-time (RT) patterns, the analysis of Evoked EEG components (ERPs), or fMRI BOLD signal, provide only limited information about when and how specific representations are used in the service of goal-directed action. Also, these techniques are of limited value when trying to explain performance on the level of individual trials—either because they rely on differences between conditions, or because of the level of inherent noise. Moreover, it is of particular theoretical importance to characterize the dynamic behavior of those abstract representations that control “lower-level” stimulus or response-related codes. Yet, because such codes are not tied to specific stimuli or responses, they are particularly difficult to pin down with existing methods.

We used a standard task-switching paradigm (see Figures 1A and 4). On each trial, participants (1) encoded an auditory task cue (with two possible cues per task) in order to (2) establish one of two possible task sets (color vs. orientation), which then are used to (3) localize the task-relevant object among a set of eight objects, (4) ignoring the currently irrelevant object, and (5) translate the response-related stimulus feature (e.g., a specific color) into an actual response (e.g., left versus right key). Via a linear classifier, we examined to what degree the spatial power distribution in four bands of the EEG frequency spectrum (i.e., delta: 2-3Hz, theta: 4-7Hz, alpha: 8-12Hz, beta: 13-31Hz) expressed information about each of these five aspects, using trial-by-trial and millisecond-by-millisecond data.

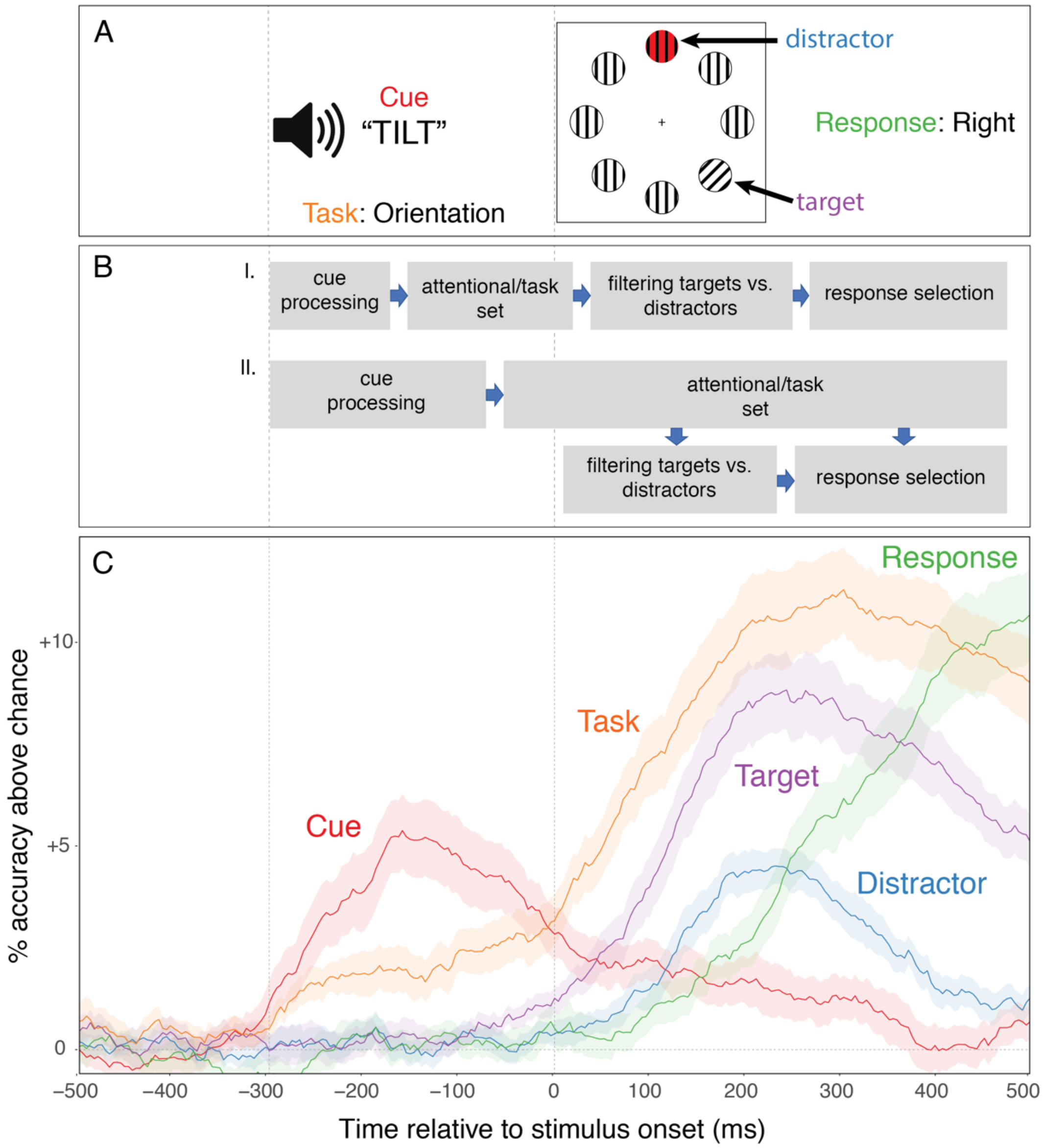

Figure 1. Timeline of task sequence and decoding results. Dotted lines indicate cue (-300 ms) and stimulus onset (0 ms) in all three panels. **A.** Stimulus sequence with each to-be-decoded task element is highlighted. **B.** Timeline of competing models, one with serial progression from the attentional set to stimulus and response processing (I), and a simultaneous model of high-level and low-level representations (II). C. Decoding accuracy of each task element across time, normalized to the same chance level.

Previous work has shown that the locus of spatial attention and working memory can be decoded from the EEG signal[1]. However, so far we know little about how the time course with which locations of task-relevant (target) and task-irrelevant (distractor) information can be distinguished. Eye-tracking based estimates from task-switching experiments have indicated a surprisingly long (about 300 ms) initial period of indifference between target and distractor information[2]. Further, we are particularly interested in the degree to which we can decode abstract attentional-set representations, and if so, how these representations behave relative to lower-level codes that represent cues, stimuli, or responses. For example, according to one prominent model, people may not actually rely on abstract attentional/task settings, but instead use conjunctions of cue and stimulus representations to resolve ambiguity between competing tasks[3]. A further, unresolved question is to what degree attentional sets are activated prior to lower-level processes in order to tune these towards goal-relevant information, or whether they are activated in parallel with lower-level representations, "molding" goal-relevant processing as it occurs (see Figure 1B).

Figure 1C shows how decoding accuracy unfolds for the different stimulus aspects. The auditory task cues are represented only initially, but become much less important once the stimulus appears—a result that is inconsistent with the cue-stimulus conjunction model[3]. In contrast, in line with the notion that the cue is used to retrieve the attentional setting[4], decoding of the attentional set (task) begins during the prestimulus period. Decoding of the attention set increases once the stimulus appears and remains strong throughout the post-stimulus period. In additional analyses, we verified that decoding of the attentional set generalized across features/responses, thus ensuring the abstract nature of the represented information (Figure 6). In parallel to the attentional set we can also decode the target position, and to a lesser degree of the distractor position—both peaking between 200 and 300 ms. Strikingly, the difference between target and distractor location emerged much earlier than the approximately 300 ms time point suggested by previous eye-tracking evidence[2]. In parallel with target processing, decoding of the feature value and/or response code (which cannot be dissociated in the current desing, but see SM) ramps up and peaks near 500 ms.

Decoding was performed on a trial-by-trial manner. Therefore, we can use an additional output, the posterior probability, as an indicator of the classifier confidence to examine to what degree RT variability is explained by lower-level codes (i.e., target/distractor position or feature/response), or by the more abstract, attentional-set representation (see SM for more detail). As shown in Figure 2A, beyond decoding accuracy of the target position and of the feature/response, the attentional set serves as a powerful predictor: The more strongly it is expressed on a given trial, the faster the RT. Figure 3 also demonstrates the strength of the relationship between the decodability of the different codes and performance (slow vs. fast RTs), but also how switching of attentional sets affects representations. Specifically, the attentional set representation was weakened, but not delayed on switch trials.

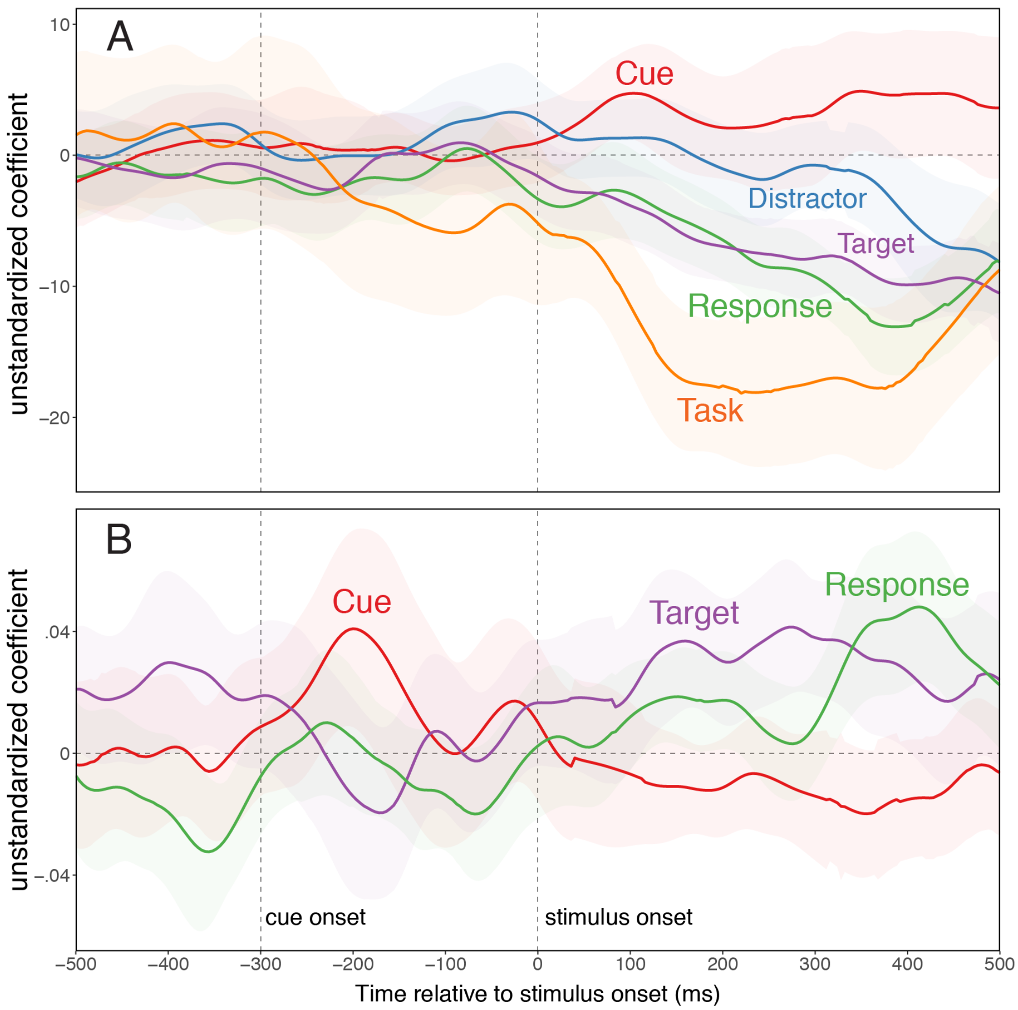

Figure 2. Relating trial-by-trial classifier confidence for each element to reaction time (A), and to task-related evidence (B). Lines denote estimates from mixed-effects models, repeated for each task element and time point. Bands indicate 95% confidence intervals around the estimates.

As indicated in Figure 1C, the task representation can be decoded in the prestimulus phase, suggesting some degree of pre-stimulus preparation over and above the representation of the superficial task cue. However, most of the predictive power arises only after the stimulus is presented, and in parallel with the other lower-level codes. This suggests that the attentional setting promotes simultaneous, task-specific processing. In fact, when examining relationships between decodability of the task and the lower-level features (Figure 2B), we see that in the post-stimulus phase, task-set decodability is initially related to target position decodability, after which also a relationship with feature/response decodability emerges. While this evidence allows no conclusions about causal direction, it is consistent with the view that attentional sets mold simultaneous lower-level processing in a goal-appropriate manner.

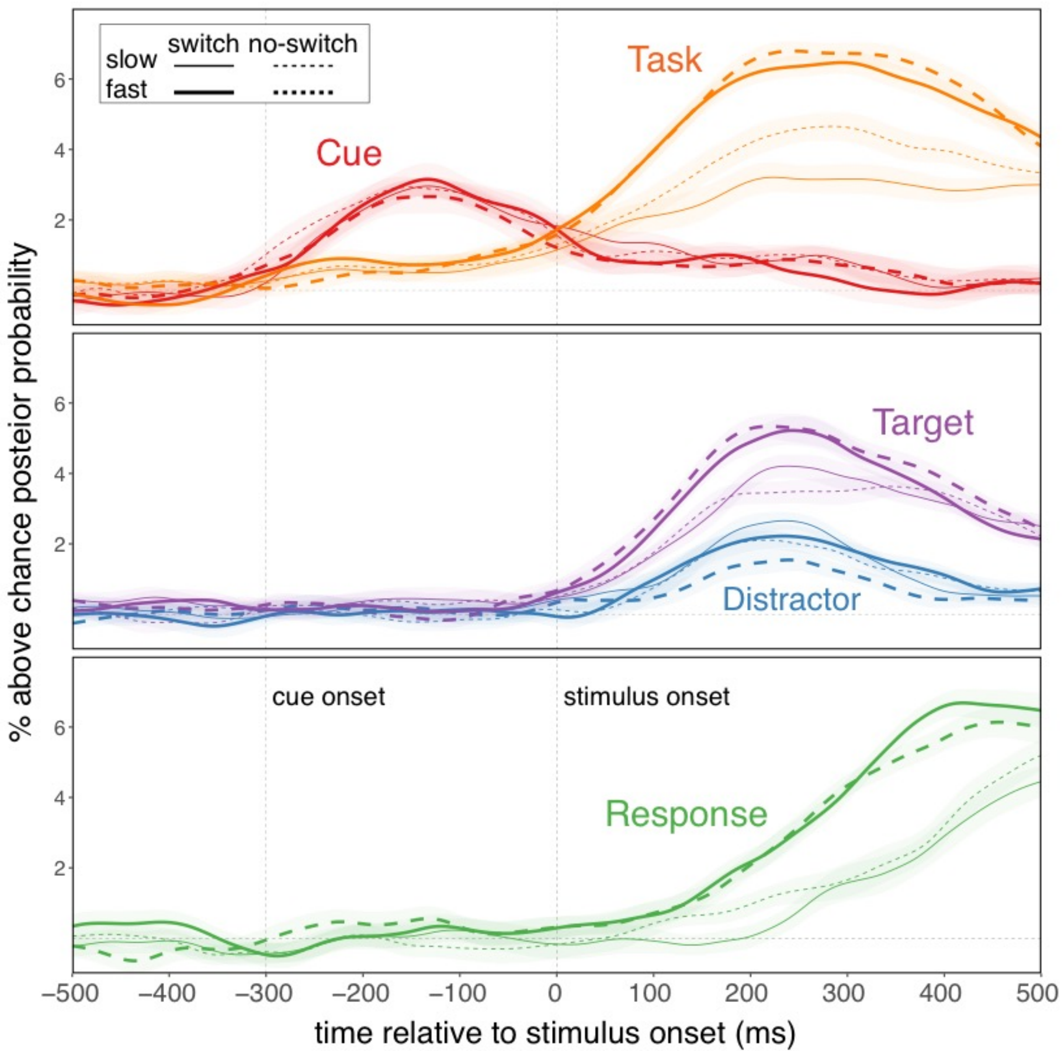

Figure 3. Average classifier confidence over time, separated by fast vs. slow RTs (determined via median split) and switch vs. no-switch trials. Task elements are separated into three panels for clarity. Bands indicate within-subject confidence intervals calculated at each time point and fast/slow condition. Notably, most elements show a fast/slow difference, and task evidence selectively shows a switch effect in slow trials.

Using data-driven, EEG decoding analyses we clarified the temporal dynamics of both lower-level stimulus/response codes and abstract, attentional settings. We found that processing of task-relevant stimulus information emerged much earlier than suggested by previous research. We also found that the strength of the abstract attentional set representation (rather than of superficial cue information) promotes fluent performance, and that it seems to do so by shaping simultaneously active, lower-level representations. We believe that these, and related methods have enormous potential for uncovering the temporal dynamics of hidden representations.

Extended Methods

Participants

A total of $n = 22$ participants from the surrounding community in Eugene, OR participated in this experiment. They were compensated at a rate of $10 per hour, with additional incentives based on performance on the task (explained below).

Procedure

The basic stimulus procedure followed a cued task switching paradigm (Rogers & Monsell, 1995) and followed the general setup of paradigms we have used previously with eye tracking (Mayr, Kuhns, & Reiter, 2013; Kikumoto, Hubbard, & Mayr, 2015). On each trial, an auditory cue indicated which of 2 tasks the participant had to complete, and upon stimulus onset, participants had to covertly attend to a unique stimulus among an array. The array consisted of 8 circular gratings (diameter of each ~ 2.4 degrees) in a larger circular arrangement (diameter ~ 12.5 degrees). Most of the stimuli were vertical, black and white gratings that should always be ignored. Each task involved attending to an item in the array that differed on a single dimension. For the Color task, one needed to pay attention to the only colored (vertical) grating, which could be two different shades of reddish-orange. For the Orientation task, one attended to the black and white grating that was tilted 30 degrees either to the left or right. Participants then made a speeded response using a left or right key on the keyboard ("z" and "?", respectively) depending on the unique item that was presented. For the Color task, if the item was in the more yellow color, they pressed the left (z) key, and if it was the redder color they pressed the right (?) key. For the Orientation task, if it was tilted to the left, they pressed the left (z) key, and if it was tilted to the right they pressed the right (?) key. Both potentially task-relevant items were present on every trial, so it was crucial to pay attention to the cue before stimulus onset. It has been previously established that a substantial portion of task-switch costs can be attributed to cue switches instead of task switches *per se* (Mayr, 2006). One way to eliminate this possibility is to pair each task with two different cues and alternate those cues across trials, resulting in cue switches on every trial. Accordingly, each task was paired with two auditory cues: “color” or “hue” for the Color task, and “tilt” or “lean” for the Orientation task. The audio files were generated from the built-in text-to-speech in Mac OS X 10.10, then edited so that their duration was exactly 300 ms.

The stimulus sequence is depicted in Figure 4. Each trial began with a 700 ms prestimulus interval with a fixation cross in the center of the screen. The auditory cue was presented in the last 300 ms of this interval, so the stimulus array appeared as soon as the cue completed. On every trial, the stimulus array consisted of 6 vertical black-and-white gratings, one colored vertical grating (either reddish or yellowish), and one tilted black-and-white grating (tilted 30 degrees to the left or right). Participants were seated approximately 70 cm from the screen, and instructed to keep their eyes at fixation and not blink throughout the trial; trials containing eye movements were detected via electrooculogram (EOG) and excluded from the analysis. This is important, as eye movements and blinks contribute considerable noise to the EEG signal. Thus any ability to detect a signal relating to the stimulus position is a result of covert attention and not something to do with the eye movements. After identifying the correct item, participants responded using the index finger of their left or right hand. They were instructed to respond as quickly and accurately as possible. The stimuli remained on screen until a response was made. If they made a mistake, an error tone was emitted for 100 ms. Lastly, there was a jittered inter-trial interval (ITI) between 750 and 937 ms, where participants were allowed to blink before the next trial began. The experiment began with 2 practice single-task blocks of 20 trials each (one for each task, order counter-balanced), followed by a practice block containing task switches, then 22 blocks of 64 trials each. In order to incentivize them to respond quickly and accurately, they were rewarded a small amount (0.5 cents) for each trial where they were faster than the 75% percentile of their previous RT distribution, but only if they maintained at least 90% accuracy for a given block. At the end of each block they were given feedback regarding their average RT and accuracy for that block. The RT distribution was determined separately for each task and switch condition after the first mixed-task block, and updated with each trial. Task switches occurred on 50% of trials, and the location of the cued item (target) and irrelevant item (distractor) was randomly determined on each trial, with the constraint that they occurred along the different parts of the array with roughly equal frequency. Since putting this constraint on the 8 unique positions would result in too many conditions to balance, the array was broken into 4 bins (2 positions per bin) and targets and distractors were constrained to occupy each bin with roughly equal frequency. The particular stimulus for each task (e.g., left-tilt or right-tilt for the Orientation task) was randomly determined on each trial.

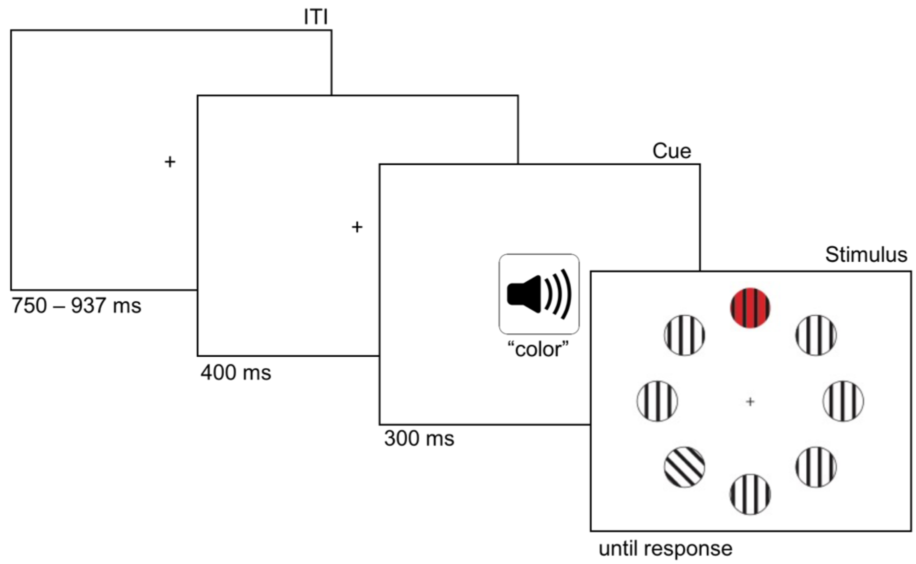

Figure 4. Trial sequence. After a jittered ITI and 400ms fixation, an auditory cue indicated the task (the speaker image was not presented), followed immediately by the stimulus onset. Every trial contained both a target-relevant and irrelevant stimulus.

EEG Recording and Preprocessing

Electroencephalographic (EEG) activity was recorded from 20 tin electrodes held in place by an elastic cap (Electrocap International) using the International 10/20 system. The 10/20 sites F3, Fz, F4, T3, C3, CZ, C4, T4, P3, PZ, P4, T5, T6, O1, and O2 were used along with five nonstandard sites: OL midway between T5 and O1; OR midway between T6 and O2; PO3 midway between P3 and OL; PO4 midway between P4 and OR; and POz midway between PO3 and PO4. The left-mastoid was used as reference for all recording sites. Data were re-referenced off-line to the average of all scalp electrodes. Electrodes placed ~1cm to the left and right of the external canthi of each eye recorded horizontal electrooculogram (EOG) to measure horizontal saccades. To detect blinks, vertical EOG was recorded from an electrode placed beneath the left eye and reference to the left mastoid. The EEG and EOG were amplified with an SA Instrumentation amplifier with a bandpass filter of 0.01–80 Hz and were digitized at 250 Hz in LabView 6.1

running on a PC. Preprocessing was performed using the Signal Processing and EEGLAB (Delorme & Makeig, 2004) toolboxes in MATLAB. Trials including blinks (>80uv, window size = 200 ms, window step = 50ms), large eye movements (>1°, window size = 200 ms, window step = 10ms), and blocking of signals (range = -0.0 5uv to 0.05uv, window size = 200 ms) within the interval of -700 to +400 ms relative to the stimulus were rejected and excluded from further analysis, resulting in an average of 180 trials (12.4%) rejected across participants.

Results

Behavior

Of the 22 participants who participated in the experiment, one was excluded for having EEG artifacts in excess of 30% of trials, and one was excluded due to an experimenter error that resulted in data loss, leaving a total of n = 20 for analysis. For all analyses reported, we first excluded practice trials, error trials (3.8% across participants), trials after errors, and trials that were rejected based on the criteria in the EEG preprocessing (e.g., blinks, artifacts, eye movements, explained above). We then submitted the RTs to a repeated-measures ANOVA with factors Task (coded Color = 1, Orientation = 0) and Switch. This yielded a modest, but highly significant switch effect ($M_{no\text{-}switch}$ = 634 ms, M_{switch} = 659 ms; $F(1,19) = 40.7$, $p < .001$), but no effect of Task, ($F(1,19) = .03$, $p = .86$), and only a marginal Task x Switch interaction, $F(1,19) = 3.33$, $p = .08$. The small magnitude of the switch effect is unsurprising, given the relatively long response-stimulus interval (RSI) that ranged between 1450 and 1637 ms. Having such a long interval is known to reduce (but does not eliminate) the switch effect (c.f., Mayr et al., 2013). In designing the paradigm, this fact had to be balanced with the constraint that participants needed sufficient time to blink between trials.

EEG analysis

After initial preprocessing and identification of artifacts, the single-trial EEG data were decomposed into a time-frequency representation via wavelet decomposition. The power spectrum of EEG signal was obtained from fast Fourier transform, which was then convolved with the power spectrum of complex Morlet wavelets, defined by $e^{i2\pi tf} e^{-t2/(2*\sigma 2)}$ where t is time, f is frequency, and σ is the width of each frequency band, set according to $n/2\pi f$ where n increased logarithmically from 3 to 8. This was repeated for the frequency bands between 2 and 31 Hz in logarithmically-spaced steps. The incremental number of wavelet cycles was used to balance between both temporally-based and frequency-based precision (Cohen, 2014). Logarithmic scaling of the frequency bands was used to keep the width across each band approximately equal. This was all performed in the frequency domain, then brought back into the temporal domain using inverse Fourier transform. A frequency band-specific estimate at each time point was defined as the squared magnitude of the convolved signal $Z(\mathrm{real}([z(t)]^2 + \mathrm{imag}[z(t)]2)$ for power, and as arctangent of $Z(\mathrm{imag}[z(t)] / \mathrm{real}([z(t)])$ for phase. Only power was considered for the present investigation, and for simplicity we focused on frequency bands that are most often presented in the literature: delta (2-3 Hz), theta (4-7 Hz), alpha (8-12 Hz), and beta (13-31 Hz). For each frequency band, we averaged the power signal across the range of interest.

Decoding Procedure

Performing the wavelet convolution as described above resulted in an estimate of power of each frequency band for each of the 20 electrodes across time. For each trial, we extracted a window centered around stimulus onset, starting 500 ms before and extending 500 ms after the onset. The end of this interval corresponds to the 30th percentile of the RT distribution, ensuring that at least 70% of trials are still in progress at that point. This is important when decoding information towards the end of the interval, where we want to capture a substantial portion of the data, not just trials that have exceptionally long RTs. Thus we had an estimate of power in a given frequency band (delta/theta/alpha/beta) at a particular point in time, for a given trial for each of the 20 electrodes. Previous work has established that different types of information are encoded in brain oscillations at particular frequency bands (e.g., Buschman et al., 2012; Fries, 2005; Engel, Fries, & Singer, 2001; Foster et al., 2016) which motivated decomposing the raw EEG signal into the separate bands. However, in the present investigation we were agnostic to which bands encode which type of information, and thus concatenated all 4 bands together in the decoding analysis. We also repeated the main decoding analysis using simply the raw EEG signal across the 20 electrodes at each timepoint, and found the results were generally similar, but (as expected) noisier.

We were, however, interested in the evolution of the information in time, and accordingly performed the analyses separately for each timepoint. Thus in the decoding analyses, the features consisted of the estimate of power for each electrode at a single point in time, repeated for each frequency band (20 electrodes x 1 timepoint x 4 bands = 80 features). Prior to decoding, the EEG data were z-scored so that the mean of each trial's data was 0. Thus we examined the extent to which the spatial pattern of the EEG power across the scalp (across the 4 frequency bands) was predictive of each task feature. As discussed in the main text, the features we considered were the auditory cue ("color"/"hue" or "tilt"/"lean"), the task (Color or Orientation), the target position (partitioned into 4 bins, coded 1-4), the distractor position (bins 1-4), and the response (left vs. right). Note that in the current paradigm, we cannot distinguish between the manual response and the unique task stimulus (e.g., left-tilted grating, or reddish grating) as they were completely confounded. Additionally, as we wanted to isolate the discriminability of each feature regardless of any task differences, for each feature, we performed the decoding separately within each task (except, of course, in decoding the task itself). The results from these analyses were then averaged. In both EEG and fMRI, spatial patterns of activation tend to be idiosyncratic across individuals, and thus we performed all analyses separately for each subject and averaged the results across the group (c.f., Foster et al., 2016). The machine learning algorithm we used was L2-regularized logistic regression, as implemented in the scikit-learn package in Python (Pedregosa et al., 2011), with a tolerance of 1×10^{-4} and the inverse of the regularization strength (C) set to 1.0. Note that we repeated the main decoding analyses with naïve Bayes, support vector machines, and random forests and the results were consistent across the different algorithms. Multi-class classification (which was the case for target and distractor positions), was implemented as a series of binary classifications. We used a 4-fold cross-validation procedure where approximately 75% of trials were used in the training set, and the remaining 25% of trials were used as the test set, and this was repeated until each trial had an opportunity to be part of the test set. In the main analysis, we report the average decoding accuracy, averaged within subject, timepoint, and factors of interest (e.g., task), then across subjects. In addition to the decoding accuracy, we also utilized the posterior probabilities generated from the classifier on each trial. Specifically, for each item in the test set, the classifier generates a posterior probability (using the predict_proba function in scikit-learn) for each class, indicating the classifier's confidence that the test observation belongs to each class. The class with the highest posterior thus corresponds to the guess given by the classifier, and the sum of all probabilities for a single test observation sum to 1. Thus for each trial, we can obtain a continuous measure of the classifier confidence that a certain feature (e.g., task, target position) is a particular value (color task, position 4), and this can be repeated for each feature of interest. Importantly, we always retained the posterior associated with the correct answer, regardless of whether the classifier guessed correctly or not. We used these posteriors in two ways in order to relate them to participants' performance on the task. First, we averaged them in a similar manner as the classification accuracy, but within factors that were related to performance but unknown to the classifier (e.g., switch trials, or fast responses, Figure 3). Secondly, we used these trial-by-trial indicators of classifier confidence as predictors in linear mixed-effects models in order to predict response time (RT) across trials, described in more detail below.

Prediction Analyses

In the first phase of the analysis, we examine the extent to which we can recover information about the task context from the pattern of the EEG power spectra. In the prediction analyses, we utilize the posterior probabilities (classifier confidence) from the first phase and examine the extent to which the information recovered is predictive of trial-by-trial performance on the task. Importantly, when training the classifier, we give no information other the the labels of the factor being decoded (cue/task/target/distractor/response). While the decoding accuracy alone is persuasive, particularly when considering the emergence of difference pieces of information across time, it is still difficult to interpret exactly what is recovered from the EEG signal. A more convincing case is made if that information then relates to performance on the task itself, or distinguishes between different trial types unknown to the classifier. This suggests that we are capturing something about the extent to which participants are representing the given task feature. Accordingly, after performing the decoding analyses as discussed above, we retained the posterior probabilities for each trial as it is included in the test set. This is repeated for each feature of interest: cue, task, target position, distractor position, and response. Specifically, we retained the probability associated with the correct class, regardless of the guess from the classifier. Thus we include both probabilities where the classifier had high confidence and got the answer correct, as well as trials where the classifier had low confidence in the correct answer (i.e., misclassifications). This provides a relatively unbiased method of examining the extent to which the information recovered via the decoding analysis relates to task performance; if the classifier is incorrect on a particular trial, it may be because the participant did not have a robust representation of the task context, which would then lead to reduced performance (i.e., slow RTs). If the classifier confidence is related to some contingency in the task itself and not the participants' representation of the context, then it is not likely to relate to performance in a systematic manner. The probabilities for each of the 5 task features for each trial were first logit transformed, since they were not normally distributed, then used as predictors in linear mixed-effects models predicting RT, with a random intercept and random slopes for each predictor (unless otherwise noted) for each subject. These analyses were carried out using the lme4 package in R (Bates et al., 2014). This procedure was performed on a timepoint-by-timepoint level, where a separate model was fit at each time, as well as a trial-by-trial analysis, with posteriors averaged across time for each trial. In the former case, the implicit assumption is that the different task factors can be represented simultaneously, and we are thus accounting for the unique contributions of each towards predicting the outcome (e.g., RT). In the latter case, we instead acknowledge that information emerges at different stages across the trial, and perhaps the strength of each one (at the appropriate point in time) predicts the behavioral outcome. For this analysis, we look at the group-averaged decoding results, and for each feature find the timepoint with the maximum decoding accuracy. We then choose a 150 ms window centered around that point, and average the (logit-transformed) posteriors across that window, separately for each feature. We also used this same general procedure to examine the extent that the different pieces of information are related to each other (Figures 2B and 8), again using linear mixed-effects models, but using the posteriors for one feature (e.g., task) to predict another (e.g., target position).

Decoding Results

As described in the main text, we examined the extent to which information for each task feature is represented across time. To observe this emergence across time, we decoded the task feature at each individual timepoint, thus the results at any given point are independent from the adjacent timepoints. From our previous work using eye tracking in similar paradigms (Mayr et al., 2013; Kikumoto et al., 2015), we expected that there would be timing differences in the processing of the Color versus Orientation task. Generally, color is processed more rapidly than most other features. This was confirmed in exploratory analyses showing a more rapid deployment of attention to the color target compared to the orientation target. Given these task differences, we did not want the classifier to pick up on contingencies that were based on the task alone. Similarly, when decoding the cue, which has 4 possible values, looking at decoding accuracy alone makes it difficult to distinguish how much the classifier is picking up on between- versus within-task contingencies. Accordingly, for all features except for task itself, we preformed decoding separately within each task. Thus the solution that was fit in the training set could not be due to task differences. For Figure 1, we then simply averaged the decoding accuracy across both tasks. For each feature, we confirmed that there were no substantial differences, where there was a complete absence of information for one task. Thus all features had a 50% chance rate, except for the location of the target/distractor. For the target/distractor location, we decoded the position based on the bin that each item appeared in (with 2 unique positions per bin). Since the task was designed such that the target and distractor bins were balanced, this ensured that the target and distractor occupied each combination of bins with equal frequency (including sharing the same bin). This way, we can rule out the contention that successful distractor decoding may be due to the classifier actually decoding "not target position". Thus chance accuracy for the target and distractor is 25%, and in Figure 1 in the main text, these are simply shifted up so that chance level is the same as the other features.

As evident in Figure 1C, the decoding accuracy for each feature follows what we would expect in the information processing stream. Just after the cue onset at -300 ms, we see increased decoding accuracy for the cue. Since the cue itself is only informative in its relation to the task set, we would not expect such information to persist for an extended time, and in the decoding accuracy, we see that it greatly reduced by stimulus onset and completely absent by 400 ms. At a similar time, we see a ramping up for information for the task, mostly after the cue information as died down. This is expected to be a persisting representation that lasts the duration of the trial, and we see that the information does indeed stay largely above chance throughout the interval. Right at stimulus onset, we see an increase in information towards the target position (bin) that peaks around 250 ms post-stimulus, and begins to die down towards the end of the interval. We see a similar pattern but with lower overall accuracy for the distractor position, which again dies down towards the end of the interval. Lastly, information for the response (or stimulus) builds starting around stimulus onset, but doesn't peak until the end of the interval, when information for most other features have begun to decline.

While the emergence of information across time makes sense, with the decoding of the task in particular, this analysis alone cannot rule out the possibility that the classifier is not decoding an abstract attentional set, but perhaps lower-level features such as the response—for instance, if the color task is easier to attend to, or easier to make the decision once the item is attended, then the classifier could simply pick up on this contingency and use it as a means of (indirectly) decoding the task. If, however, we are indeed picking up on an abstract attentional or task set, then we would expect this representation to generalize across the particular (left vs. right) responses. Accordingly, we performed a cross-classification analysis where at each timepoint we trained the classifier to distinguish the task on trials that ultimately required a left response, and tested the classifier on trials requiring a right response (and *vice versa*). If the classifier is picking up on an abstract representation, then the particular response should not matter, and we should still see robust decoding accuracy with the same temporal profile as in Figure 1C. As can be seen in Figure 5, when cross-classifying across responses, we get a slightly lower, but still above-chance decoding of the task, compared to classifying within left/right responses. We also performed the converse analysis, shown in Figure 7—decoding the response itself (left vs. right) but training on one task, and testing on the other task. If we are indeed picking up on a left versus right manual response, then this should generalize across the particular task in which it occurs, which appears to be the case. Note that in figures 5 and 6 we extend the interval to 700 ms since we are specifically examining the generalizability of these representations that occur later in the trial, but the pattern does not change drastically after the 500 ms mark.

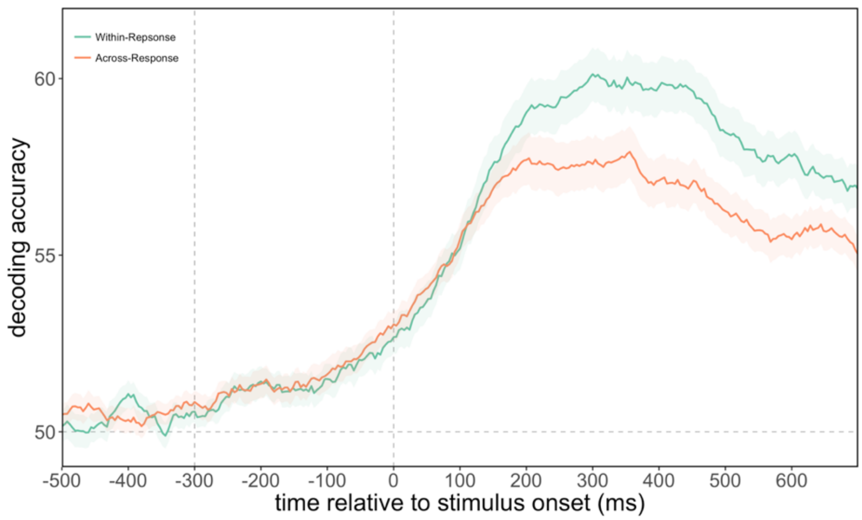

Figure 5. Cross-classification of task across unique responses. Classifiers were trained to distinguish task (Color/Orientation) in left responses, then tested on trials with right responses, and vice-versa. For completeness, the within-response classification is also displayed.

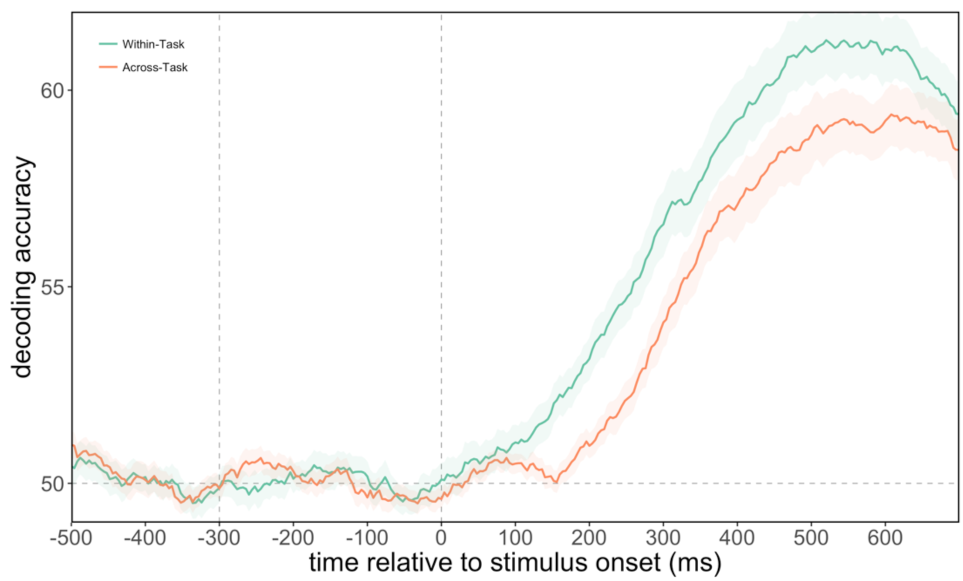

Figure 6. Cross-classification of response across tasks. Classifiers were trained to distinguish response (left/right) in one task, then tested on the other task, and vice-versa. The within-task decoding is also shown for completeness.

Classifier Confidence by Performance

In the following analyses, we turn to the posterior probabilities for each task feature and each trial, and examine the extent that they are predictive of behavior. As a first step, we examine these probabilities in a similar manner as the classification accuracy, but no we average the results based on factors that were never given to the classifier. We saw in the behavioral results that we have a reliable switch effect, so we wanted to see the extent to which each feature was modulated by switch versus no-switch trials. It is possible that a change in task will cause a disruption in processing of all task features. Alternatively, the behavioral switch effect may only be related to modulations in certain task features. Similarly, we also acknowledge that within switch and no-switch trials, there will be variability in task performance not necessarily related to the task factors (e.g., lapses in attention, mind-wandering, fatigue). These factors may also result in a more general perturbation of all task-relevant information, or more specific factors. Accordingly, we first did a median split on RTs in each individual, separately within switch and no-switch trials. We then averaged the posterior probabilities from the main classification analysis separately within each switch and RT condition (i.e., fast vs. slow RTs). Note that except for decoding task, classification was performed within tasks separately, and thus the classifier had no explicit knowledge of the switch condition or the fast vs. slow RTs. Thus any differences in the trial-by-trial confidence in the task features should arise from differences in the clarity of information that results from performance (e.g., slow responses or switch trials). These results are presented in Figure 3 in the main text. It reveals that evidence for task, response, and target position are quite different between fast and slow trials, while evidence for the task is selectively modulated across switch and no-switch conditions, but only in the slow trials. This pattern of results makes sense, suggesting that when one is more on-task, it is associated with a sharper (i.e., more easily decodable) representation of the target, the task, and ultimately the response. Conversely, being on-task does not necessarily modulate processing of the cue, which is not surprising given that the cue comes before stimulus onset and only serves as an indicator for the task set. What was surprising was the lack of a difference for the distractor position, which we expected might be increased on slow RT and switch trials. The relatively low decoding accuracy in the overall results, however, indicate that there is relatively little information from the distractor that can be recovered. It should be noted that since the posteriors bear some relationship to the decoding accuracy, averaging the decoding accuracy in this manner produces very similar results, except they are noisier due to the fact that accuracy is binarized instead of continuous.

Predicting Performance from Classifier Confidence

As a last step, we wanted to examine the relationship between the classifier confidence and performance in a more fine-grained manner. Specifically, we wanted to establish the extent to which information about each task element was predictive of the ultimate response time on each trial. We did this using the same posterior probabilities examined above, but used them as predictors in a linear mixed effects model predicting RT. First, RTs were prewhitened by removing any linear or quadratic trends over the course of the experiment. Next, as the posterior probabilities are bounded and not normally distributed, we logit-transformed them prior to the analysis. In the first step, we examined how the representation each task element across time contributes to the ultimate response time. Accordingly, for each timepoint and each task element we ran a separate mixed-effects model, using the (logit-transformed) probability of a given task element (e.g., task, target position) and the prewhitened RT data as the dependent variable. Each model included a random intercept for each participant as well as a random slope. The estimates for each task element across time are plotted in Figure 2A in the main text. Bands indicate 95% confidence intervals around the estimate. Note that since the dependent variable is RT, coefficients going in the negative direction indicate better performance. What is most apparent is the prediction associated with the task confidence, which shows the largest fluctuation, beginning slowly around cue onset, but increasing dramatically after stimulus onset. Given that it goes in the negative direction tells us that higher confidence in the task representation is associated with faster responses, particularly in the interval around 200-400 ms post-stimulus. Similarly, we see a negative-going prediction for the target position and the response, starting to increase right after stimulus onset, but peaking a bit later after the task predictability. Similar to the results above, the cue carries little predictive power, but if anything, shows a trend towards slower responses. This makes sense, as a robust representation of the cue (the auditory stimulus) would only hinder task performance after stimulus onset, as one needs to translate the cue processing to the actual task set. The most surprising pattern is for that of the distractor, which is predictive of *shorter* RTs. This predictability occurs relatively late in the trial compared to the other task elements. Note that in this analysis, we ran a separate model for each task element, disregarding the other elements. Since the decoding was also carried out in this manner, we felt this was the most appropriate. We did, however, perform another analysis including all elements in the same model, assessing the predictability for each, above-and-beyond the influence of the other elements, and this model yielded very similar results.

As a next step, rather than tracking the evolution across time, we wanted to better capture the contribution of each task-relevant element to performance on a trial-by-trial level. It is certainly reasonable to assume that different sources of information are more robustly maintained at different parts of the trial when they are actively being used—for instance, there would be no expectation of response-relevant information until after the task and target position have been established. With this in mind, we turned to the original decoding analysis, and identified for each task feature the time of maximum decodability. We then chose a fixed window of 150 ms centered around this point for each element separately, then averaged the (logit-transformed) posterior probabilities across this window for each trial. For the response, the maximum decoding accuracy was towards the end of the 500 ms interval, so we simply averaged the period from 350-500 ms. Thus for each trial we had an average probability across a particular window in time for each task element. We then used these averaged (logit-transformed) probabilities in a single mixed-effects model, with each probability as a predictor, and the prewhitened RT as the dependent variable. In this case it was appropriate to include all elements in the same model, since we were not making assumptions about them being simultaneously represented in time. This model included a random intercept and a random slope for all predictors for each participant. The estimates from this model are depicted in Figure 7. Note that we see a similar pattern as in Figure 2A, where task seems to carry the most influence (and predicting faster responses), followed by the target position and the response. Again, the cue and distractor position carry little predictive power with respect to the response time.

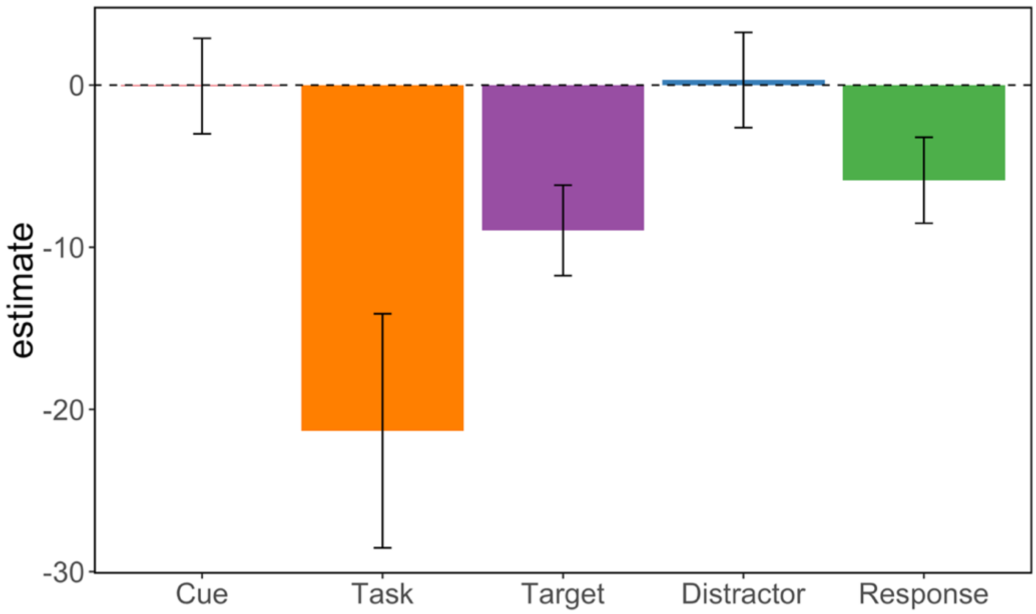

Figure 7. Performance-related predictability across trials. For each element, the posterior probability was averaged over a 150 ms including the point of maximum decodability. These average probabilities were included as predictors in a single mixed-effects model predicting RT across trials. Error bars denote 95% confidence intervals around the estimates.

From the analyses above, it is clear that the extent to which information about the task can be decoded from the EEG signal has some relation to performance on the task itself. This was also true for the processing of the target position, and for the response. This prompts the question of whether what is decoded as "task" may subsume some information related to the other two factors. In other words, is it an abstract attentional set that arises independently of the lower-level factors, or is there some contingency where the lower-level information could masquerade as this abstract set in the decoding analysis? To address this question, we performed similar timepoint-by-timepoint regressions as above, but this time using the (logit-transformed) task probability as a predictor, and the other factors (e.g., target position, response) as the dependent variables. The result gives a timepoint-by-timepoint representation of how strongly each of the other factors is related to the task decoding. If the decoding of task is really just related to these lower-level factors, then we would see very strong positive relationships in the classifier confidences. These results are presented in Figure 2B in the main text, and repeated in Figure 8 to also include the distractor-related information. As evident in Figure 8, the target position and the response have the highest associations with the task-related confidence. Importantly, we see that these associations emerge at different points in time, with the target position being related to task early in the trial, and the response shortly thereafter. It is also notable that these associations are limited to specific windows in time—if we were truly decoding the same information across contexts, then we would expect strong associations across the entire interval. Notably, in Figure 1C it is clear that there is above-chance decoding of task, target, and response before the associations between them emerge. We also see a modest positive association with the cue, but only in the pretrial interval, which also makes sense. Overall this analysis suggests that there is probably some shared information among the attentional set and the lower-level factors, as suggested in the simultaneous model discussed in the main text, but we are also capturing distinct information from each element. The findings from the cross-classification analyses also support the notion that we are capturing an abstract set that can generalize across particular responses.

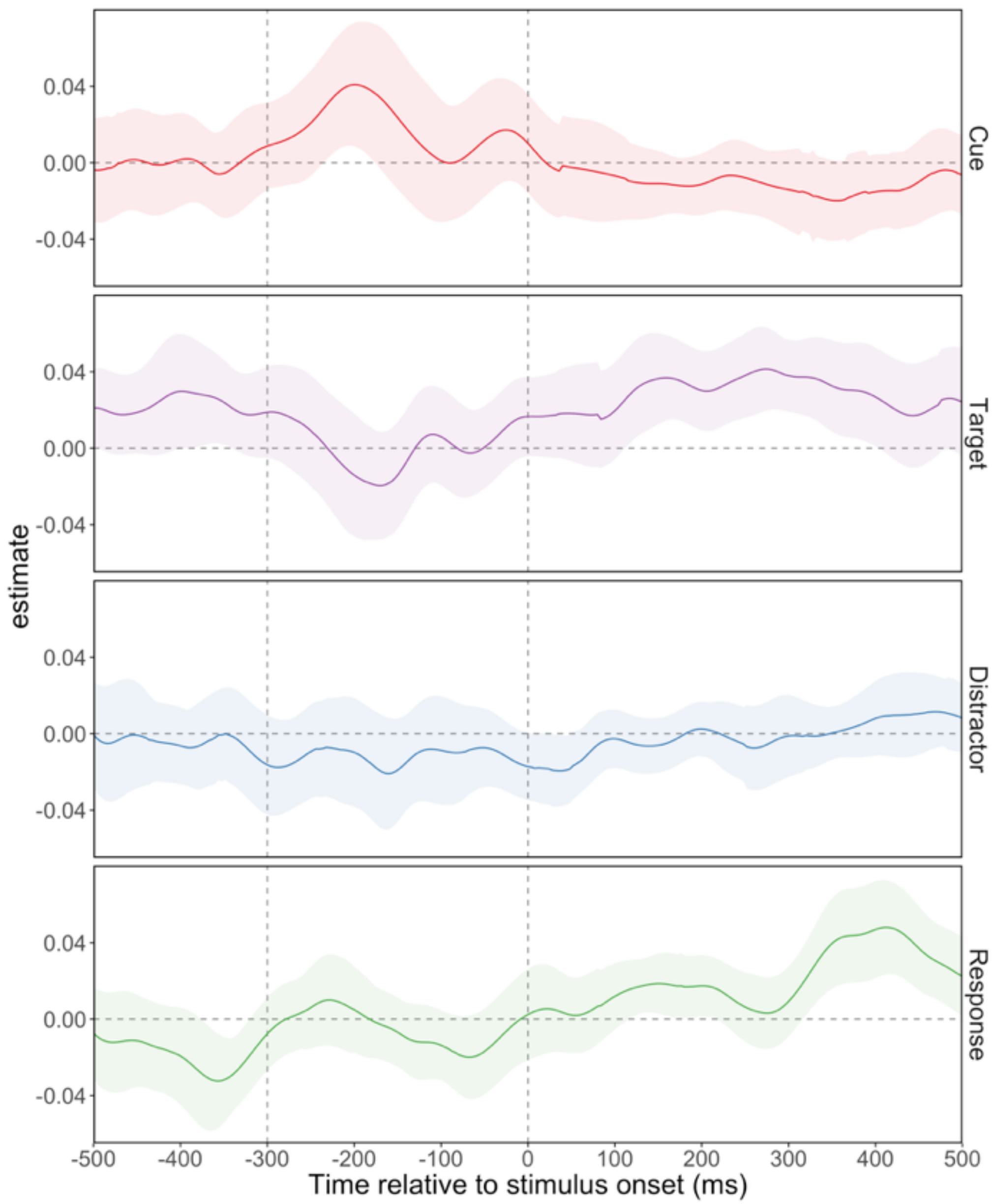

Figure 8. Associations between the posterior probabilities for each factor and task evidence at each timepoint. Bands denote 95% confidence intervals around the estimates.

CHAPTER V

CONCLUSION

The present project was aimed at clarifying the fluctuations of executive control as they occur within a complex task environment. As discussed above, there is a body of evidence suggesting that executive control operates, at least in part, by pushing the cognitive system between broad processing states. These states, in turn, influence how lower-level systems (attention, memory) operate with respect to their reliance on information in the environment (Aston-Jones & Cohen, 2005; O'Reilly, 2006; Lindenberger & Mayr, 2014). Specifically, we examine two such modes of processing, one which robustly maintains representations independently of environmental influences, and another which specifically seeks out information from the environment to update those representations. As we discussed, each mode has its strengths and weaknesses depending on the task context—focused attention is beneficial when filtering out irrelevant information is necessary, but can also lead to overly-rigid behavior. Conversely, flexible behavior can be desirable in some contexts, but can lead to distractibility. The contention that older adults are over reliant on the updating mode would explain why they suffer on so many focused-attention tasks, but not so much on ones requiring flexibility (Kray & Lindenberger, 2000; Linenberger & Mayr, 2014).

While there is no shortage of work examining executive control in general, there is relatively little specifically examining these broad states, and thus the present work makes a contribution in this regard. Using paradigms that provide crisp distinctions between these modes allowed us to examine how and when switches to the updating state occur, and, using age-comparative samples allowed us to determine specifically how age differences are manifested. Chapter II used a paradigm that enforced shifts to maintenance and updating modes, and by looking at performance time-locked to the forced updates, we saw that older adults seemed to persist in this updating mode for several trials. Younger adults also suffered from the updates (as expected), but to a much lesser degree. Using eye tracking we were able to examine the allocation of attention prior to response, to characterize specifically how people were distracted on these trials. We saw that in the older adults, updates led to profound changes in attentional allocation, where the eyes showed considerable distractibility from irrelevant information in both the early and late portions of the trial, and persisting for several trials. Younger adults, on the other hand, saw much more transient distractibility that was much more quickly resolved. In general, the patterns were consistent with the notion that older adults were more frequently in the updating state, and specifically that the forced updates led to them being stuck in that state, while younger adults were more quick to bounce back. In Chapter III, we turned to the more challenging question regarding updating that occurs more naturally, and not necessarily enforced by the task. By focusing on fixations towards irrelevant cues as an index of updating, we found that updates did occur in response to certain task factors (e.g., conflict, errors), but also as a result of the more global task context (mixed versus single-task blocks), but also independently of those factors (passive carry-over from one trial to the next). Consistent with other work, we found that older adults were particularly sensitive to the global context (Spieler, Mayr, & Lagrone, 2006). We also found they showed a stronger tendency towards passive carry-over, which is consistent with the notion that they stay stuck in the updating state over longer epochs compared to younger adults. The younger adults, by contrast, seemed to be more swayed by the local task factors, suggesting they were entering this state more in response to the task demands. Lastly, in Chapter IV we used neuroimaging (EEG) to lay the groundwork for more directly addressing these issues, by providing a means of measuring the trial-by-trial and moment-by-moment processing of the task environment. In both the decoding analyses, we found that we were able to extract distinct elements of the task at different points in time, and in the prediction analyses we showed that this information was behaviorally-relevant. Moving forward, we can examine how these representations are altered in the updating or maintenance states. For instance, in Chapter II, we examine how attention towards targets and distractors are changed in the updating state, with the assumption that a weaker representation of the task set in the updating state would lead to more distractor fixations and/or slower responses. Using the basic setup as in Chapter IV but adding the same interruption procedure, we could directly examine how the task-set representation is modulated in these post-interruption trials without having to rely on the RT or eye movement measures. Similarly, we could contrast how these representations change under different global task demands, such as mixed versus single-task blocks. Lastly, by adding age-comparative samples, we could have a more sensitive metric by which to gauge the age differences. This more sensitive measure could also enable us to examine outstanding questions regarding the dynamics of these global states. For instance, it is still unclear whether the endogenously-generated updating (as seen in Chapter III) may be periodic in nature, occurring at some regular frequency. While the age differences seem clear, there is also the possibility that there are stable individual differences in the frequency of updating, which could have direct consequences on real-world outcomes (e.g., academic performance). There is also the outstanding issue whether the increased updating in older adults is a direct consequence of biological changes (e.g., decrease in dopamine production, Bäckman, et al., 2006) that drive updating automatically, or a strategy on the part of the individual to deal with compromised abilities, which may or may not be biologically influenced (e.g., Gazzaley, 2013; Braver et al., 2001). At the moment, the RT and eye tracking data are just too coarse to address these questions, and thus the neuroimaging measures provide some hope. Nevertheless, the present project was successful in providing a more thorough examination of these executive control dynamics, and in identifying directions for future work.

APPENDIX

FULL MODEL SUMMARIES FROM CHAPTER II

Table 1. Updating Effect in Experiment 1 by Conflict and Task

	estimate	*t*
Endo, No-Conflict		
Age	465	8.46
Experimental	72	1.31
Update	57	3.71
Age x Experimental	122	1.10
Age x Update	81	2.64
Experimental x Update	28	0.92
Age x Experimental x Update	81	1.32
Endo, Conflict		
Age	612	8.11
Experimental	176	2.34
Update	85	4.17
Age x Experimental	249	1.65
Age x Update	124	3.03
Experimental x Update	114	2.78
Age x Experimental x Update	184	2.24
Exo, No-Conflict		
Age	392	7.55
Experimental	16	0.31
Update	166	7.49
Age x Experimental	-13	-0.13
Age x Update	218	4.91
Experimental x Update	42	0.94
Age x Experimental x Update	-27	-0.31
Exo, Conflict		
Age	430	7.20
Experimental	129	2.16
Update	258	8.25
Age x Experimental	134	1.12
Age x Update	248	3.97
Experimental x Update	235	3.75
Age x Experimental x Update	167	1.34

Table 2. Linear Effect of Maintenance Trials in Experiment 2

	estimate	*t*
Endo Task		
Age	489	7.70
Experimental	95	1.50
Conflict	118	8.54
Linear	2	1.00
Quadratic	-1	-0.73
Age x Experimental	135	1.06
Age x Conflict	115	4.16
Experimental x Conflict	57	2.06
Age x Linear	**-1**	**-0.25**
Age x Quadratic	0	-0.20
Experimental x Linear	-2	-0.51
Experimental x Quadratic	-1	-0.50
Conflict x Linear	2	0.65
Conflict x Quadratic	0	-0.02
Age x Experimental x Conflict	77	1.39
Age x Experimental x Linear	**0**	**-0.02**
Age x Experimental x Quadratic	-3	-0.83
Age x Conflict x Linear	2	0.26
Age x Conflict x Quadratic	0	0.13
Experimental x Conflict x Linear	-15	-2.34
Experimental x Conflict x Quadratic	4	1.13
Age x Experimental x Conflict x Linear	**-14**	**-1.10**
Age x Experimental x Conflict x Quadratic	3	0.45
Exo Task		
Age	296	7.98
Experimental	1	0.03
Conflict	14	1.83
Linear	-8	-4.32
Quadratic	2	2.53
Age x Experimental	25	0.33
Age x Conflict	22	1.43
Experimental x Conflict	1	0.05
Age x Linear	**-11**	**-3.11**
Age x Quadratic	3	1.85
Experimental x Linear	-12	-3.45
Experimental x Quadratic	3	2.34
Conflict x Linear	-7	-2.69
Conflict x Quadratic	2	1.48
Age x Experimental x Conflict	44	1.41
Age x Experimental x Linear	**-12**	**-1.71**
Age x Experimental x Quadratic	2	0.87
Age x Conflict x Linear	-8	-1.55
Age x Conflict x Quadratic	1	0.49
Experimental x Conflict x Linear	-9	-1.65
Experimental x Conflict x Quadratic	5	1.82
Age x Experimental x Conflict x Linear	**14**	**1.32**
Age x Experimental x Conflict x Quadratic	-1	-0.16

Table 3. Comparison with No-Interruption Controls in Experiment 1

	estimate	*t*
Both Tasks		
Experimental	93	1.54
Task	445	15.98
Conflict	116	10.70
Experimental x Task	163	2.94
Experimental x Conflict	47	2.16
Task x Conflict	134	8.68
Experimental x Task x Conflict	102	3.32
Endo Task		
Experimental	175	2.18
Conflict	184	12.56
Experimental x Conflict	97	3.33
Exo Task		
Experimental	13	0.26
Conflict	49	3.95
Experimental x Conflict	-2	-0.06

Table 4. Updating Effect in Experiment 2 by Conflict

	estimate	*t*
No-Conflict		
Age	262	5.46
Task	237	12.54
Update	156	8.13
Age x Task	82	2.16
Age x Update	**138**	**3.61**
Task x Update	-62	-3.96
Age x Task x Update	33	1.05
Conflict		
Age	332	5.65
Task	350	11.88
Update	267	10.65
Age x Task	173	2.94
Age x Update	**177**	**3.54**
Task x Update	-130	-6.08
Age x Task x Update	7	0.16

Table 5. Linear Effect of Maintenance Trials in Experiment 2

	estimate	*t*
Both Tasks		
Age	219	5.26
Task	336	39.36
Conflict	99	8.48
Linear	-4	-1.97
Quadratic	2	2.02
Age x Task	129	7.54
Age x Conflict	62	2.67
Task x Conflict	140	8.21
Age x Linear	**-7**	**-1.97**
Age x Quadratic	2	1.32
Task x Linear	6	2.51
Task x Quadratic	0	0.36
Conflict x Linear	**-6**	**-2.65**
Conflict x Quadratic	1	0.91
Age x Task x Conflict	70	2.04
Age x Task x Linear	-4	-0.85
Age x Task x Quadratic	-3	-1.03
Age x Conflict x Linear	**-5**	**-1.15**
Age x Conflict x Quadratic	-1	-0.31
Task x Conflict x Linear	1	0.25
Task x Conflict x Quadratic	2	0.87
Age x Task x Conflict x Linear	7	0.80
Age x Task x Conflict x Quadratic	5	0.83
Endo Task		
Old	270	5.27
Conflict	159	7.97
Linear	-18	-4.72
Quadratic	9	6.55
Old x Conflict	90	2.25
Age x Linear	**-22**	**-2.90**
Age x Quadratic	6	2.33
Conflict x Linear	**-7**	**-2.22**
Conflict x Quadratic	4	2.38
Age x Conflict x Linear	**-1**	**-0.08**
Age x Conflict x Quadratic	2	0.73
Exo Task		
Old	145	3.91
Conflict	19	1.53
Linear	-33	-9.36
Quadratic	12	8.57
Old x Conflict	29	1.20
Age x Linear	**-19**	**-2.74**
Age x Quadratic	8	2.76
Conflict x Linear	**-20**	**-7.67**
Conflict x Quadratic	5	4.25
Age x Conflict x Linear	**-6**	**-1.23**
Age x Conflict x Quadratic	-4	-1.43

REFERENCES CITED

Chapter I

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review of Neuroscience, 28*, 403-450.

Banich, M. T. (2009). Executive Function: The Search for an Integrated Account. *Current Directions in Psychological Science*, *18*(2), 89–94.

Dreisbach, G., & Goschke, T. (2004). How Positive Affect Modulates Cognitive Control: Reduced Perseveration at the Cost of Increased Distractibility. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(2), 343–353.

Durstewitz, D., & Seamans, J. K. (2008). The Dual-State Theory of Prefrontal Cortex Dopamine Function with Relevance to Catechol-O-Methyltransferase Genotypes and Schizophrenia. *Biological Psychiatry*, *64*(9), 739–749

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Attention, Perception, & Psychophysics, 16*(1), 143-149.

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2002). Cognitive Neuroscience, New York: W. W.

Kikumoto, A., Hubbard, J., & Mayr, U. (2015). Dynamics of task-set carry-over: Evidence from eye-movement analyses. *Psychonomic Bulletin & Review*, *23*(3), 899–906.

Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging*, *15*(1), 126–147.

Lindenberger, U., & Mayr, U. (2014). Cognitive aging: Is there a dark side to environmental support? *Trends in Cognitive Sciences*, *18*(1), 7–15.

Mayr, U. (2001). Age differences in the selection of mental sets: The role of inhibition, stimulus ambiguity, and response-set overlap. *Psychology and Aging*, *16*(1), 96–109.

Mayr, U., Kuhns, D., & Hubbard, J. (2014). Long-term memory and the control of attentional control. *Cognitive Psychology*, *72*(C), 1–26.

Mayr, U., Kuhns, D., & Rieter, M. (2013). Eye movements reveal dynamics of task control. *Journal of Experimental Psychology: General, 142*(2), 489.

O'Reilly, R. C. (2006). Biologically based computational models of high-level cognition. *Science*, *314*(5796), 91–94.

Rogers, W. A., Hertzog, C., & Fisk, A. D. (2000). An individual differences analysis of ability and strategy influences: Age-related differences in associative learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 26*, 359-394.

Salthouse, T. A. (2005). Relations Between Cognitive Abilities and Measures of Executive Functioning. *Neuropsychology*, *19*(4), 532–545.

Spieler, D. H., Mayr, U., & Lagrone, S. (2006). Outsourcing cognitive control to the environment: Adult age differences in the use of task cues. *Psychonomic Bulletin & Review*, *13*(5), 787–793.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology, 18*(6), 643.

Chapter II

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390 – 412.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48

Braver, T. S., Barch, D. M., Keys, B. A., Keys, B. A., Carter, C. S., … Reed, B. R. (2001). Context processing in older adults: Evidence for a theory relating cognitive control to neurobiology in healthy aging. *Journal of Experimental Psychology: General*, *130*(4), 746.

Braver, T. S., & West, R. (2008). Working memory, executive control, and aging. The handbook of aging and cognition, 3, 311-372.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers*, *34*(4), 613–617.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorial in Quantitative Methods for Psychology, 1*(1), 4–45.

Gazzaley, A. (2013). *Top-down modulation deficit in the aging brain: An emerging theory of cognitive aging*. R.T. Knight, D.T. Stuss. Principles of Frontal Lobe Function.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*(4), 411–428.

Kikumoto, A., Hubbard, J., & Mayr, U. (2015). Dynamics of task-set carry-over: Evidence from eye-movement analyses. *Psychonomic Bulletin & Review*, *23*(3), 899–906.

Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging*, *15*(1), 126–147.

Lindenberger, U., & Mayr, U. (2014). Cognitive aging: Is there a dark side to environmental support? *Trends in Cognitive Sciences*, *18*(1), 7–15.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*(4), 492–527.

Mayr, U. (2001). Age differences in the selection of mental sets: The role of inhibition, stimulus ambiguity, and response-set overlap. *Psychology and Aging*, *16*(1), 96–109.

Mayr, U., & Bryck, R. L. (2005). Sticky Rules: Integration Between Abstract Rules and Specific Actions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(2), 337–350.

Mayr, U., Kliegl, R., & Krampe, R. T. (1996). Sequential and coordinative processing dynamics in figural transformations across the life span. *Cognition*, *59*(1), 61–90.

Mayr, U., Kuhns, D., & Hubbard, J. (2014). Long-term memory and the control of attentional control. *Cognitive Psychology*, *72*(C), 1–26.

Mayr, U., Kuhns, D., & Rieter, M. (2013). Eye movements reveal dynamics of task control., *142*(2), 489–509.

O'Reilly, R. C. (2006). Biologically based computational models of high-level cognition. *Science*, *314*(5796), 91–94.

Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings.

Rogers, W. A., Hertzog, C., & Fisk, A. D. (2000). An individual differences analysis of ability and strategy influences: Age-related differences in associative learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 26*, 359-394.

Royall, D. R., Palmer, R., Chiodo, L. K., & Polk, M. J. (2004). Declining Executive Control in Normal Aging Predicts Change in Functional Status: The Freedom House Study. *Journal of the American Geriatrics Society*, *52*(3), 346–352.

Salthouse, T. A. (2005). Relations Between Cognitive Abilities and Measures of Executive Functioning. *Neuropsychology*, *19*(4), 532–545.

Spieler, D. H., Mayr, U., & Lagrone, S. (2006). Outsourcing cognitive control to the environment: Adult age differences in the use of task cues. *Psychonomic Bulletin & Review*, *13*(5), 787–793.

Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998). Our Eyes do Not Always Go Where we Want Them to Go: Capture of the Eyes by New Objects. *Psychological Science*, *9*(5), 379–385.

Verhaeghen, P. (2011). Aging and Executive Control: Reports of a Demise Greatly Exaggerated. *Current Directions in Psychological Science*, *20*(3), 174–180.

Chapter III

Altmann, E. M. & Gray, W. D. (2002). Forgetting to remember: The functional relationship of decay and interference. *Psychological Science, 13*, 27-33

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48

Bäckman, L., Nyberg, L., Lindenberger, U., Li, S.-C., & Farde, L. (2006). The correlative triad among aging, dopamine, and cognition: Current status and future prospects. *Neuroscience and Biobehavioral Reviews*, *30*(6), 791–807.

Braver, T. S., & West, R. (2008). Working memory, executive control, and aging. The handbook of aging and cognition, 3, 311-372.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers*, *34*(4), 613–617.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorial in Quantitative Methods for Psychology, 1*(1), 4–45.

Dutilh, G., Forstmann, B. U., Vandekerckhove, J., & Wagenmakers, E. J. (2013). A diffusion model account of age differences in posterror slowing. *Psychology and Aging, 28*(1), 64.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: a dynamical model of saccade generation during reading. *Psychological Review, 112*(4), 777.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Attention, Perception, & Psychophysics, 16*(1), 143-149.

Esterman, M., Noonan, S. K., Rosenberg, M., & DeGutis, J. (2012). In the zone or zoning out? Tracking behavioral and neural fluctuations during sustained attention. *Cerebral Cortex, 23*(11), 2712 - 2723.

Gazzaley, A. (2013). *Top-down modulation deficit in the aging brain: An emerging theory of cognitive aging*. R.T. Knight, D.T. Stuss. Principles of Frontal Lobe Function.

Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(2), 252–269.

Jepma, M., & Nieuwenhuis, S. (2011). Pupil Diameter Predicts Changes in the Exploration–Exploitation Trade-off: Evidence for the Adaptive Gain Theory. *Journal of Cognitive Neuroscience*, *23*(7), 1587–1596.

Kikumoto, A., Hubbard, J., & Mayr, U. (2015). Dynamics of task-set carry-over: Evidence from eye-movement analyses. *Psychonomic Bulletin & Review*, *23*(3), 899–906.

Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging*, *15*(1), 126–147.

Lindenberger, U., & Mayr, U. (2014). Cognitive aging: Is there a dark side to environmental support? *Trends in Cognitive Sciences*, *18*(1), 7–15.

Lu, C. H., & Proctor, R. W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic bulletin & review*, *2*(2), 174-207.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*(2), 163-203.

Mayr, U. (2001). Age differences in the selection of mental sets: The role of inhibition, stimulus ambiguity, and response-set overlap. *Psychology and Aging*, *16*(1), 96–109.

Mayr, U., Awh, E., & Laurey, P. (2003). Conflict adaptation effects in the absence of executive control. *Nature neuroscience, 6*(5), 450-452.

Mayr, U., Kliegl, R., & Krampe, R. T. (1996). Sequential and coordinative processing dynamics in figural transformations across the life span. *Cognition*, *59*(1), 61–90.

Mayr, U., Kuhns, D., & Hubbard, J. (2014). Long-term memory and the control of attentional control. *Cognitive Psychology*, *72*(C), 1–26.

Mayr, U., Kuhns, D., & Rieter, M. (2013). Eye movements reveal dynamics of task control., *142*(2), 489–509.

Nieuwenhuis, S., & Jepma, M. (2011). Investigating the role of the noradrenergic system in human cognition. *Decision making. Attention & performance, 23*, 367-385.

O'Reilly, R. C. (2006). Biologically based computational models of high-level cognition. *Science*, *314*(5796), 91–94.

Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology and Aging, 19*(2), 278.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies. *Psychological Bulletin, 132*(1), 1.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of experimental psychology: General, 124*(2), 207.

Rogers, W. A., Hertzog, C., & Fisk, A. D. (2000). An individual differences analysis of ability and strategy influences: Age-related differences in associative learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 26*, 359-394.

Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews Neuroscience*, *10*(3), 211–223.

Spieler, D. H., Mayr, U., & Lagrone, S. (2006). Outsourcing cognitive control to the environment: Adult age differences in the use of task cues. *Psychonomic Bulletin & Review*, *13*(5), 787–793.

Starns, J. J., & Ratcliff, R. (2010). The effects of aging on the speed–accuracy compromise: Boundary optimality in the diffusion model. *Psychology and Aging, 25*(2), 377.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology, 18*(6), 643.

Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998). Our Eyes do Not Always Go Where we Want Them to Go: Capture of the Eyes by New Objects. *Psychological Science*, *9*(5), 379–385.

Verhaeghen, P. (2011). Aging and Executive Control: Reports of a Demise Greatly Exaggerated. *Current Directions in Psychological Science*, *20*(3), 174–180.

Chapter IV – Main Text

1. Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K., & Awh, E. *Journal of Neurophysiology* **115**(1), 168-177 (2016).

2. Mayr, U., Kuhns, D., & Rieter, M. *Journal of Experimental Psychology: General* **142**(2), 489 (2013).

3. Logan, G. D., & Bundesen, C. *Psychonomic Bulletin & Review* **11**(5), 832-840 (2004).

4. Mayr, U., & Kliegl, R. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **29**(3), 362 (2003).

Chapter IV – Extended Methods

Buschman, T. J., Denovellis, E. L., Diogo, C., Bullock, D., & Miller, E. K. (2012). Synchronous Oscillatory Neural Ensembles for Rules in the Prefrontal Cortex. *Neuron, 76*(4), 838–846.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48

Cohen, M. X. (2014). Analyzing neural time series data: theory and practice. MIT Press.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorial in Quantitative Methods for Psychology, 1*(1), 4–45.

Engel, a K., Fries, P., & Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. Nature Reviews. Neuroscience, 2(10), 704–16.

Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K., & Awh, E. (2016). The topography of alpha-band activity tracks the content of spatial working memory. Journal of neurophysiology, 115(1), 168-177.

Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. Trends in Cognitive Sciences, 9(10), 474–480.

Kikumoto, A., Hubbard, J., & Mayr, U. (2015). Dynamics of task-set carry-over: Evidence from eye-movement analyses. *Psychonomic Bulletin & Review*, *23*(3), 899–906.

Mayr, U. (2006). What matters in the cued task-switching paradigm: Tasks or cues?. Psychonomic bulletin & review, 13(5), 794-799.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12(Oct), 2825-2830.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of experimental psychology: General, 124*(2), 207.

Chapter V

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review of Neuroscience, 28*, 403-450.

Bäckman, L., Nyberg, L., Lindenberger, U., Li, S.-C., & Farde, L. (2006). The correlative triad among aging, dopamine, and cognition: Current status and future prospects. *Neuroscience and Biobehavioral Reviews*, *30*(6), 791–807.

Braver, T. S., Barch, D. M., Keys, B. A., Keys, B. A., Carter, C. S., … Reed, B. R. (2001). Context processing in older adults: Evidence for a theory relating cognitive control to neurobiology in healthy aging. *Journal of Experimental Psychology: General*, *130*(4), 746.

Gazzaley, A. (2013). *Top-down modulation deficit in the aging brain: An emerging theory of cognitive aging*. R.T. Knight, D.T. Stuss. Principles of Frontal Lobe Function.

Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging*, *15*(1), 126–147.

Lindenberger, U., & Mayr, U. (2014). Cognitive aging: Is there a dark side to environmental support? *Trends in Cognitive Sciences*, *18*(1), 7–15.

O'Reilly, R. C. (2006). Biologically based computational models of high-level cognition. *Science*, *314*(5796), 91–94.

Spieler, D. H., Mayr, U., & Lagrone, S. (2006). Outsourcing cognitive control to the environment: Adult age differences in the use of task cues. *Psychonomic Bulletin & Review*, *13*(5), 787–793.