

Context-Specific Adjustments of Cognitive Flexibility

by

Audrey Siqu-Liu

Department of Psychology and Neuroscience  
Duke University

Date: \_\_\_\_\_

Approved:

\_\_\_\_\_  
Tobias Egner, Supervisor

\_\_\_\_\_  
Scott Heuttel

\_\_\_\_\_  
Marty Woldorff

\_\_\_\_\_  
Felipe DeBrigard

Dissertation submitted in partial fulfillment of  
the requirements for the degree of Doctor of Philosophy  
in the Department of  
Psychology and Neuroscience in the Graduate School  
of Duke University

2023

ABSTRACT

Context-Specific Adjustments of Cognitive Flexibility

by

Audrey Siqu-Liu

Department of Psychology and Neuroscience  
Duke University

Date: \_\_\_\_\_

Approved:

\_\_\_\_\_  
Tobias Egner, Supervisor

\_\_\_\_\_  
Scott Heuttel

\_\_\_\_\_  
Marty G. Woldorff

\_\_\_\_\_  
Felipe DeBrigard

An abstract of a dissertation submitted in partial  
fulfillment of the requirements for the degree  
of Doctor of Philosophy in the Department of  
Psychology and Neuroscience in the Graduate School of  
Duke University

2023

Copyright by  
Audrey Siqu-Liu  
2023

## **Abstract**

The stability-flexibility dilemma describes the challenge of balancing the antagonistic goals of focusing on the current task-set (cognitive stability) and updating that task-set in response to changes in the environment (cognitive flexibility). Dynamic adjustments of cognitive flexibility are observed in cued task-switching paradigms, wherein switch costs, or the performance costs associated with switching between tasks, have been shown to decrease as the proportion of switch trials within a block increase. This effect is referred to as the list-wide proportion switch (LWPS) effect, and presents evidence of meta-flexibility, or people's ability to find an optimal level of flexibility based on contextual demands. While context-sensitive control adaptations have been extensively researched in the conflict literature, fewer studies have been dedicated to investigating such flexibility adaptations. Consequently, the underlying mechanisms of meta-flexibility remain unknown. Across four behavioral experiments, Chapter 2 teases apart the different levels of learning that may contribute to list-wide flexibility adaptations, controlling for stimulus- and task-level associations. Chapter 3 investigates the EEG neural signatures of meta-flexibility. In Chapter 4, we test for whether learned flexibility adjustments benefit from memory consolidation, like other forms of associative learning. Lastly, Chapter 5 utilized three novel behavioral paradigms to investigate different conditions under which flexibility learning transfers or fails to

transfer. Collectively, the results in this dissertation suggest that flexibility adjustments to contextual demands occurs rapidly and transfers reliably across novel stimuli, such that, in high switch frequency blocks, participants could perform more rapid task switches even on trials involving items never seen before. However, both behavioral and neural evidence suggest that flexibility learning is also task specific, such that switch performance boosts to tasks that appear more often as switch trials do not generalize to other tasks that occur in the same temporal contexts.

## **Dedication**

To my mother, Yalin Luo.

# Contents

Abstract .....	iv
List of Tables .....	xiii
List of Figures .....	xiv
Acknowledgements .....	xvi
1. Introduction .....	1
1.1 What is cognitive flexibility?.....	1
1.1.1 Task sets, switch costs, and switch-readiness .....	4
1.1.2 The Stability-Flexibility Dilemma .....	6
1.2 Processes that contribute to switch costs .....	10
1.2.1 Task-Set Reconfiguration .....	10
1.2.2 Task Set Inertia.....	12
1.2.3 Associative Retrieval of Task-Sets.....	13
1.2.4 Feature integration account of switch costs.....	14
1.2.5 Switch cost components that may be sensitive to context .....	17
2. Contextual adaptation of cognitive flexibility is driven by task- and item-level learning .....	18
2.1 Experiment 1 .....	21
2.1.1 Method .....	21
2.1.1.1 Participants .....	21
2.1.1.2 Stimuli.....	22
2.1.1.3 Procedure .....	22

2.1.1.4 Design .....	25
2.1.2 Results and Discussion .....	25
2.2 Experiment 2 Introduction .....	29
2.2.1 Method .....	31
2.2.1.1 Participants .....	31
2.2.1.2 Stimuli.....	31
2.2.1.3 Procedure .....	31
2.2.1.4 Design.....	32
2.2.2 Results and Discussion .....	32
2.3 Experiment 3a Introduction .....	35
2.3.1 Method.....	37
2.3.1.1 Participants .....	37
2.3.1.2 Stimuli and Procedures.....	37
2.3.1.3 Design.....	38
2.3.2 Results and Discussion .....	38
2.4 Experiment 3b Introduction.....	41
2.4.1 Method.....	42
2.4.1.1 Participants .....	42
2.4.1.2 Stimuli and Procedures.....	42
2.4.1.3 Design .....	44
2.4.2 Results and Discussion .....	44
2.5 Experiment 4 Introduction .....	48



2.5.1 Method .....	50
2.5.1.1 Participants .....	50
2.5.1.2 Stimuli and Procedures .....	50
2.5.1.3 Design .....	51
2.5.2 Results and Discussion .....	52
2.6 General Discussion .....	55
2.6.1 Task sets form the boundaries of control strategies .....	58
2.6.2 How are task-specific switch readiness adjustments implemented? .....	61
2.6.3 Possible Neural Mechanisms .....	67
2.6.4 Conclusions .....	68
3. Neural dynamics of context-sensitive adjustments in cognitive flexibility .....	69
3.1 Methods .....	74
3.1.1 Participants .....	74
3.1.2 Stimuli .....	75
3.1.3 Procedure .....	75
3.1.4 Behavioral Analysis .....	79
3.1.5 EEG Acquisition and Preprocessing .....	79
3.1.6 ERP Analyses .....	80
3.1.7 Time Frequency Analyses .....	82
3.2 Results .....	84
3.2.1 Behavioral Results .....	84
3.2.2 ERP Results .....	86

3.2.2.1 Task Sequence Effects.....	86
3.2.2.2 Interaction Effect of Task Sequence and PS.....	90
3.2.2.3 Time Frequency Decomposition: Alpha (8-14 Hz) Power .....	92
3.3 Discussion.....	94
3.3.1 Conclusions .....	104
4. Minimal effects of consolidation on learned switch-readiness .....	104
4.1 Experiment 1 .....	109
4.1.1 Method.....	109
4.1.1.1 Sample Size .....	109
4.1.1.2 Participants .....	109
4.1.1.3 Stimuli.....	110
4.1.1.4 Procedure .....	110
4.1.1.5 Data Analysis.....	116
4.1.2 Results.....	117
4.1.2.1 Biased, PS-20/PS-80 Items .....	118
4.1.2.2 Unbiased, PS-50 Items .....	121
4.1.2.3 Discussion .....	124
4.2 Experiment 2 .....	124
4.2.1 Method.....	125
4.2.1.1 Participants .....	125
4.2.1.2 Procedure .....	126
4.2.2 Results.....	126

4.2.2.1 Biased, PS-20/PS-80 Items .....	126
4.2.2.2 Unbiased, PS-50 Items .....	130
4.2.2.3 Discussion .....	132
4.3 Between-Experiment Analysis.....	133
4.4 Exploratory Analyses.....	134
4.5 General Discussion.....	137
4.5.1 Benefits of consolidation may occur for reward-learning, but not control learning .....	141
4.5.2 Benefits of consolidation may depend on the type of control learning.....	145
4.5.3 Conclusions .....	148
5. Task sets define boundaries of learned cognitive flexibility in list-wide switch proportion manipulations.....	149
5.1 Experiment 1 .....	150
5.1.1 Method.....	151
5.1.1.1 Participants .....	151
5.1.1.2 Stimuli.....	152
5.1.1.3 Procedures.....	152
5.1.2 Results .....	154
5.1.3 Discussion.....	155
5.2 Experiment 2 .....	156
5.2.1 Method.....	158
5.2.1.1 Participants .....	158
5.2.1.2 Stimuli.....	158

5.2.1.3 Procedures.....	159
5.2.2 Results.....	160
5.2.3 Discussion.....	161
5.3 Experiment 3.....	162
5.3.1 Method.....	164
5.3.1.1 Participants.....	164
5.3.1.2 Stimuli.....	165
5.3.1.3 Procedures.....	165
5.3.2 Results.....	167
5.3.3 Discussion.....	171
General Discussion.....	172
6. Conclusions.....	176
6.1 Potential mechanisms for task-specific switch readiness learning.....	176
6.2 Task-sets as an event-file boundary for flexibility learning in cued-switch paradigms.....	178
6.3 Limitations and Future Directions.....	180
Appendix A.....	183
Appendix B.....	185
References.....	188
Biography.....	209

## List of Tables

Table 1: Response Times (ms) and Accuracy (%) with Standard Errors in Experiment 1 and 2 as a Function of CSI and Switch Proportion.....	34
Table 2: Mean Response Times (ms) and Accuracy (%) with Standard Errors in Experiment 3a and 3b as a Function of Task Bias and Switch Proportion.....	48
Table 3: Mean Response Times (ms) and Accuracy (%) with Standard Error in Experiment 4 as a Function of Task Bias and Switch Proportion .....	55
Table 4: Mean Response Times (ms) and Error (percentage) as a function of the proportion of switch trials (PS) and task sequence .....	85
Table 5: Inferential Statistics .....	85
Table 6: Excluded Trial Counts and Final Number Remaining in Smallest and Largest Cells.....	183
Table 7: Frequencies of Task Presentations as Switch and Repeat <i>per Different Switch Frequency Blocks</i> .....	184
Table 8: Across Experiments ANOVA with Block Number (1-8) instead of PS factor for Reaction Time .....	186

## List of Figures

Figure 1: Simplified task switching paradigm.....	7
Figure 2: Experimental protocol. ....	23
Figure 3: Experiment 1 mean RTs and switch costs in the (A) short CSI and (B) long CSI Conditions.....	27
Figure 4: Experiment 2 mean RTs and switch costs collapsed across three tasks. ....	33
Figure 5: Experiment 3a mean RTs and switch costs for (A) biased and (B) unbiased transfer tasks.....	39
Figure 6: Experiment 3b mean RTs and switch costs for (A) biased and (B) unbiased transfer tasks.....	45
Figure 7: Experiment 4 mean RTs and switch costs for A) biased and B) unbiased stimuli sets.....	53
Figure 8: Experimental protocol. ....	76
Figure 9: Behavioral results. ....	86
Figure 10: Cue-locked early effects of task sequence.....	88
Figure 11: Cue-locked longer-latency effects of task-sequence.....	90
Figure 12: Cue-locked interaction effect of block-wise PS and task sequence. ....	92
Figure 13: Cue-locked alpha effects.....	94
Figure 14: Experiment design.....	115
Figure 15: Behavioral results from Experiment 1 for biased (PS-20/80) items. ....	120
Figure 16: Behavioral results from Experiment 1 for unbiased (PS-50) items.....	123
Figure 17: Behavioral results from Experiment 2 for biased (PS-20/80) items. ....	129
Figure 18: Behavioral results from Experiment 2 for unbiased (PS-50) items.....	132

Figure 19: Example stimulus and trial timing of Experiment 1 protocol.....	154
Figure 20: Experiment 1 results.....	155
Figure 21: Experiment 2 results.....	162
Figure 22: Experiment 3 results.....	170
Figure 23: Across experiment switch cost results.....	185

## Acknowledgements

First, I could not be more grateful to my advisor, Tobias Egner, for being such a consistent, patient, and responsive mentor and a brilliant role model in science. I always tell anyone who would listen that I have the best PI. I would also like to thank my co-advisor, Marty Woldorff, for his guidance in research and teaching and for fostering a friendly and engaging lab environment that made me feel at home early in graduate school. I also thank my other committee members for their feedback throughout developing this thesis: my committee chair, Scott Heuttel, who has been a model for undergraduate teaching, and Felipe DeBrigard, who welcomed me into his lab as a rotation student.

I have received so much help along the way from my inspirational colleagues. Thank you, Jack Dolgin, Peter Whitehead, Christina Bejjani, Khoi Vo, Eva Gjorgieva, and Matthew Bachman, for showing me the ropes when I started and for being such great friends throughout. Ken Roberts, for all the hours of your time I took up with questions about EEG. Tanya Wen, for being an amazing source of research inspiration. Candice Wang and Raphael Geddert, for your day-to-day companionship and friendship in and out of Egner Lab. I also thank members of my lovely cohort, Rachael Wright, Allie Sinclair, Kevin O'Neill, Pranjal Gupta, Wenxi Xiao, and all other members of CCN who made my time at Duke.



Finally, I want to thank my family for their unwavering love and support, especially Louis Wallen, without whom none of this would have been remotely possible.

# 1. Introduction

## 1.1 *What is cognitive flexibility?*

Cognitive control refers to the human ability to configure the cognitive system to perform specific tasks through adjustments in attentional selection, response biasing, and working memory maintenance of contextual information (Botvinick et al., 2001). Controlled (or effortful and “willed”) processes are defined in contrast to automatic (or habitual) processes that often need to be overridden to achieve correct task performance (W. Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). A classic example of this is the Stroop task, where correct performance depends on suppressing the automatic reading of words to instead complete the more effortful task of stating the color of fonts.

Control processes also enable flexible task-switching behavior, which requires participants to engage in the effortful process of activating a different set of rules defining stimulus-response associations, i.e. undergo task-set reconfiguration, in trials where the task alternates (Stephen Monsell, 2003; Rogers & Monsell, 1995). For example, while driving a car one may have to disengage from conversing with the passenger to focus on driving when maneuvering a tricky turn. After reaching a stretch of unobstructed highway, one may then judge that it is appropriate to return to the conversation again. Detecting cues in the environment that signal changes in goal states and subsequently activating the relevant task requires cognitive control.

While a large body of literature has been dedicated to delineating the effects of control, less work has focused on how the recruitment of control itself is regulated. One does not often know how much control a task requires before acting; rather it is the experience of a demanding task that prompts activation of cognitive resources. The conflict monitoring theory (Botvinick et al., 2001) proposes that cognitive control is recruited in response to the level of conflict detected in information processing pathways, that is, the degree to which strong competing representations are activated during task performance. This idea that control recruitment is commensurate with the level of experienced task demand is supported by a large body of empirical findings (for reviews, see Bugg & Crump, 2012; Duthoo et al., 2014; Egner, 2007; Egner, 2014).

Most of the evidence on control adaptations to context and task demands comes from conflict paradigms, such as the Stroop, Flanker, and Simon tasks, wherein the target stimulus or stimulus feature can be congruent or incongruent with a distractor. These paradigms often generate a main effect of congruency, wherein congruent stimuli (e.g., the word RED presented in red letters) tend to produce faster and more accurate responses than incongruent stimuli (e.g., the word RED presented in blue letters). The main effect of congruency usually interacts reliably with previous-trial congruency, such that smaller congruency effects would be found in the current trial when the previous trial was incongruent rather than congruent (see Egner, 2007 and Duthoo et al, 2014 for reviews). This well-established congruency sequence effect is commonly used to study

trial-by-trial control regulations, as it is thought to reflect upregulation of control following a high-conflict trial, which then leads to smoother conflict resolution in the current trial (Egner, 2007).

Other studies have manipulated the proportion of congruent trials associated with different temporal or physical contexts or stimuli (for reviews see Braem et al., 2019; Bugg & Crump, 2012; Chiu & Egner, 2019). For example, Lowe & Mitterer (1982) created mostly congruent or mostly incongruent temporal contexts by manipulating the proportion of congruent trials within a single block (or “list”), creating blocks wherein 0%, 50%, or 75% of trials included congruent stimuli. Other works have similarly looked at congruency effects when different physical contexts (such as the location of stimulus presentation) or specific stimulus items (such as a color word) were associated with higher/lower frequency of congruent trial occurrence. Collectively this body of work has demonstrated that people learn that specific contexts or stimuli are associated with high or low conflict and subsequently adapt attentional states to match these demands, resulting in smaller congruency effects in high conflict contexts and vice versa. Higher levels of conflict drive greater protection of working memory content, so that the irrelevant task dimension (distracters) interferes less with processing of the relevant dimension.

Just as higher levels of experienced conflict appear to drive greater shielding of the relevant task dimension, there is evidence to suggest that greater demand for

switching drives adjustments in the readiness to switch. Control processes associated with task-switching may thus conform to the same principles of frequency learning and adaptation as the processes underpinning conflict-control. Though well-documented and extensively researched in the conflict-control literature, control-adaptation effects and their associated learning processes are not yet well-established in the cognitive flexibility literature. This thesis attempts to bridge this gap by enhancing understanding of how flexibility is learned and adapted to contexts and task demands.

### **1.1.1 Task sets, switch costs, and switch-readiness**

We begin the review of the task switching literature by defining several key terms that will be used repeatedly: task-sets, switch costs, and switch readiness. “Task-set” or “task” is used interchangeably to refer to a rule that specifies a set of task-relevant stimuli or stimulus features and their associated responses (Kiesel et al., 2010; Stephen Monsell, 2003; Rogers & Monsell, 1995). Implementing a task-set is thought to require 1) the attentional selection of relevant stimulus features, 2) activation of their respective responses, and 3) shielding these stimulus-response rules from potential interference by task-irrelevant information (Dreisbach & Haider, 2008; Dreisbach & Wenke, 2011; Meiran, 2010).

Implementation of two or more task-sets is commonly studied in cued-task switching paradigms (for reviews, see Kiesel et al., 2010, 2010; Monsell, 2003; Vandierendonck et al., 2010). Figure 1 A-B portrays a typical cued task switching

paradigm where participants switch between a letter task (vowel or consonant?) and a digit task (odd or even?). A word cue informs participants of which task to activate. When implementing the letter task, participants must selectively attend the relevant stimulus feature (“A”), retrieve the correct response (vowel = right button press), and shield the letter task from interference by the irrelevant stimulus-feature (“3”) and stimulus-response associations relevant to the un-cued task (odd = left button press). Repeat trials are preceded by the same task while switch trials are preceded by the alternate task.

Cued task switching paradigms reliably find that switching between tasks is associated with greater reaction times and error rates compared to repeating the same task. This difference between switch and repeat trials is referred to as the “switch cost” (Figure 1C) (Kiesel et al., 2010; Meiran et al., 2000; Stephen Monsell, 2003; Siqi-Liu & Egner, 2020; Vandierendonck et al., 2010). The switch cost arises from a variety of different processes that make switching between two tasks more difficult than repeating a task (reviewed below in Section I). Regardless of different underlying processes, the size of the switch cost is typically assumed to be indicative of one’s current level of cognitive flexibility, or “switch readiness.” That is, when switch costs are high, flexibility is low, such that switch cost is an inverse measure of switch readiness (Braem & Egner, 2018; Dreisbach & Fröber, 2019).

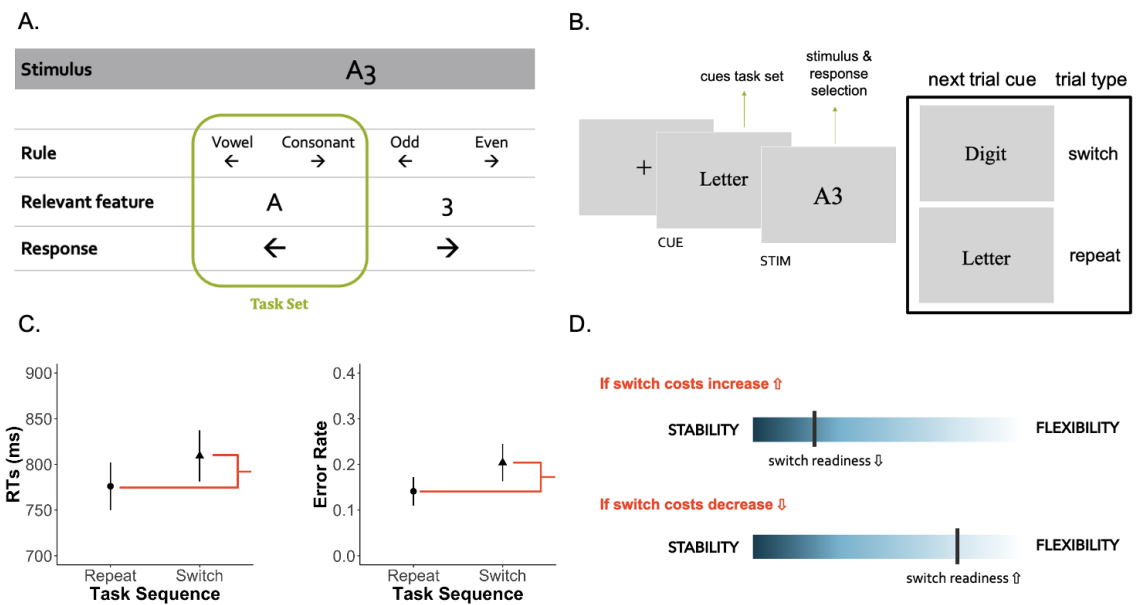
### 1.1.2 The Stability-Flexibility Dilemma

Finding a context-appropriate level of switch-readiness constitutes a central goal of adaptive flexibility. This section will introduce the stability-flexibility dilemma (Goschke, 2000, 2013) as a theoretical framework for understanding such flexibility adaptations, and the hypothesized neural underpinnings and potential clinical significance of navigating the tradeoff between stability and flexibility.

The stability-flexibility dilemma refers to the challenge of balancing the antagonistic goals of cognitive stability and flexibility (see Figure 1D). On the one hand, cognitive stability is required to shield an ongoing task from distraction to efficiently accomplish our goals; on the other hand, we also need to remain sensitive to possible changes in the environment to quickly update our cognitive strategies, requiring cognitive flexibility. Too rigid goal shielding comes at the cost of reduced flexibility and perseverance, such as when a novice driver is so focused on the task of changing into the right lane that he neglects that the traffic light has changed from green to red. Conversely, an overly flexible processing mode may render the agent easily distractible, as when a driver takes their eyes off the road to check a new notification on their phone.

Switch-readiness can be conceived of as a set-point on this stability-flexibility continuum, reflecting a meta-control parameter termed the “updating threshold” (Goschke, 2003, 2013). When switch readiness is high, task switches are less costly (flexibility is high), but this necessitates poor task-shielding against interference

(stability is low). When switch readiness is low, switching between tasks is more difficult, but in turn the current task-set is better protected against interference (but see Geddert & Egner, 2022) and performance is enhanced to a greater degree when tasks repeat (stability is high). Flexibly adjusting this meta-control parameter to match environmental demands would allow one to meet the challenges posed by the stability-flexibility dilemma.



**Figure 1: Simplified task switching paradigm** A) In an example paradigm where participants switch between letter and digit tasks, the green box highlights the rule, relevant feature, and correct response as defined by the letter task for an example stimulus “A3.” B) A typical cued task switching paradigm where a word preceding the stimulus cues the letter task set. The next trial is a repeat trial if the same task (letter in this case) is cued again; it is a switch trial if a different task (digit) is cued. C) Switch costs in response times and error rates is marked in red. Data from Siqi-Liu & Egner (2020). D) Illustration of how switch costs are thought to be inversely related to switch-readiness, a set point on the stability-flexibility continuum.



The functional tradeoff between stability and flexibility described by this dilemma is thought to be closely tied to dopamine (DA) states in the prefrontal cortex (PFC) and striatum (reviewed in Cools & D'Esposito, 2011). Deficient DA function is implicated in a wide variety of neurobehavioral disorders associated with deficits in working memory and cognitive flexibility, including Parkinson's disease, ADHD, and schizophrenia. Dopamine may fulfill qualitatively different functions in the PFC and striatum, thus contributing to the divergent symptomologies. Accumulated cross-species evidence points to the importance of optimal levels of dopamine in the PFC for working memory encoding and maintenance (Sawaguchi & Goldman-Rakic, 1991; Vijayraghavan et al., 2007). On the other hand, optimal levels of DA in the striatum may facilitate flexible updating of goal representations (Braver & Cohen, 2000; Dodds et al., 2009) via modulating (or "gating") connectivity between PFC and task-relevant representations in the posterior cortex (van Schouwenburg et al., 2010). These divergent functions of dopamine could generally be considered functionally opponent in neurotypical cognitive performance: while optimal dopamine levels in the PFC may promote maintenance of currently active working memory representations (i.e., cognitive stability) at the expense of slow updating, optimal dopamine levels in the striatum may promote a low updating threshold (i.e., cognitive flexibility) at the expense of greater distractibility (Cools & D'Esposito, 2011).

To successfully navigate this stability-flexibility dilemma, the brain needs to continuously adapt a contextually optimal level of cognitive flexibility. In other words, *switch readiness must be sensitive to contextual demands*. Greater task set protection (lower switch readiness) is suitable for stable environments with low demands for switching, while greater flexibility (higher switch readiness) is optimal for more volatile contexts involving frequent task-switches.

Measures of such adaptive flexibility (i.e., contextual changes in switch cost) may prove to be more clinically relevant than pure task-switching performance (i.e., mean switch cost), which have traditionally correlated poorly with clinical measures. While individuals with ASD exhibit cognitive flexibility deficits in daily life, including rigid behavior and difficulty in adapting to changes in daily routines (Kenworthy et al., 2010), they do not necessarily exhibit corresponding performance deficits in laboratory task-switching paradigms (Geurts et al., 2009). For example, (de Vries & Geurts, 2012) found that children with ASD do not show impaired performance compared to controls in an ecologically valid switch task which employs social and emotional stimuli. However, there is also some evidence of impaired task-switching performance (Cepeda et al., 2000; King et al., 2007) and differential brain activity during switching (Dibbets et al., 2010) in children with ADHD. Nonetheless, measuring people's ability to adjust between stability and flexibility based on contextual demands has the potential to reveal a new, and more ecologically and clinically valid, dimension of flexible behaviors.

## **1.2 Processes that contribute to switch costs**

To understand adaptive switch-readiness, this section will first cover possible processes that contribute to switch costs. Switching to a new task is thought to be more costly because it is associated with at least three different processes that repeating the same task does not require (Monsell, 2003): task-set reconfiguration, task-set inertia, and associative retrieval. An overview of these processes is followed by a discussion of the feature-integration account of switch costs, which challenges the assumption that switching between tasks involve controlled processes. Lastly, the section will be concluded with a brief discussion on which of these processes that contribute to switch costs could be affected or modulated by contextual demands.

### **1.2.1 Task-Set Reconfiguration**

Switching between tasks (or task-set updating) may require reconfiguration, that is, the act of replacing of the previously active task-set with a new set of stimulus-response rules (Meiran, 1996; Stephen Monsell & Mizon, 2006; Rogers & Monsell, 1995). Task-set reconfiguration (TSR) could include shifting attention between stimulus attributes or features (e.g., shifting from the letter attribute “A” to the digit attribute “3”), and retrieving goal states and relevant response mappings into working memory (Monsell, 2003).

The *preparation effect* is often taken as evidence of the contribution of TSR to switch costs (Meiran, 1996; Monsell & Mizon, 2006). Meiran (1996) manipulated the

interval between the task cue and stimuli (cue-stimulus interval, or CSI), giving participants either more or less time to prepare for a task-switch or repeat. Across five experiments, Meiran (1996) demonstrated that switch costs were greatly reduced when participants were given sufficient time to prepare for an upcoming task in the long CSI conditions. This preparation effect (also referred to as the reduction in switch cost effect in Monsell & Mizon, 2006) held true even when controlling for the response-stimulus interval (RSI), or the time between current trial stimulus onset and the response to the previous trial, indicating that it is due to proactive preparation for the upcoming task before task-execution rather than passive, dissipating interference from the previous trial (Allport et al., 1994).

Further increasing the CSI to 1.2 seconds, Rogers & Monsell (1995) demonstrated that switch cost reductions asymptote and switch costs are not completely abolished even with ample time to prepare. This persisting switch cost is sometimes referred to as *residual cost*. One potential account for residual switch costs is that some aspect of TSR cannot be completed until exogenously triggered by task-relevant stimulus attributes (Rogers & Monsell, 1995), in contrast to other aspects that could be accomplished endogenously, or before stimulus onset. An alternative account, the intentional-activation hypothesis (De Jong, 2000) suggests that there is no distinction between exogenous and endogenous switch costs; rather, residual switch costs arise from a mix of trials where subjects successfully complete TSR before stimulus onset and trials where

they 'fail to engage' and must perform TSR after stimulus onset. The latter account is supported by discrete state mixed-model results of TSR as an all or none process in prepared switch RTs (De Jong, 2000) and is consistent with the idea that TSR involves long-term memory retrieval attempts of task-rules (Mayr & Kliegl, 2000).

### **1.2.2 Task Set Inertia**

Second, on task switch trials one may have to inhibit interference from the previously active task set (also conceived as overcoming "task-set inertia") (Allport et al., 1994; Stephen Monsell, 2003). Studies have shown that, surprisingly, when switching between one dominant, well-practiced, or automatic task (e.g., word reading in a Stroop paradigm) and a nondominant, effortful task (e.g., color naming), the dominant task is associated with higher switch costs than the nondominant task (Allport & Wylie, 1999; Allport et al., 1994; Arbuthnott, 2008; Ellefson et al., 2006; Wu et al., 2015; Yeung & Monsell, 2003).

A TSR account alone cannot explain this switch-cost asymmetry, since the well-practiced dominant task set should be easier and faster to activate. Allport et al. (1994) proposed that performance of the nondominant task is associated with strong inhibition of the dominant task; this inhibition then has to be overcome, resulting in larger switch costs when switching to the dominant task in subsequent trials. The idea that the easier (more automatic) task is inhibited has since been disputed by findings of reduced asymmetric switch costs after implementing certain controls for episodic memory effects

(Stephen Monsell et al., 2000; D. W. Schneider & Anderson, 2010; Wong & Leboe, 2009; Yeung & Monsell, 2003). However, the consensus remains that activation of a task set persists after trial termination (hence exhibiting 'inertia') and can influence performance on the subsequent trial. Evidencing this persisting activation, studies that employ variable RSIs typically find that interference from the previous task passively dissipates with increased time between the current and previous trial (Meiran et al., 2000; Ruthruff et al., 2001; Sohn & Anderson, 2001). The above-mentioned controls for RSI when manipulating cue-stimulus interval (CSI) were implemented to separate the effects of task-set reconfiguration and task-set inertia.

### **1.2.3 Associative Retrieval of Task-Sets**

Lastly, *associative retrieval* of all task-sets that were previously associated with a given stimulus can interfere with the execution of the current task set (Allport & Wylie, 2000; Allport et al., 1994; Waszak et al., 2003). Associative retrieval refers to the idea that stimuli become associated with the tasks in which they occur. Interference from these task-set associations contributes to task-mixing costs, or behavioral costs associated with mixed- compared to single- task blocks regardless of switch or repeat trial identity (Allport & Wylie, 2000; Waszak et al., 2003). Crucially, Waszak et al. (2003) suggests that associative interference also contributes to switch costs (assuming that the same stimuli or stimulus categories are shared by the tasks). Researchers found that stimuli presented in both tasks were associated with higher switch costs compared to those

presented in only one task. Further, this associative interference is strongest when current task activation is weak, such as on switch trials. These results suggest that, in tasks with reoccurring stimuli, one must overcome interference from other task-sets associated with the stimuli to enact the current task-set.

#### **1.2.4 Feature integration account of switch costs**

Associative interference is thought to predominantly operate after stimulus onset, unlike TSR (Monsell, 2003). A traditional challenge to interpreting switch costs as requiring control implementation is that the same behavioral results cited as evidence for supposed control recruitment could be explained via stimulus-driven associative learning (or contingencies) without alluding to under-specified higher-level functions (Hommel, 2004; Schmidt & De Houwer, 2011; Schmidt & Lemercier, 2019).

For example, in the realm of conflict-control Hommel et al. (2004) demonstrated that the congruency sequence effect could be attributed to stimulus and response-feature integration effects (though see also Akçay & Hazeltine, 2007, 2011; Hazeltine et al., 2011; Ullsperger et al., 2005). Namely, transitions from congruent to incongruent trials are more likely to involve partial, rather than complete, repetitions of stimuli features and responses compared to incongruent-congruent, incongruent-incongruent, or congruent-congruent transitions.

This *feature-integration account* (Hommel, 2004) builds upon prior research on “event files,” or episodic memory representations of stimulus and response features that

co-occur in time or on any given trial (Hommel, 1998). Event file theory posits that, when any given feature in an event file reoccurs, the associated response features in the same file are primed. This implies that performance is facilitated on trials involving complete repetitions or complete alterations of stimulus features and responses, and performance is impaired on trials involving partial repetitions, or when several but not all stimulus and response features repeat (Hommel, 1998; Mayr et al., 2003). Event file theory is also directly relevant to the associative retrieval account of switch costs, as the underlying assumption for the associative retrieval account is that stimuli and co-occurring tasks are bound in an event file, such that reencountering a particular stimulus automatically retrieves associated task-sets.

In the task switching literature, similar attempts have been made to explain switch costs in terms of feature-integration biases, rather than relying on the higher-order notion of task set reconfiguration and activation. Notably, Schmidt & Liefoghe (2016) systematically examined an extensive set of possible sources of feature integration biases. First, task-repetitions could be confounded with cue-repetitions, which would render repeat trials easier simply because additional cue-processing is unnecessary. Schmidt & Liefoghe (2016, Exp 2) controlled for cue-repetition by assigning two cues to each task such that cues alternate on every trial even while task repeats. The authors found that switch costs were reduced after controlling for cue-repeats, but not



completely abolished (for similar results, see also Arrington & Logan, 2004; Logan & Bundesen, 2003, 2004; Logan & Schneider, 2006; D. W. Schneider & Logan, 2005).

In contrast to conflict tasks, task-switching paradigms are comparatively immune to simple stimulus-response feature integration effects, because task, stimulus, and response switches can occur independently. That is, a task switch does not have to be accompanied by a stimulus or response switch. However, a stimulus repetition always entails a response repetition when task repeats (complete repetition) and a response alternation when task switches (partial repetition). Lastly, switch costs could also be confounded by conceptual response repetitions, or repetitions of the correct stimulus category instead of the actual button press. For example, an odd → consonant categorization is a conceptual response alternation even if odd numbers and consonants are associated with the same physical button press. Conceptual repetitions are only possible in task repeats, creating yet another level of repetition that could confer an advantage to repeat trials. Like before, Schmidt & Liefoghe (2016, Exp 1) found that controlling for both stimulus- physical and conceptual response integration biases reduced but did not eliminate switch costs.

Evidence of persistent switch costs even after rigorous controls for feature-integration biases suggests that some notion of task-set reconfiguration and activation remains necessary for explaining why task-switches are more difficult than repeats. The tentative consensus in the current literature is that, at the very least, preparation effects

(Meiran, 1996; Stephen Monsell & Mizon, 2006) constitute relatively incontrovertible evidence that task-switching performance benefits from endogenous task-set reconfiguration processes, implying that switching between tasks cannot be explained by variance in stimulus-feature repetitions exclusively.

### **1.2.5 Switch cost components that may be sensitive to context**

Collectively, task-set reconfiguration, task-set inertia, and associative interference are thought to contribute to switch costs. Task-set inertia is thought to dissipate passively (Meiran et al., 2000; Ruthruff et al., 2001; Sohn & Anderson, 2001) and associative interference and its resolution occur after stimulus onset. Thus, TSR, which is thought to reflect proactive preparation for the upcoming task-set, is most likely to be sensitive to *strategic* modulations of attentional modes in response to task and context demands. However, switch-readiness adjustments do not necessarily have to occur in a top-down manner. As this thesis will review in the upcoming sections, it is also possible for context-appropriate switch readiness to be cued in a bottom up manner by tasks (Siqi-Liu & Egner, 2020) or stimuli (Chiu & Egner, 2017). Switch readiness modulations could thus be mediated by any one of the processes underlying switch costs, and which one plays a larger role may depend on the specifics of the task protocol.

## 2. Contextual adaptation of cognitive flexibility is driven by task- and item-level learning

As outlined in Chapter 1, successfully navigating the shifting-shielding dilemma can be conceptualized as learning to strategically adjust one's updating threshold to suit changes in environmental demand for relatively more or less cognitive flexibility (Goschke, 2003). Importantly, behavioral evidence for these types of dynamic adjustments in switch readiness has been obtained in cued task-switching protocols that manipulate the frequency (and thus, likelihood) of switch trials between blocks of trials. Specifically, several studies have shown that the magnitude of switch costs tends to scale inversely with the frequency that task switches occur within a given block of trials (Bonnin, Gaonac'h, & Bouquet, 2011; Dreisbach & Haider, 2006; Dreisbach, Haider, & Kluwe, 2002; Duthoo, De Baene, Wühr, & Notebaert, 2012; Mayr, 2006; Monsell & Mizon, 2006) or at a specific spatial location (Crump & Logan, 2010; Leboe et al., 2008). For instance, Monsell & Mizon (2006, Experiment 4) varied switch proportions from 25% to 50% to 75% between blocks of trials and observed the greatest switch costs at a switch frequency of 25% and the smallest switch costs at a switch frequency of 75%. We will here refer to these block-based modulations of switch cost as the list-wide proportion switch effect (LWPSE), leaning on a similar nomenclature in the congruency effect literature (e.g., Bugg & Chanani, 2011; Bugg & Crump, 2012). While the above demonstrations of a LWPSE provides basic evidence that people can adapt their switch

readiness to varying task statistics, the exact scope of this adaptation, as well the particulars of the underlying learning processes, are presently not known.

In the present study, we ask in particular what kind of learning drives these effects, and we distinguish between three ways in which changes in updating threshold could become associated with features of low vs. high frequency switch blocks: the list-wide level (producing sustained and generalizable changes in flexibility), the task-set level (where a particular level of switch readiness becomes associated with a specific task-set), and the item level (where a particular level of switch readiness becomes associated with specific task stimuli). To investigate the kinds of learning that drive the LWPSE, we ask several questions that have not been previously addressed in the literature: First, because previous studies that found these context-sensitive switch cost modulations only required that participants switch between two tasks, it is not clear to what degree the LWPSE reflects a generic change in cognitive flexibility or task-specific preparation processes. In other words, reduced switch costs in high proportion switch blocks could reflect participants preparing for the particular alternate task, rather than general preparation for a task switch (to any other task). Intuitively, the latter would stand as stronger evidence for a genuine adjustment of cognitive flexibility, since flexible engagement with a changing environment requires increased aptitude to respond to events that are often unexpected.

Second, it is not yet known to what extent the LWPSE is driven by associating switch readiness with the global switch likelihood of the current block context (list-wide learning), or by using the specific task-sets and/or task stimuli (also referred to as “items”) as cues for adjusting switch readiness: in prior studies, in high switch frequency blocks all tasks and all task stimuli were also presented more frequently as switch versus repeat trials (and *vice versa* for low switch frequency blocks). Therefore, any reductions in switch costs that were observed could have resulted from participants’ learning of task- and/or item-specific associations with switch frequencies instead of linking the temporal, list-wide context to a greater need for flexibility.

In the current paper, we present a series of four experiments that shed light on these unanswered questions about the scope and mechanisms of meta-control over the stability-flexibility tradeoff, as indexed by the LWPSE. Experiment 1 attempts to replicate the LWPSE using the design of Monsell & Mizon (2006) with a different stimulus set. Using three instead of two tasks, Experiment 2 tests whether the LWPSE is preserved when participants do not know which specific alternate task they will switch to. To tease apart list- and task- level biases, Experiment 3a and 3b probed for the generalization of the LWPSE to an unbiased “transfer task” that occurred equally often as switch and repeat trials, presented in blocks with overall high or low switch bias. Following a similar logic, Experiment 4 used switch proportion biased versus unbiased stimuli to investigate whether the LWPSE can be observed in the absence of item-level

biases. The data and materials for all experiments are available at <https://osf.io/5cxam/>, and none of the experiments were preregistered.

## **2.1 Experiment 1**

The first experiment was a conceptual replication of Monsell & Mizon's (2006) Experiment 4. We sought to replicate the switch proportion dependent switch cost to validate a basic task protocol with which to assess the determinants of the LWPSE in the subsequent experiments. Specifically, participants performed cued letter and digit categorization tasks under within-subject manipulations of task sequence (task repeat vs. task switch trials), CSI (short: 190 ms or long: 840 ms) and the proportion of switch trials per block (30%, 50%, or 70%). The CSI factor was included because the pattern of results in prior work suggested that the LWPSE may be CSI-dependent, with maximal effects of switch proportion evident at short CSIs (cf. Monsell & Mizon, 2006). In other words, participants may rely more on context in aiding their task-set updating strategy when they have less time to utilize the trial-by-trial cue for task set reconfiguration. We therefore expected to find reduced switch costs with an increasing proportion of switch trials to be most pronounced in the short CSI condition.

### **2.1.1 Method**

#### **2.1.1.1 Participants**

A power analysis based on the effect size of the smallest switch cost modulation (switch cost difference between the 50% and 75% switch condition) in Monsell & Mizon

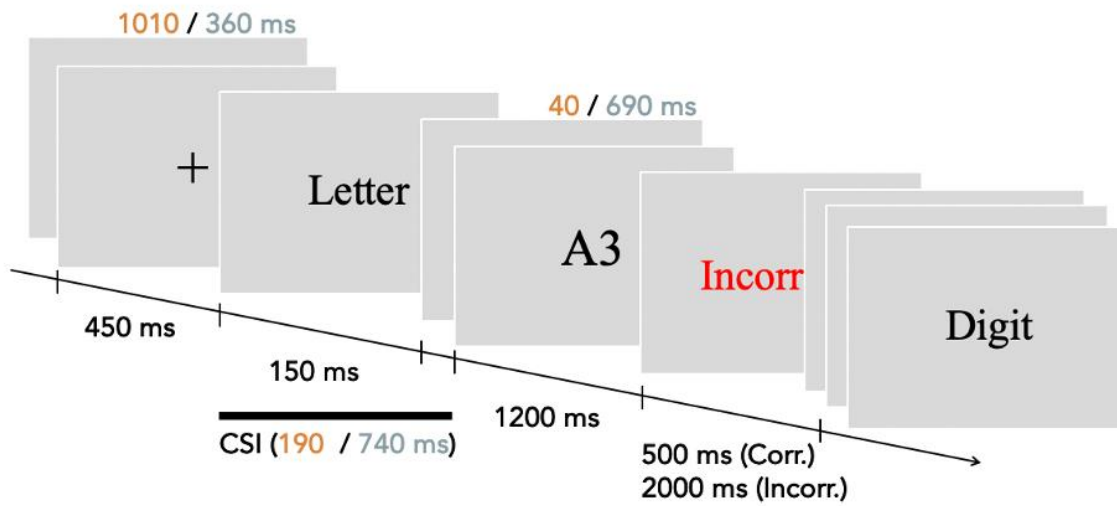
(2006, Experiment 4) suggested a total sample size of 26 to achieve 0.95 power. To be conservative and to take into account larger participant exclusion rates for online testing, we roughly doubled this estimate and recruited 56 participants from MTurk. The experiment lasted ~60 minutes and 16 participants were excluded from data analysis for lower than 75% overall accuracy on the task, leaving a final sample size of 40.

### **2.1.1.2 Stimuli**

Task stimuli consisted of a letter and a digit displayed simultaneously at either side of the center of the screen for each trial. The letter was randomly selected from 'A', 'E', 'I', 'U', 'G', 'K', 'M', or 'R' and the digit was randomly selected from '2', '3', '4', '5', '6', '7', '8', or '9'. Whether the letter or the digit was presented on the left or right was randomized across trials.

### **2.1.1.3 Procedure**

Experiment procedures roughly followed experiment 4 of Monsell & Mizon (2006). Each trial began with a blank interval of 1010 ms (short CSI condition) or 360 ms (long CSI condition), followed by a 450 ms long fixation display, a cue display lasting 150 ms, and another blank interval of either 40 ms (in the short CSI condition) or 690 ms (in the long CSI condition). Finally, the task stimuli appeared and remained on screen for 1200 ms. The lengths of the blank intervals were varied so that the RSI, or the sum of the blank intervals, fixation display, and cue display, was a constant 1650 ms for both short and long CSI trials (see Figure 2).



**Figure 2: Experimental protocol.** The basic trial structure of Experiments 1-4 involved a task cue, followed by a cue-to-stimulus-interval (CSI) preceding a target stimulus display, followed by accuracy feedback.

Participants were required to perform a letter classification task (“Is the letter a vowel or consonant?”) if they saw the cues “Letter” or “Alphabet” and to perform a digit classification task (“Is the digit odd or even?”) if they saw the cues “Digit” or “Number.” The 2:1 cue-to-task mapping allowed us to change the cue on every trial, regardless of whether the task was switched or repeated, thus eliminating the contribution of possible response time benefits that come from repeating cues on task repeat trials (e.g. Logan & Bundesen, 2003; Mayr & Kliegl, 2003) to our computation of task switch costs. Participants had to press the ‘d’ or ‘k’ key to categorize the stimuli as vowel/consonant or odd/even. Participants were randomly assigned to different response mappings for each task. Correct responses were followed by a 500 ms blank screen, and incorrect responses were followed by the word “ERROR” displayed for 2000



ms. Responses made while the task stimulus was not onscreen were considered incorrect.

Each participant completed 18 blocks of 31 trials. All trials except the first in each block were coded as belonging to either the task switch (preceded by a different task) or task repeat (preceded by the same task) condition. The percentage of switch trials per block was 30, 50, or 70; there were six blocks of each switch proportion condition. The trial sequence for each block was generated pseudo-randomly according to an algorithm that ensured each task was presented an approximately equal number of times. In the 30% switch block, each task was presented either 4 or 5 times as a switch trial and 11 or 10 times as a repeat trial, creating a 9:21 switch to repeat ratio. In the 70% switch block, the number of switch v. repeat trials per task was reversed, creating a 21:9 switch to repeat ratio; in the 50% switch block, each task was presented either 7 or 8 times as repeat and as switch trials. For a table depicting switch/repeat frequencies for each task in this and subsequent experiments, refer to Appendix A, Table 7. All 6 blocks of the same switch proportion were presented consecutively to increase the saliency of the switch/repeat context, but the presentation order of the chunk of blocks with the same switch proportion was counterbalanced across participants. CSI alternated from block to block beginning with the short CSI. Before starting the main experiment, participants completed two short CSI blocks and one long CSI block for practice. All practice blocks had 50% switch proportion.

#### 2.1.1.4 Design

The experiment followed a 2 (task sequence: switch vs. repeat)  $\times$  2 (CSI: long vs. short)  $\times$  3 (switch proportion: 30% vs. 50% vs. 70%) repeated-measures factorial design.

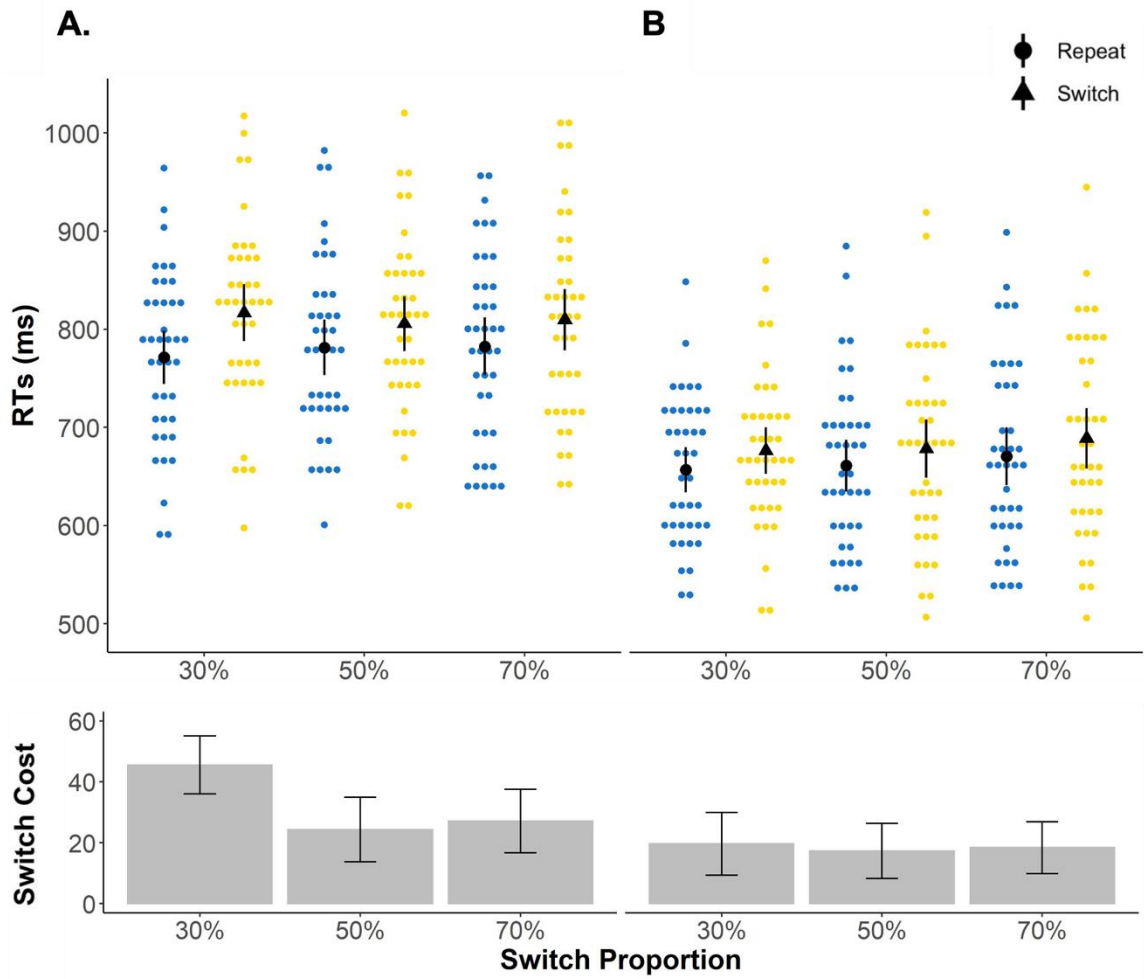
#### 2.1.2 Results and Discussion

For assessing performance accuracy, we analyzed data from all trials after excluding practice blocks and the first trial of each block. For RT analyses, we additionally excluded incorrect trials, and trials following incorrect trials. After applying these exclusion criteria, trials with response times (RT) outside 1.5 times the interquartile RT range of the remaining sample were filtered out for the RT analyses. Descriptive statistics are displayed in Table 1. Excluded trial counts and the number of remaining trials per smallest and largest cells are included in Appendix A, Table 6.

We ran a repeated-measures analysis of variance (ANOVA) with the independent variables of task sequence (switch vs. repeat), CSI (long vs. short), and switch proportion (30% vs. 50% vs. 70%). Replicating classic effects in the task switching literature, we observed a main effect of task sequence (i.e., switch costs), as reflected in slower RTs for switch trials ( $M_{\text{switch}} = 734.59$  ms) compared to repeat trials ( $M_{\text{repeat}} = 712.15$  ms),  $F(1,39) = 84.02$ ,  $p < .0001$ ,  $\eta_p^2 = .68$ ; a main effect of CSI, as short CSIs yielded longer RTs ( $M_{\text{short}} = 784.74$  ms) than long CSIs ( $M_{\text{long}} = 665.79$  ms),  $F(1, 39) = 312.72$ ,  $p < .0001$ ,  $\eta_p^2 = .89$ ; and a task sequence  $\times$  CSI interaction ( $F(1, 39) = 10.77$ ,  $p = .002$ ,  $\eta_p^2 = .22$ ), wherein

short CSI trials produced larger switch costs ( $M_{\text{switchcost}} = 32.33$  ms) than long CSI trials ( $M_{\text{switchcost}} = 18.45$  ms).

More crucial to the focus of the current study, there was a significant interaction effect of task sequence  $\times$  switch proportion,  $F(2, 78) = 3.87$ ,  $p = .02$ ,  $\eta_p^2 = .09$ , as switch cost was greater in the 30% switch condition ( $M_{\text{switchcost}} = 32.6$  ms) than in the 50% ( $M_{\text{switchcost}} = 20.84$  ms) and 70% ( $M_{\text{switchcost}} = 22.73$  ms) switch conditions. Moreover, as anticipated, the interaction effect of task sequence  $\times$  switch proportion was driven by a modulation of switch cost by switch proportion in the short but not in the long CSI condition (see **Figure 3**), as supported by a three way interaction between task sequence  $\times$  CSI  $\times$  switch proportion ( $F(2, 78) = 3.14$ ,  $p = 0.05$ ,  $\eta_p^2 = .07$ ). Post hoc tests revealed that, in the short CSI condition, switch cost for the 30% switch condition ( $M_{\text{switchcost}} = 45.56$  ms) was significantly larger than the 50% switch condition ( $M_{\text{switchcost}} = 24.34$  ms,  $p = .004$ ) and the 70% switch condition ( $M_{\text{switchcost}} = 27.10$  ms,  $p = .025$ ). On the other hand, in the long CSI condition, there were no significant switch cost differences between any of the three different switch proportions ( $p = 1$ ). No other main or interaction effects were significant.



**Figure 3: Experiment 1 mean RTs and switch costs in the (A) short CSI and (B) long CSI Conditions. Upper panels depict group mean RTs in black circles (repeat) or triangles (switch) and individual mean RTs in blue (repeat) or yellow (switch) dots. Lower panels depict mean switch costs in bars. All error bars indicate confidence intervals ( $1.96 \times$  standard error). For trials with short CSI, switch cost in the 30% switch condition was significantly greater than the 50% and 70% conditions. Trials with long CSIs did not show significant switch cost moderations by switch proportion.**

An identical ANOVA was run on subject's mean accuracies. There was an expected main effect of task sequence ( $F(1,39) = 28.63, p < .0001, \eta_p^2 = .42$ ) as participants performed with lower accuracy on switch ( $M_{\text{accuracy}} = 0.83$ ) compared to repeat trials

( $M_{\text{accuracy}} = 0.88$ ). We also observed a main effect of CSI, as trials with shorter CSI periods ( $M_{\text{accuracy}} = 0.83$ ) produced significantly lower accuracy rates ( $F(1,39) = 34.63, p < .0001, \eta_p^2 = .47$ ) compared to long CSI trials ( $M_{\text{accuracy}} = 0.89$ ). There was also a significant task sequence  $\times$  CSI interaction ( $F(1,39) = 6.19, p = .02, \eta_p^2 = .14$ ), wherein short CSI trials were associated with larger accuracy switch costs (repeat - switch) ( $M_{\text{switchcost}} = 0.05$ ) than long CSI trials ( $M_{\text{switchcost}} = 0.03$ ). Unlike the RT data, all other effects were nonsignificant.

Experiment 1 successfully replicated the key results of Experiment 4 of Monsell & Mizon (2006) in RTs: switch costs were reduced in conditions where switching was more frequent, but only for short CSIs. However, it is noteworthy that, unlike in Experiment 4 of Monsell & Mizon (2006), switch cost reductions in the current experiment seemed to be mainly driven by increases in repeat trial RTs, rather than decreases in switch trial RTs. This pattern of results is observed across all four experiments in this study and is discussed in depth in the General Discussion, where a probable explanation of the lack of switch trial RT improvements is offered.

Another caveat to interpreting our results is that switch frequency may be confounded with run length, i.e., the number of consecutive task repeats (Bonnin et al., 2011). Since run lengths are longer in low switch frequency blocks, repeated exposure to the same task could promote within-run RT speeding and produce greater task-set inertia that requires more laborious inhibition when participants finally encounter a switch trial, leading to RT slowing. However, finding switch cost adjustments even after

restricting their analysis to the first three positions in a run, Bonnin et al. (2011) demonstrated that run length is not the primary contributor to LWPSE.

Nonetheless, the switch cost reductions we observed suggest that participants employ the statistics of control demands – the incidence of switch trials in the different blocks – to guide their cognitive strategies. Moreover, this context-sensitive adjustment in switch-readiness is only evident under conditions where the cue-to-target interval is too short to engage in substantial task-set reconfiguration prior to target onset on a trial-by-trial basis. In the following experiments, we sought to characterize more closely the scope and sources of this form of learned cognitive flexibility.

## ***2.2 Experiment 2 Introduction***

Experiment 1 replicated the basic effect of switch proportion (Monsell & Mizon, 2006), demonstrating that the costs of switching are lower when switches are more frequent. However, since participants were only switching between two tasks, they always knew which particular task they would be switching to when they expected a switch. Therefore, the results of Experiment 1 may reflect specific preparation for the particular alternate task instead of a general adjustment of switch-readiness or cognitive flexibility. To test whether the changes in switch cost reflect a modulation of generalizable switch readiness rather than better preparation for a specific alternate task, we adapted the design of Experiment 1 to involve three tasks instead of two (see Chiu & Egner (2017) for an equivalent approach in the context of an item-specific switch

proportion manipulation). If switch cost were still moderated by switch proportion in Experiment 2, this would constitute evidence that participants can use context to facilitate task switching even when they do not know what task they are switching to. That three task paradigms make actively anticipating the upcoming task more difficult in turn implies that inhibition (or lack thereof) of the previous task (Mayr & Keele, 2000) should have larger influences on the size of switch costs than anticipatory task-set reconfiguration (Rogers & Monsell, 1995; Monsell, 2003).

In Experiment 2, a color classification task was included as the third task, in addition to the letter and digit tasks. Participants were cued from trial-to-trial as to which of the three tasks to perform. Only a short CSI (200 ms) was used because Experiment 1 demonstrated that a long CSI eliminated the effect of switch proportion on switch costs. Additionally, only 30% and 70% switch blocks were used, because the switch cost difference between 50% and 70% switch proportion blocks was nonsignificant in Experiment. The 50% condition is also more difficult to compare to the other two conditions, as that condition has a greater level of overall task uncertainty (0.5 in a two-task design) compared to the 30% and 70% blocks, which are equated in terms of uncertainty (0.3 in a two-task design).

## **2.2.1 Method**

### **2.2.1.1 Participants**

58 participants were recruited from MTurk. The experiment lasted ~30 minutes, and participants were compensated \$3.00 if they performed at 50% accuracy or above. 17 participants were excluded from data analysis for lower than 65% overall accuracy on the task, leaving a final sample size of 41. The accuracy threshold for participant inclusion was modified from Experiment 1 (where it was 75%) due to the increased difficulty associated with switching between three rather than two tasks.

### **2.2.1.2 Stimuli**

Task stimuli were the same as before, except that a colored frame was placed around the letter/digit pair. The color of the frame was randomly selected from four warm colors (shades of red, orange, and yellow) and four cold colors (shades of green, blue, and purple).

### **2.2.1.3 Procedure**

Experiment procedures were largely identical to Experiment 1, with a few changes to timing in order to accommodate the third task and the increased difficulty of switching between three tasks. First, 10 ms were added to the cue display period, bringing the CSI to 200 ms. Second, the target stimulus display period was increased by 100 ms (totaling 1300 ms) to give participants more time to respond (as before, correct responses had to be made while the stimuli were on screen). Finally, cued by the word



“Color” or “Hue,” the color task was added requiring participants to indicate whether the color of the frame was cold or warm (cf. Cooper et al., 2019). As in Experiment 1, the cue was always switched between consecutive trials. Participants were randomly assigned to different response mappings for each task. The experiment consisted of 8 blocks of 31 trials each. Four blocks had a 30% switch proportion and the other four had a 70% switch proportion. Trial sequence and the block presentation order were generated in the same way as in Experiment 1.

#### **2.2.1.4 Design**

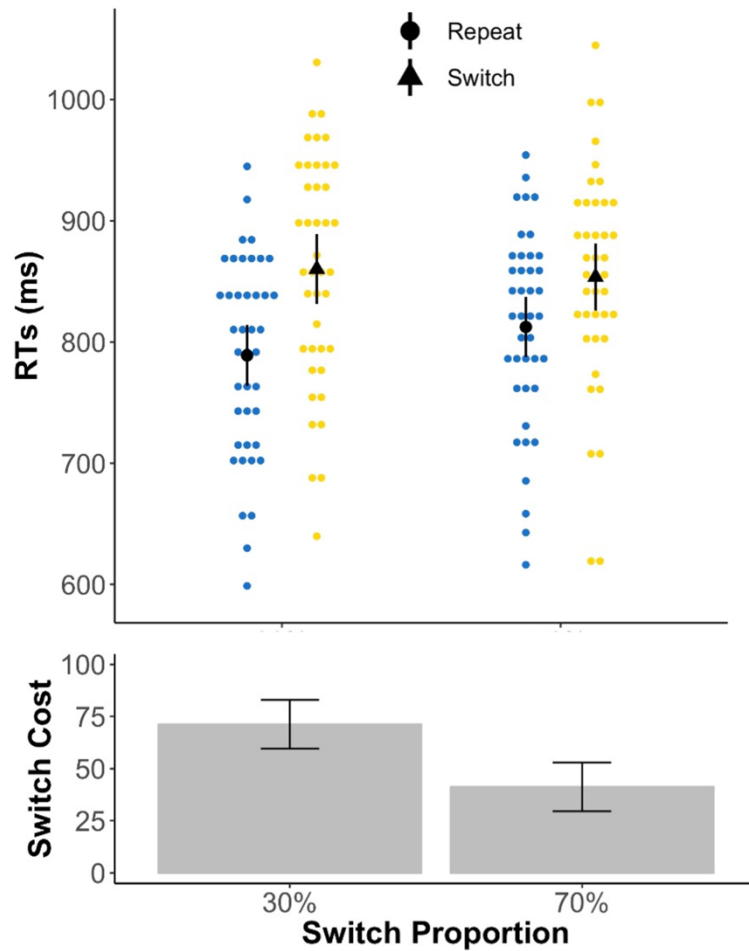
The experiment followed a 2 (task sequence: switch vs. repeat) x 3 (switch proportion: 30% vs. 70%) repeated-measures factorial design.

#### **2.2.2 Results and Discussion**

The same exclusion criteria as in Experiment 1 were applied. Descriptive statistics are displayed in **Table 1**, side-by-side with those of Experiment 1.

We ran a repeated-measures ANOVA with the independent variables of task sequence (switch vs. repeat) and switch proportion (30% vs. 70%). The main effect of task sequence on RT (i.e., the switch cost) was significant,  $F(1, 40) = 149.59$ ,  $p < .0001$ ,  $\eta_p^2 = .79$ , as participants responded slower on switch ( $M_{\text{switch}} = 852.07$  ms) compared to repeat trials ( $M_{\text{repeat}} = 795.26$  ms). There was no main effect of switch proportion. Crucially, we found an interaction effect between task sequence and switch proportion,  $F(1, 40) = 15.60$ ,  $p = .0003$ ,  $\eta_p^2 = .28$ . Switch cost was significantly higher in the 30%

( $M_{\text{switchcost}} = 71.26$  ms) than in the 70% switch condition ( $M_{\text{switchcost}} = 41.26$  ms) (see **Figure 4**).



**Figure 4: Experiment 2 mean RTs and switch costs collapsed across three tasks. Upper panel depicts group mean RTs in black circles (repeat) and triangles (switch) and individual mean RTs in blue (repeat) and yellow (switch) dots. Lower panel depicts mean switch costs in bars. All error bars indicate confidence intervals ( $1.96 \times$  standard error). Switch cost in the 30% switch condition was significantly greater than the 70% condition.**

The accuracy analysis only found a significant effect of task sequence where accuracy for switch trials ( $M_{\text{switch}} = 0.78$ ) was lower than repeat trials ( $M_{\text{repeat}} = 0.86$ ),

$F(1, 40) = 64.10, p < .0001, \eta_p^2 = .62$ . The main effect of switch proportion and the interaction of task sequence and switch proportion that were significant in the RT data were nonsignificant in the accuracy data.

In Experiment 2, the three-task design makes anticipating switching to one particular alternate task impossible. The observation of a significant modulation of switch cost by switch proportion therefore provides evidence that the reduction of switch cost in high switch blocks is not limited to situations in which the task that will be switched to is known. Thus, Experiment 2 indicates that the LWPSE cannot be explained via trial-by-trial preparation for a specific alternative task.

**Table 1: Response Times (ms) and Accuracy (%) with Standard Errors in Experiment 1 and 2 as a Function of CSI and Switch Proportion**

	<i>Exp 1 (2 tasks)</i>						<i>Exp 2 (3 tasks)</i>			
	30%		50%		70%		30%		70%	
	<i>RT</i>	<i>Acc</i>	<i>RT</i>	<i>Acc</i>	<i>RT</i>	<i>Acc</i>	<i>RT</i>	<i>Acc</i>	<i>RT</i>	<i>Acc</i>
<b>190 ms CSI</b>										
<b>Switch</b>	816.9 <sup>1</sup> (14.9)	81.9 (1.8)	805.7 (14.3)	79.4 (2.5)	809.7 (15.9)	78.7 (2.4)	860.2 (14.7)	78.0 (1.8)	853.7 (14.1)	78.5 (1.6)
<b>Repeat</b>	771.4 (13.8)	87.42 (1.5)	781.3 (14.4)	84.9 (2.2)	782.6 (15.1)	84.0 (2.1)	788.9 (12.8)	86.5 (1.6)	812.4 (12.6)	85.2 (1.7)
<b>740 ms CSI</b>										
<b>Switch</b>	676.5 (12.2)	88.1 (2.3)	678.4 (15.2)	87.2 (2.0)	688.9 (15.8)	85.9 (2.0)				
<b>Repeat</b>	656.8 (11.7)	90.8 (1.8)	661.1 (13.4)	89.7 (1.6)	670.5 (15.0)	90.1 (1.8)				

<sup>1</sup> Data refers to group mean RTs (excluding error trials, trials following error trials, and RT outliers) and percentage accurate with SE in parentheses.

*Note.* Data refers to group mean RTs (excluding error trials, trials following error trials, and RT outliers) and percentage accurate with SE in parentheses.

### **2.3 Experiment 3a Introduction**

Thus far, we have demonstrated that switch readiness can be adjusted based on the frequency of switches within blocks (Experiments 1 and 2), and that reduced switch costs under high switch frequency do not depend on there being only a single alternate task to switch to (Experiment 2). However, it is unclear whether these adjustments in switch-readiness reflect learning processes that are tied to the specific tasks, which all occurred at biased switch frequencies, or whether participants instead entered some kind of global flexible state in a high switch proportion context, which should be transferable to tasks that are not biased in their switch frequency. Experiments 1 and 2 could not answer this key question because the 30% and 70% switch blocks were biased both at the block (list-wide) level *and* at the task level. That is, in blocks where the overall proportion of switch trials was higher, each task was also presented more frequently as switch rather than repeat trials. (The reverse was true for low proportion switch blocks.) So, for instance, even though in high switch proportion blocks in Experiment 2 participants could not anticipate which task they would be switching to, any one of the three tasks had in fact been encountered more frequently as switch than repeat trials in that block.

Though paradigms that associate biased switch proportions with spatial locations (Leboe et al., 2008, Exp 2; Crump & Logan, 2010) are not contaminated by task- and item-level biases (i.e. tasks and stimuli are presented equally likely across the two biased spatial locations), list-wide paradigms are not controlled in the same way. Our novel manipulations allow us to investigate whether contextual adaptation to list-wide temporal context can occur when we control for task (Exp 3a and 3b) and item (Exp 4) bias.

To find out whether there is a global, transferable effect of block-wise switch proportion that is independent of task-level switch proportion, in Experiment 3a we introduced an unbiased third task, presented alongside two heavily biased tasks that drove the 30% or 70% switch proportion in each block. The unbiased task always appeared equally often as switch and repeat trials, regardless of the block-wise switch proportion (for an equivalent approach in the realm of conflict-control, see (Bugg, 2014; Bugg & Chanani, 2011; Hutchison, 2011)). If participants do enter a globally more cognitively flexible state in high switch proportion blocks, then the unbiased task should also show reduced switch cost. On the other hand, if the switch proportion effect operates at the level of specific task sets, then we should not observe a transfer of this effect to the unbiased task.

## **2.3.1 Method**

### **2.3.1.1 Participants**

59 participants were recruited from MTurk. The experiment lasted ~45 minutes and participants were compensated \$4.50 if they performed at 50% accuracy or above. 18 participants were excluded from data analysis for lower than 65% overall accuracy on the task, leaving a final sample size of 41.

### **2.3.1.2 Stimuli and Procedures**

Task stimuli, trial components, and timing were identical to Experiment 2. Participants were randomly assigned to different response mappings for each task and the identity of the unbiased task (i.e. whether it was the letter, digit, or number task).

The experiment consisted of 16 blocks of 31 trials each. The number of blocks was doubled compared to Experiment 2 in order to ensure that each participant was exposed to 160 trials of the unbiased task (there were 10 trials per block). Eight blocks had a 30% switch proportion and eight blocks had a 70% switch proportion. Trial sequences were pseudo-randomly generated. In the 30% switch proportion block, the two biased tasks were each presented twice as switch trials and 8 times as repeat trials and the unbiased task was presented 5 times as switch and 5 times as repeat trials, creating an overall switch to repeat ratio of 9:21. In the 70% switch proportion block, the number of switch v. repeat trials was reversed for the biased tasks, while the unbiased task was still presented 5 times as switch and 5 times as repeat trials. Participants were

randomly assigned to conditions where either the letter, digit, or color task appeared as the unbiased task.

### 2.3.1.3 Design

The experiment followed a 2 (task sequence: switch vs. repeat) x 2 (task bias: biased vs. unbiased) x 3 (switch proportion: 30% vs. 70%) repeated-measures factorial design.

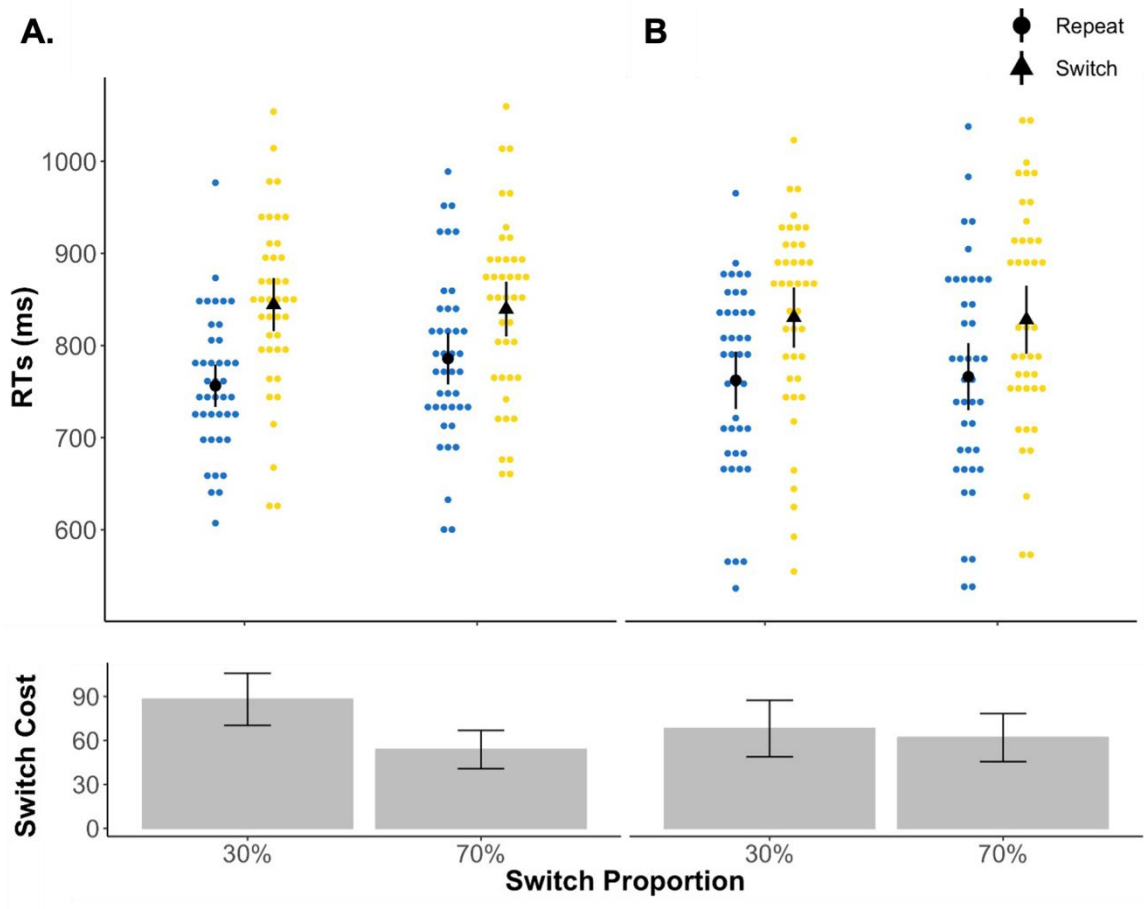
## 2.3.2 Results and Discussion

The same exclusion criteria (including RT exclusion) as Experiment 1 and 2 were applied. Descriptive statistics are displayed in Table 2.

We ran a repeated-measures ANOVA with the independent variables of task sequence (switch vs. repeat), task bias (biased vs. unbiased task), and switch proportion (30% vs. 70%). We observed a main effect of task sequence on RT,  $F(1, 40) = 127.63$ ,  $p < .0001$ ,  $\eta_p^2 = .76$ , as responses on switch trials were slower ( $M_{\text{switch}} = 829.25$  ms) compared to repeat trials ( $M_{\text{repeat}} = 758.51$  ms). As in Experiments 1 and 2, there was a significant interaction effect of task sequence x switch proportion,  $F(1,40) = 8.34$ ,  $p = .006$ ,  $\eta_p^2 = .17$ , wherein switch costs were higher in the 30% switch condition ( $M_{\text{switchcost}} = 78.06$  ms) compared to the 70% switch condition ( $M_{\text{switchcost}} = 57.82$  ms).

Most pertinently, we observed a three way interaction between task sequence, switch proportion, and task bias ( $F(1, 40) = 4.16$ ,  $p = .048$ ,  $\eta_p^2 = .09$ ), which was due to the fact that the interaction effect of task sequence x switch proportion was driven by the

biased tasks only (see Figure 5). Post hoc tests confirmed that, in the biased task condition, switch cost in the 30% switch condition ( $M_{\text{switchcost}} = 88.01 \text{ ms}$ ) was significantly higher than in the 70% switch condition ( $M_{\text{switchcost}} = 53.75 \text{ ms}$ ,  $F(1,40) = 11.53$ ,  $p = .002$ ,  $\eta_p^2 = .22$ ).



**Figure 5: Experiment 3a mean RTs and switch costs for (A) biased and (B) unbiased transfer tasks. Upper panels depict group mean RTs in black circles (repeat) and triangles (switch) and individual mean RTs in blue (repeat) and yellow (switch) dots. Lower panels depict mean switch costs in bars. All error bars indicate confidence intervals ( $1.96 \times$  standard error). In the biased task condition, switch cost in the 30% switch condition was significantly greater than the 70% condition. Switch cost differences were not significant in the unbiased transfer task.**



To ensure that the three-way interaction was not due to greater power for detecting effects of the biased tasks (there were two biased tasks, which meant there were more total biased task trials), we ran three separate ANOVAs on each of the two biased tasks and the transfer task. The task sequence  $\times$  switch proportion interaction effect was significant for both biased tasks ( $F(1,40) = 4.14, p = .05, \eta_p^2 = .09$ ;  $F(1,40) = 11.39, p = 0.002, \eta_p^2 = .22$ ), but not for the unbiased task ( $F(1,40) = .43, p = .52, \eta_p^2 = .01$ ).

As before, the accuracy data showed a main effect of task sequence, reflecting significant switch costs ( $M_{\text{repeat}} = .87; - M_{\text{switch}} = 0.80$ ),  $F(1, 40) = 49.43, p < .0001, \eta_p^2 = .55$ . There was also an interaction between task sequence and task bias (biased v. unbiased task), wherein participants exhibited lower mean switch cost in the transfer task ( $M_{\text{switchcost}} = 0.045$ ) compared to the biased tasks ( $M_{\text{switchcost}} = 0.071$ ),  $F(1, 40) = 4.54, p < .04, \eta_p^2 = .10$ . All other main effects and interactions were nonsignificant, including the crucial three-way interaction between task sequence, switch proportion, and task bias we observed in the RT data.

Experiment 3a found that the inverse relationship between switch cost and switch proportion did not hold for a transfer task that did not exhibit a switch frequency bias, even while it could be found in the biased tasks that occurred in the same block. One factor to consider when interpreting Experiment 3a's results is that manipulating task-level switch proportion biases may have introduced inadvertent determinacies in the task-sequence that could have affected switch cost. In the 30% switch proportion

blocks, in order to keep overall switch proportion low while presenting the transfer task half of the time as switch and half of the time as repeat trials, the transfer task had to be presented as a switch trial every time participants were switching *away* from one of the biased tasks. In other words, in order to present the transfer task enough times on switch trials, every switch *from* one of the other tasks had to be utilized as a switch *to* the transfer task. This pattern within the task sequence may have made it easier for participants to switch to the transfer task, compared to either of the biased tasks, in the 30% switch probability condition, thereby lowering switch cost for only the transfer task in the low switch likelihood condition, where one would have expected greater switch costs, thus potentially counteracting the detection of a possible LWPSE. We designed Experiment 3b to control for this possibility.

## **2.4 Experiment 3b Introduction**

In order to control for the potential confound posed by the inadvertent determinative task sequences in Experiment 3a, in Experiment 3b we reduced the number of transfer task trials in each block, from a third of all trials to around one fifth. Note that while the relatively rare occurrence of the transfer tasks may generally affect performance on transfer task trials, it should do so in the same manner across the two switch frequency conditions. Since we are separately comparing how the transfer and biased tasks performance changes in different switch frequency contexts, the differential number of task presentations between biased and transfer tasks should therefore not

affect our comparisons of interest. Importantly, reducing trial numbers of the transfer task allowed us to create a scenario where only a minority of the switches away from either of the biased tasks represented a switch to the transfer task. In this manner, Experiment 3b tested whether Experiment 3a's results – the lack of transfer of block-wise proportion biases to the unbiased transfer task – would still hold after removing potential confounds due to predictable transitions between tasks.

## **2.4.1 Method**

### **2.4.1.1 Participants**

120 participants were recruited from MTurk. The experiment lasted ~44 minutes and participants were compensated \$5.60 if they performed at 65% accuracy or above. 59 participants were excluded from data analysis for lower than 65% overall accuracy on the task, leaving a final sample size of 61.

### **2.4.1.2 Stimuli and Procedures**

Task stimuli, trial components, and timing were identical to Experiment 3a, except that error feedback was only presented for 500 ms instead of 2000 ms to cut down on experiment time. (In Experiment 3a, participants were shown a blank screen for 500 ms after stimulus offset if their response was correct or the word "ERROR" in red for 2000 ms if their response was incorrect.) Participants were randomly assigned to different response mappings for each task and the identity of the unbiased task (i.e. whether it was the letter, digit, or number task).

The experiment consisted of 18 blocks of 39 trials each. The number of blocks was increased compared to Experiment 3a in order to ensure that each participant was exposed to 144 trials of the rarely occurring transfer task (there were 8 transfer task trials per block). Nine blocks had a 39% switch proportion and nine blocks had a 61% switch proportion. Trial sequences were pseudo-randomly generated. In the low switch proportion block, the two biased tasks were each presented either 5 or 6 times as switch trials and 10 or 9 times as repeat trials and the unbiased task was presented 4 times as switch and 4 times as repeat trials, creating an overall switch to repeat ratio of 15:23, or 39% switch trials. Thus, in low switch proportion blocks, a switch away from either of the biased tasks was associated with a switch to the transfer task in ~33-40% of cases. In the high switch proportion block, the number of switch v. repeat trials was reversed for the biased tasks, while the unbiased task was still presented 4 times as switch and 4 times as repeat trials, creating an overall switch to repeat ratio of 23:15, or 61% switch trials. Thus, in high switch proportion blocks, a switch away from either of the biased tasks was associated with a switch to the transfer task in ~17-18% of cases. Participants were randomly assigned to conditions where either the letter, digit, or color task appeared as the unbiased task.

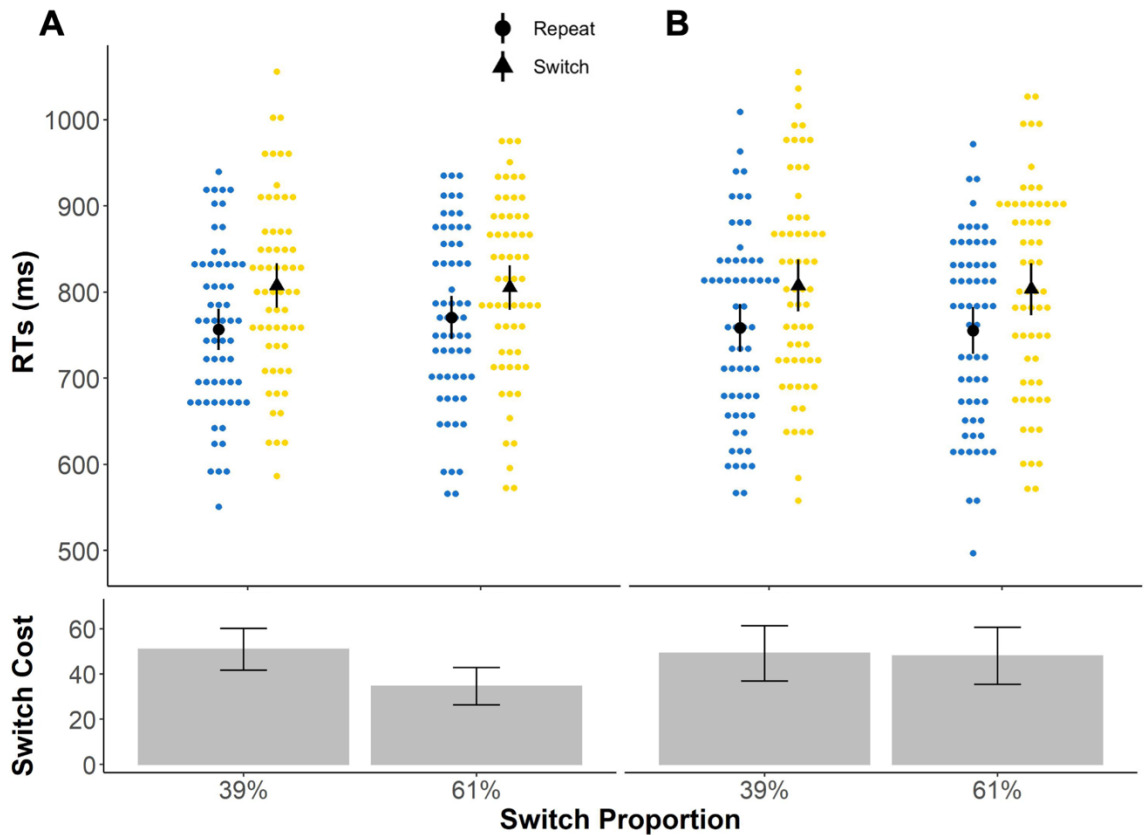
### 2.4.1.3 Design

The experiment followed a 2 (task sequence: switch vs. repeat) x 2 (task bias: biased vs. unbiased) x 3 (switch proportion: low vs. high) repeated-measures factorial design.

### 2.4.2 Results and Discussion

The same exclusion criteria (including RT exclusion) as Experiment 3a were applied. Descriptive statistics are displayed in Table 2, side-by-side with those of Experiment 3a.

We ran a repeated-measures ANOVA with the independent variables of task sequence (switch vs. repeat), task bias (biased vs. unbiased task), and switch proportion (low vs. high). We observed a main effect of task sequence on RT,  $F(1, 60) = 162.52$ ,  $p < .0001$ ,  $\eta_p^2 = .73$ , as responses on switch trials were slower ( $M_{\text{switch}} = 802.32$  ms) compared to repeat trials ( $M_{\text{repeat}} = 756.07$  ms). The interaction effect between task sequence  $\times$  switch proportion only reached marginal significance,  $F(1,60) = 3.75$ ,  $p = .06$ ,  $\eta_p^2 = .06$ , wherein switch costs were marginally higher in the low switch condition ( $M_{\text{switchcost}} = 49.96$  ms) compared to the high switch condition ( $M_{\text{switchcost}} = 41.28$  ms) (See Figure 6).



**Figure 6: Experiment 3b mean RTs and switch costs for (A) biased and (B) unbiased transfer tasks. Upper panels depict group mean RTs in black circles (repeat) and triangles (switch) and individual mean RTs in blue (repeat) and yellow (switch) dots. Lower panels depict mean switch costs in bars. All error bars indicate confidence intervals ( $1.96 \times$  standard error). In the biased task condition, switch cost in the 39% switch condition was significantly greater than the 61% condition. Switch cost differences were not significant in the unbiased transfer task.**

The previously observed three-way interaction between task sequence, switch proportion, and task bias reached marginal significance ( $F(1, 60) = 3.44, p = .07, \eta_p^2 = .05$ ). Figure 6 shows that, as in Experiment 3a, the trending interaction effect of task sequence  $\times$  switch proportion was driven by the biased tasks only. Post hoc tests confirmed that, in the biased task condition, switch cost in the low switch proportion condition ( $M_{\text{switchcost}}$

= 50.83 ms) was significantly higher than in the high switch proportion condition ( $M_{\text{switchcost}} = 34.58$  ms,  $F(1,60) = 13.63$ ,  $p < .0005$ ,  $\eta_p^2 = .19$ ).

To ensure that this marginally significant three-way interaction was not due to greater power for detecting effects of the biased tasks, which had much larger trial counts, we ran three separate ANOVAs on each of the two biased tasks and the unbiased task. As in Experiment 3a, the task sequence  $\times$  switch proportion interaction effect was significant for both biased tasks ( $F(1,60) = 6.81$ ,  $p = .01$ ,  $\eta_p^2 = .10$ ;  $F(1,60) = 8.72$ ,  $p = .004$ ,  $\eta_p^2 = .13$ ), but not for the unbiased task ( $F(1,60) = .02$ ,  $p = .88$ ,  $\eta_p^2 = .0004$ ).

The accuracy data showed a main effect of task sequence, reflecting significant switch costs ( $M_{\text{repeat}} = .86$ ;  $-M_{\text{switch}} = 0.79$ ),  $F(1, 60) = 91.32$ ,  $p < .0001$ ,  $\eta_p^2 = .60$ . The three-way interaction between task sequence, switch proportion, and task bias did not reach significance. Separate ANOVAs on the two biased tasks and the transfer task found no task sequence  $\times$  switch proportion interactions. All other main effects and interactions were non-significant.

In Experiment 3a and 3b, we modelled task bias as factor with two levels, rather than capturing the manipulation as the independent variable "task" (biased task 1, biased task 2, transfer task) with three levels, for clarity and ease of presentation. In Experiment 3b, the previously marginally significant interaction between task sequence and switch proportion ( $F(1,60) = 3.75$ ,  $p = .06$ ,  $\eta_p^2 = .06$ ) became significant ( $F(1,60) = 7.91$ ,  $p = .007$ ,  $\eta_p^2 = .12$ ) and the previously marginally significant three way interaction

between task sequence, switch proportion, and task bias ( $F(1, 60) = 3.44, p = .07, \eta_p^2 = .05$ ) became nonsignificant ( $F(1, 60) = 2.17, p = .12, \eta_p^2 = .03$ ).

The results of Experiment 3b replicated those of Experiment 3a, thus suggesting that the lack of transfer effects of the LWPSE in the former was not solely an artifact of determinacies in the task transitions. Experiment 3a and 3b jointly indicate that the LWPSE does not transfer to an unbiased task inserted into biased blocks of low or high switch proportion. This suggests that the adjustments of switch-readiness mediating the effect are tied to task-specific learning of switch proportions. We speculate on how switch proportion biases could operate on the level of task-sets in the General Discussion. Consequently, these results imply that participants do not enter a globally more flexible state in high switch proportion blocks.



**Table 2: Mean Response Times (ms) and Accuracy (%) with Standard Errors in Experiment 3a and 3b as a Function of Task Bias and Switch Proportion**

	<i>Exp 3a</i>				<i>Exp 3b</i>			
	30%		70%		30%		70%	
<i>Biased Tasks</i>	<i>RT</i>	<i>Acc</i>	<i>RT</i>	<i>Acc</i>	<i>RT</i>	<i>Acc</i>	<i>RT</i>	<i>Acc</i>
Switch	844.4 <sup>2</sup> (14.7)	77.4 (2.2)	839.6 (15.2)	80.3 (2.7)	807.4 (13.2)	80.0 (1.6)	805.0 (13.2)	79.5 (1.4)
Repeat	756.4 (11.6)	87.3 (1.6)	785.8 (14.2)	87.2 (1.6)	756.6 (12.2)	86.4 (1.3)	770.5 (12.8)	85.8 (1.5)
<i>Transfer Task</i>								
Switch	830.5 (12.2)	79.8 (2.1)	828.1 (18.9)	83.6 (1.8)	807.5 (15.3)	76.5 (2.1)	803.3 (15.4)	76.5 (2.1)
Repeat	762.4 (15.9)	85.1 (1.9)	766.2 (18.6)	87.3 (1.9)	758.4 (14.0)	85.3 (1.7)	755.2 (13.8)	83.3 (1.8)

## 2.5 Experiment 4 Introduction

Experiment 3a and 3b investigated whether the adjustments of switch readiness observed in Experiments 1 and 2 occur in a global, generalizable fashion, which would allow them to transfer to an unbiased task embedded within biased blocks. Our results suggest that the adjustments of switch readiness were instead task specific, as modulation of switch costs were observed only for tasks whose switch frequency was biased. In other words, learned switch readiness seems to be directly tied to those task sets that are

---

<sup>2</sup> Data refers to group mean RTs (excluding error trials, trials following error trials, and RT outliers) and percentage accurate with SE in parentheses.

switched to more frequently. This finding, however, still leaves open two possible drivers of the list-wide proportion switch effect, as a given task set associated with frequent switching in our protocol also entails a specific set of task stimuli that also occur more (or less) frequently in the context of task repetitions or switches in different blocks. That is, all the stimuli in high switch proportion blocks are presented more often as switch than repeat trials (and the opposite is true for low switch proportion blocks). Thus, in addition to introducing task-level biases, biased blocks also create item-specific biases. Given that it has previously been shown that switch readiness can become associated with specific stimuli (Chiu & Egner, 2017; Leboe, Wong, Crump, & Stobbe, 2008, Exp 1), it is possible that the list-wide switch proportion effect investigated here is entirely due to a cumulative item-specific switch proportion effect.

In a final experiment, we therefore examined the role of biased associations at the stimulus- vs. task- level. This was accomplished by reverting to the two-task design of Experiment 1, but now adding a manipulation that divided the stimuli used in each task set into a biased set (driving the overall switch proportion) and an unbiased set that was associated with an equal number of task repetitions and switches (cf. Bugg & Chanani, 2011). This set-up allowed us to segregate possible item-specific switch proportion effects (analyzing biased stimuli) from possible task-level switch proportion effects (analyzing unbiased stimuli).

## **2.5.1 Method**

### **2.5.1.1 Participants**

Since we used a total of 16 stimuli, participants were only exposed to each letter or number one, two, or four times (depending on the experimental conditions) in each block of 60 trials. This may be a problem for Experiment 4 as it depends on an item-level manipulation of switch proportion associations; in order to establish these item-level associations, participants must encounter each letter/digit a sufficient number of times. Given power concerns due to participants' reduced chance of exposure to each stimulus condition, we doubled the target sample size compared to Experiment 1, 2, and 3a. 120 participants were recruited from MTurk. The experiment again lasted ~45 minutes and the same payment and exclusion criteria as Experiment 3 were applied, leaving a final sample size of 82.

### **2.5.1.2 Stimuli and Procedures**

Task stimuli, trial components, and timing were identical to Experiment 1. For each participant, half of the stimuli items (i.e. 2 vowels, 2 consonants, 2 odd digits, and 2 even digits) were pseudo-randomly selected as biased stimuli and the other half of the stimuli items were selected as unbiased stimuli.

The experiment consisted of 8 blocks of 61 trials each. The number of trials per block was doubled compared to previous 3 experiments to ensure that participants encountered each of the 16 stimuli items at least once as a switch and once as a repeat

trial per block. Four blocks had a 30% switch proportion and four had a 70% switch proportion. Trial sequences were pseudo-randomly generated according to the same algorithm as in Experiment 1. Additionally, the item relevant to the task cued on a given trial was predetermined to create the biased v. unbiased stimulus manipulation. For example, if the current task was the digit task, the digit in the digit and letter stimulus pair was predetermined while the letter was randomly selected from the 8 possible ones. The task-relevant items were pseudo-randomly generated via an algorithm that ensured the following: In the 30% switch block, the 8 biased items were each presented 4 times as repeat trials and once as a switch trial (total of 32 repeat and 8 switch trials with biased stimuli items) and all the unbiased items were presented once each as a switch and repeat trial except 2 which were presented twice each (total of 10 repeat and 10 switch trials with unbiased stimuli items). This created an overall switch to repeat ratio of 18:42. In the 70% switch block, the number of switch to repeat trials was reversed for the biased stimuli, but it remained 10:10 for the unbiased stimuli.

### **2.5.1.3 Design**

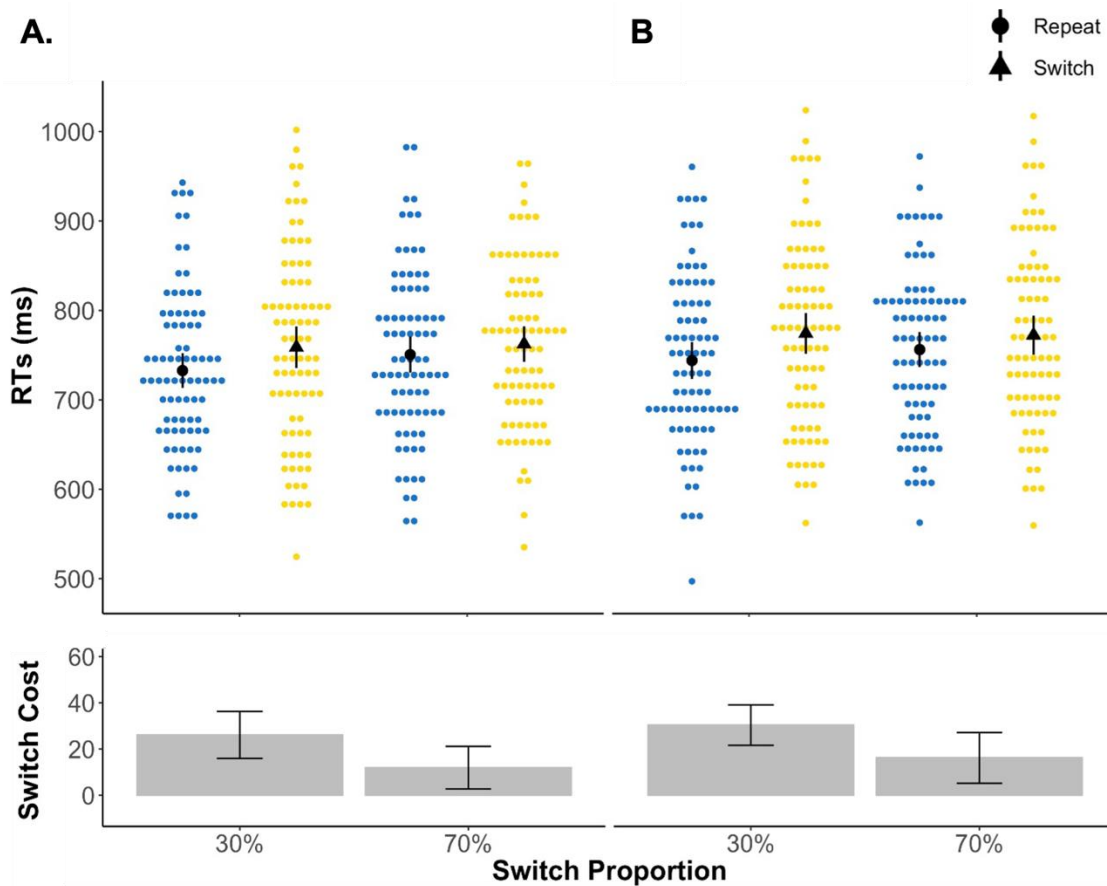
The experiment followed a 2 (task sequence: switch vs. repeat) x 2 (stimulus bias: biased vs. unbiased) x 2 (switch proportion: 30% vs. 70%) repeated-measures factorial design.

## 2.5.2 Results and Discussion

The same exclusion criteria (including RT exclusion) as in Experiment 1, 2, and 3 were applied. Descriptive statistics are displayed in Table 3.

We ran a repeated-measures ANOVA with the independent variables of task sequence (switch vs. repeat), stimulus bias (biased vs. unbiased stimulus), and switch proportion (30% vs. 70%). We observed a main effect of task sequence, as switch costs were significant ( $M_{\text{switch}} = 763.01$  ms;  $M_{\text{repeat}} = 740.07$  ms),  $F(1, 81) = 38.82$ ,  $p < .0001$ ,  $\eta_p^2 = .32$ . Moreover, RT for unbiased items ( $M = 760.51$  ms) was significantly slower than for biased items ( $M = 746.64$  ms),  $F(1, 81) = 33.77$ ,  $p < .0001$ ,  $\eta_p^2 = .29$ . There was again no main effect of switch proportion.

Importantly, we detected a task sequence  $\times$  switch proportion interaction,  $F(1,81) = 12.54$ ,  $p = .0007$ ,  $\eta_p^2 = .13$ , wherein switch cost for the 30% switch condition ( $M_{\text{switchcost}} = 28.20$  ms) was higher than for the 70% switch condition ( $M_{\text{switchcost}} = 14.04$  ms). There was no interaction between stimulus bias and switch proportion. More relevant to our assessment of differential contributions of biased vs. unbiased stimuli to the LWSPE, there was no significant three-way interaction ( $F(1,81) = 0.00$ ,  $p = >.99$ ,  $\eta_p^2 < .0001$ ) (see Figure 7).



**Figure 7: Experiment 4 mean RTs and switch costs for A) biased and B) unbiased stimuli sets. Upper panels depict group mean RTs in black circles (repeat) and triangles (switch) and individual mean RTs in blue (repeat) and yellow (switch) dots. Lower panels depict mean switch costs in bars. All error bars indicate confidence intervals ( $1.96 \times$  standard error). In both biased and unbiased stimuli sets, switch cost in the 30% switch condition was greater than in the 70% switch condition.**

To directly examine whether the switch proportion effect could be observed for unbiased stimuli, we ran separate ANOVAs for biased and unbiased items, respectively. This analysis confirmed that the task sequence  $\times$  switch proportion interaction was significant both for the biased stimuli,  $F(1,81) = 8.46$ ,  $p = .005$ ,  $\eta_p^2 = .09$ , and the

unbiased stimuli, though the effect was numerically weaker for the latter,  $F(1,81) = 4.86$ ,  $p = .03$ ,  $\eta_p^2 = .06$ .

The accuracy data showed a main effect of task sequence, reflecting significant switch costs ( $M_{\text{repeat}} = .85$ ;  $-M_{\text{switch}} = 0.80$ ),  $F(1, 81) = 113.37$ ,  $p < .0001$ ,  $\eta_p^2 = .58$ . There was also a significant effect of stimulus bias,  $F(1,81) = 18.38$ ,  $p < .0001$ ,  $\eta_p^2 = .18$ , wherein biased stimuli were associated with higher accuracy ( $M_{\text{accuracy}} = .83$ ) than unbiased stimuli ( $M_{\text{accuracy}} = .81$ ). Accuracy was also significantly higher in the 30% switch condition ( $M_{\text{accuracy}} = .823$ ) compared to the 70% switch condition ( $M_{\text{accuracy}} = .820$ ),  $F(1,81) = 4.93$ ,  $p = .03$ ,  $\eta_p^2 = .06$ . Unlike in the RT data, there was no significant interaction effect between task sequence  $\times$  switch proportion. Separate ANOVAs on the two biased tasks and the unbiased task also found no task sequence  $\times$  switch proportion interactions. All other main effects and interactions were non-significant.

In summary, though trials containing biased items produced larger switch proportion modulations of switch costs, significant switch cost adjustments were also detectable for unbiased items. Experiment 4's results therefore suggest that item-level biases are not the sole driving factor behind the LWPSE that we observed in Experiments 1, 2, and 3. The LWPSE observed in unbiased items is likely driven by task-level proportion biases. Unlike task-level bias, the presence of item-level bias is therefore not a necessary condition for making switch cost adjustments.

**Table 3: Mean Response Times (ms) and Accuracy (%) with Standard Error in Experiment 4 as a Function of Task Bias and Switch Proportion**

<i>Exp 4</i>				
	<i>30%</i>		<i>70%</i>	
<i>Biased Items</i>	<i>RT</i>	<i>Acc</i>	<i>RT</i>	<i>Acc</i>
<b>Switch</b>	758.8 (11.9) <sup>3</sup>	77.9 (1.6)	762.5 (10.2)	81.1 (1.2)
<b>Repeat</b>	732.7 (9.9)	85.1 (1.1)	750.5 (10.4)	86.5 (1.1)
<i>Unbiased Items</i>				
<b>Switch</b>	774.4 (11.6)	76.7 (1.4)	773.4 (11.1)	78.7 (1.2)
<b>Repeat</b>	744.0 (10.6)	82.7 (1.3)	756.2 (10.0)	84.5 (1.2)

## 2.6 General Discussion

This study set out to characterize the nature of the learning processes driving meta-control over cognitive flexibility, as indexed by the LWPSE. In four experiments that manipulated switch proportion, we demonstrated that the inverse relationship between switch cost and the frequency of switches in a block is driven by item- and task-level associations, rather than the global, block-level context. We find these modulations of switch costs primarily in the RT but not in accuracy data, similar to several previous studies on the LWPSE (Bonnin et al, 2011; Dreisbach & Haider, 2006; Dreisbach et al.,

---

<sup>3</sup> Data refers to group mean RTs (excluding error trials, trials following error trials, and RT outliers) and accuracy percentage with SE in parentheses.



2002). In line with previous work, Experiments 1 and 2 found reliable switch cost reductions for high proportion switch blocks when the cue-stimulus interval was short (Monsell & Mizon, 2006), and that these reductions occur even when participants cannot reliably predict which specific task they would be switching to (Chiu & Egner, 2017). Experiment 3a and 3b found that switch cost for an unbiased transfer task was unaffected by the block-wide switch-proportion context, suggesting that block-level biasing of switch likelihood does not result in a global, trans-task enhancement of cognitive flexibility. In turn, this indicates that the LWPSE is driven by task- and/or item-level learning of switch associations. Chiu & Egner (2017) found item specific switch proportion effects on biased stimuli embedded in overall unbiased (50:50 switch/repeat trials) blocks, thus demonstrating that item-level biases can elicit switch cost adjustments. Experiment 4 in the current study extends this finding by demonstrating that item-level biases are not the *only* drivers of switch cost adjustments. Rather, we also identified switch cost modulation in the unbiased (presented equally often as switch versus repeat trials) stimuli sets, indicating that task-level bias can produce adaptation in flexibility even in the absence of item-level bias.

As we alluded to in the discussion of Experiment 1, the smaller switch costs in the high switch proportion conditions across all 4 experiments were carried mainly by slower RTs to repeat trials, rather than faster RTs to switch trials (though switch trials were numerically faster in high vs. low switch proportion blocks in all four

experiments). Most previous studies reported switch costs adjustments driven by both switch and repeat trial types (e.g. Duthoo et al., 2012; Monsell & Mizon, 2006) but some have also found them driven predominantly by RT changes in repeat trials (Dreisbach & Haider, 2006). Bonnín et al. (2011) provides one possible explanation of the lack of switch trial RT adaptations: they propose that backwards inhibition (Mayr & Keele, 2000) may increase average switch trial RTs in high switch frequency blocks, since high frequency blocks contain more frequent switches back to the previous task (i.e. “ABA” task sequences) compared to low switch frequency blocks. In support of this conjecture, Bonnín et al. (2011, Experiment 1) documented numerically smaller switch RT reductions in conditions with shorter (800 ms) versus longer (1300 ms) response-stimulus intervals (RSI), the latter of which should decrease the potency of backwards inhibition effects. We employed a 810 ms RSI in the current series of experiments, which falls in the same range as the short RSI condition in Bonnín et al. (2011). Thus, even though our participants may have been exhibiting greater switch-readiness in high switch frequency blocks, the resultant performance benefits may have been masked by concurrent backward inhibition effects, an interpretation that is supported by the fact that we did not actually observe slowed switch RTs due to backwards inhibition effects in high switch frequency blocks. However, additional systematic research into backwards inhibition as another variable that can influence block-wise switch cost adjustments is clearly warranted.

### **2.6.1 Task sets form the boundaries of control strategies**

The lack of evidence for global list-wide proportion switch effects in the present study appears to clash with findings from conflict-control experiments that utilized a conceptually equivalent design of unbiased transfer items embedded within biased blocks (list-wide proportion congruency (LWPC) effects). For example, Bugg & Chanani (2011) designed a picture-word Stroop task where sets of biased items (presented more frequently as incongruent or congruent trials) were used to create mostly-congruent or mostly-incongruent blocks, and intermixed with unbiased transfer items (presented equally often as congruent and incongruent stimuli). Hence, as in our design, the biased items were biased both at the list-wide and item-specific levels, while the unbiased items were biased only at the list-wide level. Bugg & Chanani (2011) (see also Bugg, 2014; Hutchison, 2011) found that conflict (the RT difference between congruent and incongruent trials) was reduced in mostly incongruent compared to mostly congruent blocks for both stimulus sets, thus providing evidence for the existence of list-wide (global) effects of adjustments in conflict control.

Why would list-wise proportion-based adjustments in cognitive control strategy generalize in this manner for conflict control but not for task switching? We posit that this difference can be attributed to the key role that task sets play in determining processing strategies, including attentional sets (cf. Egner, 2014; Hazeltine, Lightman, Schwarb, & Schumacher, 2011; Schumacher & Hazeltine, 2016; Grant, Cookson, &

Weissman, 2020). In particular, in the case of conflict-control, the generalization of list-wide proportion effects to transfer items generally takes place within a single task set (but see Surrey, Dreisbach, & Fischer (2017; Exp 2) for an example of how context specific proportion congruence (CSPC) effects generalized to a different task set in a transfer block that maintained the same proportion congruency bias). For example, in Bugg & Chanani's (2011) picture-word Stroop task, the task is always to name the picture while ignoring the superimposed word, regardless of whether items are biased or unbiased. Since conflict-driven adaptation effects are thought to reflect the reinforcement of an ongoing task set (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Egner, 2008; Goschke, 2000), it makes sense that they would generalize to other stimuli that are processed within the same attentional set (here, "focus on pictures, ignore words"). These LWPC transfer effects from the conflict control literature are in fact directly comparable to the findings of task-level effects for unbiased items we observed in Experiment 4; here, the unbiased items for which we detected a transfer effect of the switch proportion manipulation were part of the same task set(s) within which the switch proportion bias occurred (driven by other biased items), just as in the case of the LWPC effects. Moreover, in a similar vein to our finding that task-level control effects can be observed independently of item-specific learning, Schneider (2015) found response congruency effects in a cued-task switching paradigm even when stimulus items were never repeated.

By contrast, we found no evidence for a transfer of switch cost modulation to unbiased stimuli that were associated with a distinct, unbiased task set. We are not aware of any other studies in the proportion switch domain where effects generalized in a trans-task fashion, though one study in the realm of proportion congruency (Surrey et al., 2017; Exp 2) showed transfer of CSPC effects when participants performed a new task (categorizing numbers as odd/even versus categorizing letters as consonant/vowel) in the same proportion congruence context. Our results are not directly comparable with Surrey et al. (2017) because the transfer task in our paradigm was presented with unbiased switch associations regardless of block-wide context. More future experiments would be required to delineate the exact conditions under which cross-task transfer occurs. Nonetheless, findings from the current study suggest that task sets might form the overarching cognitive structure within which frequency-driven learning of context-appropriate control strategies takes place: people are able to adapt cognitive flexibility to the specific statistics of the task(s) at hand, but they do not generalize those flexibility settings to the broader temporal (list-wide) context when having to switch to a different, unbiased task.

A potential caveat to these conclusions is that we did not assess potential transfer effects in the context of more extreme switch proportion biases than 70:30. While we (and previous studies) observed reliable switch proportion effects on the biased task sets and biased and unbiased item-levels using 70:30 (and even 61:39) trial type

proportions in the absence of any evidence for cross-task transfer, we cannot technically rule out the possibility that such effects could be observed under more extreme statistics (e.g. 90:10). However, one could also predict that a more extreme switch proportion bias would make detection of cross-task transfer less likely, as it would also make easier for participants to distinguish between the heavily biased tasks and the transfer task which remains neutral. Likewise, it is possible that cross-task transfer may develop if block numbers were increased further, and participants had more time to acquire switch bias associations.

### **2.6.2 How are task-specific switch readiness adjustments implemented?**

In assessing possible drivers of learning that may mediate the LWPSE, we have shown that strategic shifts in switch readiness are mediated by associating task sets and task stimuli with their respective switch likelihood. But how exactly are these adjustments carried out in terms of the underlying mechanisms of task-set regulation? We consider two, not mutually exclusive, possible sites of action: one related to the processes involved in switching (that is, task set reconfiguration and/or resolution of interference), and the other related task set representation (that is, the level of activation or shielding of the task rules).

In terms of switch processes, we believe that the current data are more in line with effects of switch proportion on task-set reconfiguration than on the processes of overcoming task-set inertia or resolving associative interference. First, the present study

replicated the finding that the LWSP effect is dependent on CSI in the presence of constant RSIs (Experiment 1, and Monsell & Mizon, 2006). Since the former is typically assumed to affect task-set reconfiguration processes (Rogers & Monsell, 1995) while that latter is thought to modulate effects of task-set inertia (e.g., Sohn & Anderson, 2001), this indicates that the present switch proportion effects are more likely related to modulation of reconfiguration processes than of those involved in overcoming inertia from the previous-trial task-set. Moreover, the present study employed univalent stimuli (presented side-by-side), and task stimuli associated with the uncued tasks did not systematically vary between different switch frequency conditions. The former aspect is likely to reduce associative task-set interference triggered by uncued task stimuli (Waszak, Hommel, & Allport, 2003), and even if such effects contributed to switch costs in the present study, it is not clear how they would differentially affect the different switch proportion conditions.

Monsell & Mizon (2006) proposed that the switch proportion effect observed in their study, which did not distinguish block-wide from task-set or item-related learning, may reflect either a strategy whereby participants shift from a default carryover of the prior task set to preparation of the alternative task that is implemented either before or after task cue onset (or both). The present study constrains these possibilities in important ways: first, Experiment 2 has shown that any such strategy shift is not dependent on preparation for one specific other task but can be effective in conditions of

three tasks, where no single alternate task is predictable. Second, since we did not observe transfer to an unbiased task (Experiment 3a, b), it seems unlikely that the switch frequency driven change in preparation processes occurred prior to the task cue or stimulus onset, as in that case one would anticipate transfer effects. Rather, it appears that under high switch likelihood, the shift from preparing for task-set repetition to preparing for a switch is bound to specific task cues or items, and therefore only becomes effective after task cue or stimulus onset.

If the CSI is long, the strategy (or expectation) invoked by the biased switch proportions can be overridden (Experiment 1), but at short CSIs, these effects can be observed (Experiments 1-4), such that when the actual requirements do not match the expected control settings, performance costs will be incurred (in the present case, slower task repetition performance when the task cue or item is predictive of switching). Together, these results suggest that the effects of switch proportion are mediated via fast, bottom-up (task cue- or item-triggered) priming of the context-appropriate control setting (cf. King et al., 2012). This explanation is conceptually equivalent to the account of stimulus-triggered reactive conflict-control processes mediating the item-specific proportion congruency effect (e.g., Bugg & Hutchison, 2013), but in the present case the process in question relates to task-set reconfiguration rather than conflict resolution, and can be triggered both by biased task cues and task stimuli.



In terms of effects mediated by the relative levels of activation (and/or shielding) of task-set representations, there are two *a priori* possibilities; either frequently switched-to tasks are retained in a more activated state or they become less activated (or less stable). First, it seems superficially plausible that frequently switched-to task-sets incur smaller switch costs because they are maintained at a higher level of activation (compared to conditions where they are not switched to very often). However, we believe that this is an unlikely scenario, since in that case one would predict a general RT reduction for the switch-biased tasks, regardless of whether they occur on switch or repeat trials, rather than the observed switch cost reductions that are in great part driven by repeat trial RT increases. Similarly, generally heightened activation of task-sets under high switch frequency would also predict a greater degree of task-set inertia, i.e. involuntary priming by the previous task set (Allport & Wylie, 1999), which should again result in reduced repeat trial RTs, contrary to the observed data. These factors make it unlikely that switching to tasks more frequently increases the strength of their activation.

In contrast, the possibility that task set representations become less activated (less stable or less well-shielded) under conditions of frequent switching seems more congruent with the current findings: less stable sets would be expected to be switched to and from more easily but also provide less benefits of task repetitions, which is in line with the repeat RT increases we observed in the current study under high switch

frequency conditions. The idea that greater flexibility corresponds to weaker task-set activation is also supported by recent neuroimaging evidence from Qiao, Zhang, Chen, & Egner (2017), who found via multi-voxel pattern analysis on frontoparietal cortex activity patterns that neural task-set representations are less stably encoded on task switch compared to repeat trials.

Note that the two speculative explanations that we offer, modulation of task-set reconfiguration and differential task-set activation/shielding, are not mutually exclusive – it is quite likely that both processes contribute to our observed results. In sum, our data thus suggests that instead of entering flexible versus stable global cognitive states that span the duration of a block, people seem to shift between more flexible versus more stable processing modes based on learning which specific tasks or items require greater switch-readiness or greater task-set shielding. In other words, people learn to link a context-appropriate meta-control state (e.g., an “updating threshold” setting) to task sets and stimuli and are capable of rapidly retrieving or activating that state in a bottom-up manner. These results carry practical implications for attempts to train cognitive flexibility and stability, as they demonstrate that not only external context, such as the degree of volatility in an environment, but also internal context, such as task-sets or goals, matters for learning or improving on a particular cognitive strategy.

Finally, an interesting additional question might be to what degree the learning process that we here conceptualize as reflecting the acquisition of strategic (goal-

directed) meta-control settings is an implicit or explicit one. We can only speculate on this, as we neither manipulated nor probed participants' awareness of the different switch proportion conditions. However, while the role of conscious awareness and understanding of cues for control is generally still ill-understood (see e.g., Farooqui & Manly, 2015; Bejjani, Dolgin, Zhang, & Egner, 2020), a number of related studies in the domain of conflict-control, specifically proportion congruent manipulations, have tested participants for their understanding of the experimental manipulations and have typically detected little or no explicit insight (Crump, Vaquero, & Milliken, 2008; King, Korb, & Egner, 2012; Reuss, Desender, Kiesel, and Kunde, 2014; Schmidt, Crump, Cheesman, & Besner, 2007; Bejjani, Tan, & Egner, 2020). It therefore seems likely that the same was true for the current experiments. If this assumption were correct, it would in turn indicate that meta-control, defined as adjusting the balance between complimentary control modes in a context-appropriate manner (Goschke, 2013), can in fact be accomplished by implicit learning processes. This would be consistent with perspectives that ground cognitive control in associative binding and learning processes, which do not necessarily require explicit awareness of the triggers and/or implementation of contextual adaptation in processing strategies (Egner, 2014; Abrahamse, Braem, Notebaert, & Verguts, 2016).

### 2.6.3 Possible Neural Mechanisms

How may the type of adjustments in cognitive flexibility that we document here in behavior play out in the brain? In the cognitive neuroscience literature, the effective maintenance of task sets has traditionally been associated with the dorsolateral prefrontal cortex (dlPFC) (Waskom, Kumaran, Gordon, Rissman, & Wagner, 2014; Wisniewski, Reverberi, Momennejad, Kahnt, & Haynes, 2015; Woolgar, Hampshire, Thompson, & Duncan, 2011) and, as alluded to before, recent work by Qiao et al. (2017) suggests that the stability of task set representations – as assessed by their decodability using fMRI multivoxel pattern analysis – is reduced during task switch compared to repeat trials. So one possibility is that an increase in switch-readiness under conditions of high switch-likelihood may be mediated by a strategic destabilization of dlPFC task-set representations (see our speculation on less activated task sets, above). Another, not mutually exclusive possibility is that learned changes in cognitive flexibility are mediated by the basal ganglia (BG), which are thought to play the role of a gating mechanism for updating (or not) dlPFC working memory content (Frank, Loughry, & O'Reilly, 2001; O'Reilly & Frank, 2006) including task sets (e.g. Cools, Sheridan, Jacobs, & D'Esposito, 2007). Here, greater switch-readiness may be associated with a lower threshold for gate-opening in the BG. In line with this latter possibility, it has been shown that item-specific proportion congruency effects involve prominent contributions by the caudate nucleus (Chiu, Jiang, & Egner, 2017). However, no neuroimaging study

to date has directly assessed the fMRI signatures of the LWSPE, such that a proper evaluation of the above hypotheses is left to future studies.

#### **2.6.4 Conclusions**

In summary, we presented a novel, systematic comparison of block-, task-, and item-level effects of switch proportion biases to elucidate the underlying mechanisms of the LWPSE. We observed reliable modulation of switch costs by switch proportion that was driven by participants forming associations between switch-likelihood and specific task sets and stimuli. These learned adjustments in cognitive flexibility did not generalize to an unbiased task within the same block, suggesting that task sets define a critical boundary for learning and applying suitable meta-control settings. Thus, cognitive flexibility (switch-readiness) can be strategically adapted to varying contextual demands, but the context that matters is not temporal (experimental block) but rather the task goal and task stimuli.

Since we did not find cross-task generalization of flexibility adjustments in high switch frequency blocks, the behavioral evidence from Chapter 2 suggests that there are no tonic attentional or flexibility “state” changes that underly task-switch performance benefits. Chapter 3 uses EEG, which provides a window into neural state changes, to further investigate lasting attentional adjustments (or the lack thereof) that may underly the LWPS effect.

### **3. Neural dynamics of context-sensitive adjustments in cognitive flexibility**

Theoretically, there are two rival accounts of how the behavioral LWPS effect could come about:

First, being exposed to an extended period of frequent task switches in a high-PS block could make people adopt a sustained or tonic mode of greater flexibility, which is maintained throughout the entire block. Such list-wide adjustments of control state would be consistent with behavioral evidence from conflict-control experiments, which utilize conceptually similar manipulations of the block-wise proportion of congruent versus incongruent trials in conflict tasks (e.g. Stroop task). For example, Bugg & Chanani (2011) found that conflict-control adjustments were implemented at a list-wide level and were maintained throughout the temporal context of a block, rather than varied on a trial-by-trial, stimulus-specific manner. Thus, one plausible possibility is that PS-dependent switch-readiness adjustments are implemented in a similar manner, as a change in sustained processing mode over a block of trials.

Alternatively, the switch-readiness adjustments observed in the LWPS effect could arise from phasic or transient instantiation of different levels of switch-readiness in response to specific task cues or task sets that are associated with low or high switch likelihood in low versus high-PS blocks, respectively. Chapter 2's finding that switch-readiness does not transfer to the unbiased task-set (Siqi-Liu & Egner, 2020) counter the idea of tonic state adjustments, which would predict such transfer. Rather, these data

suggest that the cognitive processes mediating the LWPS effect were instantiated only following the presentation of the relevant cues for task sets that were more often associated with switches or repetitions in each context. Studies using electroencephalographic (EEG) measures have shown that cue-evoked event-related potentials (ERPs) (Jost et al., 2008; Karayanidis et al., 2011; Lavric et al., 2008; Nicholson et al., 2006; Wong et al., 2018) and time-frequency effects in theta, alpha, and beta bands (Cooper et al., 2015, 2017, 2019; McKewen et al., 2020) differentiate between switch- and repeat- trials, constituting a “neural switch cost.” More specifically, switch cues have been associated with more positive centroparietal ERPs occurring between 200-900 ms compared to repeat cues, and effect referred to as a switch-related positivity (Barceló & Cooper, 2018; Han et al., 2018; Jost et al., 2008; Karayanidis et al., 2010, 2011; Nicholson et al., 2006; Wong et al., 2018). These neural signatures, particularly ERP correlates of switch-related activity, have been shown to be modulated by advance task-preparation during the cue-to-target interval (CTI) following exposure to an informative cue (for reviews, see De Baene & Brass, 2013; Karayanidis et al., 2010). Since ERP signatures of switch-related activity are sensitive to task preparation, it is also likely that they can be used to identify when and how strategic adaptations to context are implemented.

To the best of our knowledge, only two prior studies assessed neural signatures related to block-wise PS manipulations, one employing ERPs (Nessler, 2012), and one using functional magnetic resonance imaging (fMRI, De Baene & Brass, 2013). The fMRI

study addressed the question of which brain regions' neural switch cost activity profile may be modulated by PS context and found that a subset of regions of the frontoparietal network displayed such sensitivity (De Baene & Brass, 2013). The authors argued that the often-inconsistent findings from studies that attempt to identify brain regions underlying preparatory control may be due to the variability in switch probability across different studies. They demonstrated that only a subset of brain regions showed switch-related preparatory activity in low (30%) PS blocks were also activated in high (50%) PS blocks. However, this data provides only limited insight into how adaptations in switch-readiness are achieved in terms of how context influences control recruitment during the task-preparation process, and the time-resolution limits of fMRI meant that the study could not access whether control adjustments were accomplished in a tonic or phasic manner.

Calculating ERPs from midline electrodes, Nessler et al (2012) compared early and late switch-related parietal positivities in blocks with frequent (50% PS) versus infrequent (25% PS) switches and compared to single task blocks, utilizing either informative or uninformative cues. The uninformative cues condition was included because previous studies have demonstrated that PS had little effect on performance when participants had ample time to prepare for the upcoming target stimulus (i.e. when CTI is long [ $>790$  ms], as it was in Nessler et al, 2012). The authors focused on comparing "general switch costs," or the behavioral and neural differences between



performance on repeat trials in mixed-task versus single-task blocks, and “specific switch costs” which refer to the difference between switch and repeat trials within mixed blocks. The investigation of specific switch costs is more aligned with the goals of the current study. While Nessler et al. (2012) made important theoretical observations in regard to general switch costs, their paradigm was not optimized to elicit the LWPS effect for specific switch costs in mixed task blocks. They only found an LWPS effect when they used uninformative cues, but, even in the uninformative cue condition, no main or interaction ERP effects of PS context, or any switch-related effects during the CTI, were identified.

The current study differs from Nessler et al.’s (2012) design in several ways. First, we created experimental conditions that incentivized participants to utilize information from both the task cue and the block-wide switch proportion context. We did so by jittering target-onset relative to an informative task cue, creating a variable CTI (190 – 500 ms). Since participants did not know how much time they had to prepare for the upcoming task on the target stimulus, it would have been of value to prepare for behavioral adjustments according to block-wide switch context for optimal performance. Utilizing both an informative task cue and a CTI that should elicit behavioral LWPS effects (based on Monsell and Mizon, 2006 and Siqi-Liu and Egner, 2020), our design was set up to investigate the novel question of how PS context modulates switch-cue versus repeat-cue processing and thus allowed us to identify the timing and

characteristics of cognitive processes that contribute to meta-flexibility. Second, we employed a more data-driven approach compared to Nessler et al. (2012), conducting nonparametric analyses on data from 64 channels across the scalp rather than focusing only on midline electrodes, in order to identify potentially new neural activation patterns that underly context-sensitive flexibility adaptations. These new design features make the current study more suitable for investigating the LWPS effect.

We utilized the temporal resolution of ERPs and time frequency analyses to examine whether a tonic or a phasic account of switch-readiness adjustments more accurately describes the cascade of cognitive and neural processes that lead to the behavioral LWPS effect. We collected EEG data during a letter-digit task switching paradigm with a block-wise PS manipulation which has been shown to produce highly reliable LWPS effects (Bejjani et al., 2021; Siqi-Liu & Egner, 2020), with several minor modifications for EEG recording.

We conducted two types of analyses. First, we carried out cue-locked ERP and oscillatory EEG analyses, focused on the time between the onset of the task cue specifying the task to be performed on the upcoming task stimulus, and the onset of the task stimulus (i.e., the CTI). This allowed us to cleanly isolate activity relevant to cue-triggered preparation of the upcoming task. The tonic flexibility account would predict that PS context modulation of cue-evoked ERP amplitude differences between switch- and repeat-trials would already be evident at cue onset. By contrast, the phasic flexibility

account would predict that contextual modulation of switch-related ERPs emerge only after cue onset, since the task cue itself serves as the trigger for the cognitive processes implementing the modulation.

Second, we also investigated whether sustained control state differences between switch contexts could be identified in non-phaselocked time-frequency analyses. We analyzed oscillatory EEG power during the CTI and during the fixation period before cue-onset, primarily focusing on activity in the alpha band (8-14 Hz), given that alpha power decreases (“alpha suppression”) are commonly associated with the recruitment of attention resources to concentrate on a task (Foster et al., 2017). In a task switching context, it is possible that entering into a more stable processing mode could elicit greater alpha suppression, reflecting greater on-task concentration. Context-dependent differences in alpha power *prior* to the onset of the task cue would provide evidence for strategic adjustments in tonic (i.e., blockwise) attentional states to match environmental demands. In contrast, a lack of lasting differences in alpha power between high and low PS contexts would support the hypothesis that switch-readiness was not being implemented tonically.

### **3.1 Methods**

#### **3.1.1 Participants**

Thirty-five undergraduate students who reported no underlying medical conditions were recruited from the Duke University Department of Psychology and

Neuroscience Subject Pool. Five participants were excluded from analysis due to lower than 70% accurate task performance (1), technical or recording difficulties (3), or noisy data (1), leaving a final sample size of 30 (14 male, 16 female; mean age of 22, ranging from 18-29 years, with a standard deviation of ~3 years). All participants gave informed consent and received course credit in accordance to a protocol approved by the Duke University Institutional Review Board. All had normal or corrected-to-normal vision.

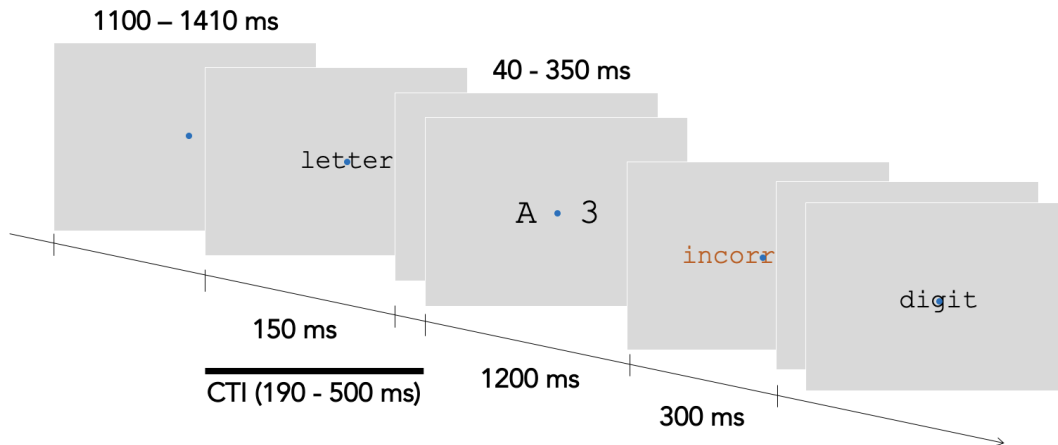
### **3.1.2 Stimuli**

Task stimuli consisted of a letter and a digit displayed simultaneously in 12 pt font on opposite sides of a fixation point at the center of the screen for each trial. The letter was randomly selected from 'A', 'E', 'I', 'U', 'G', 'K', 'M', or 'R' and the digit was randomly selected from '2', '3', '4', '5', '6', '7', '8', or '9'. Whether the letter or the digit was presented on the left or right side of fixation was randomized across trials.

### **3.1.3 Procedure**

The behavioral paradigm was programmed in Presentation from Neurobehavioral Systems (<https://www.neurobs.com/>). Each trial consisted of an inter-trial interval (ITI) of 1100-1410 ms (jittered by 10 ms bins drawn from a uniform distribution), followed by a cue display lasting 150 ms, followed by a delay of 40-350 ms (jittered by 10 ms bins matched to the ITI, such that the ITI and post-cue delay summed to 1450 ms on each trial), followed by a target stimulus presented for 1200 ms, which

was followed by a feedback screen lasting 300 ms. The CTI thus varied from 190-500 ms. The fixation dot remained at the center of the screen throughout the run (Figure 8).



**Figure 8: Experimental protocol. The trial structure involved an ITI, cue display, a delay, a target stimulus, and response feedback. In each block, 30% of trials were catch trials where the cue was not followed by a stimulus display and no response was required. CTI is marked as the sum of the cue-display interval and the jittered interval between cue-offset and target onset.**

We utilized a relatively short CTI (190-500 ms), although at the risk of significant overlap between cue- and target-evoked waveforms, because prior research (Siqi-Liu & Egner, 2020; Monsell & Mizon, 2006) did not find context-sensitive switch-cost adjustments at longer CTIs. This is likely because participants relied less on list-wide context when they have ample time to reconfigure for the upcoming task following the cue. Two measures were implemented to minimize the effect of potential cue-target overlap. First, we introduced a jitter in the delay period between the cue and target stimulus to decorrelate cue and target onsets. The length of the ITI was also jittered to

complement the delay period, thus ensuring that the response-stimulus interval, or the summed length of the ITI, cue, and delay display, was always 1600 ms. In other words, a trial with a longer delay period (and CTI) would include a correspondingly shorter ITI. The cue display period was also kept constant so that participants did not receive longer exposure to the cue on longer CTI trials. Second, to ensure that we could observe cue-evoked activity later than 190 ms (the shortest possible CTI) without overlapping target-related activity, 30% of the trials were selected to be catch trials where no target display followed the cue, and no response was required (c.f., Grent-'t-Jong & Woldorff, 2007).

Participants were required to perform a letter classification task ("Is the letter a vowel or consonant?") if they saw the cues "letter" or "abcd", and to perform a digit classification task ("Is the digit odd or even?") if they saw the cues "digit" or "number." They were informed that, on some trials, the cue will not be followed by a target stimulus (catch trials), in which case no response was required. The 2:1 cue-to-task mapping allowed us to change the cue on every trial, regardless of whether the task was switched or repeated, thus ensuring that any observed neural and behavioral differences between task switch and repeat trials would not be attributable to cue repetition effects (Logan & Bundesen, 2003; Mayr & Kliegl, 2003). Participants had to press the 'd' or 'k' key to categorize the stimuli as vowel/consonant or odd/even. Participants were randomly assigned to different response mappings for each task. The words "correct", "incorrect", and "no target" were displayed during the feedback interval on correct,

incorrect, and catch trials, respectively. Responses made while the task stimulus was no longer onscreen were considered incorrect.

Each participant completed 16 blocks of 61 trials each. The PS context of each block was 30% or 70%, and there were eight blocks of each switch proportion condition. The trial sequence for each block was generated pseudo-randomly according to an algorithm that ensured each task was presented an approximately equal number of times. In every block, 30% of the trials were randomly selected as catch trials (i.e., cue only with no target following). Accordingly, ~30% of the repeat and switch trials respectively in PS30 and PS70 blocks did not include target stimulus displays and did not require responses. Blocks of the same PS level were presented consecutively in groups of four blocks each to increase the saliency of the switch/repeat context (e.g., “AAAA BBBB AAAA BBBB”). Participants were randomly assigned to block sequences with PS30 or PS70 blocks presented first. Post-hoc analyses found no significant two- or three- way interactions between block order (PS30 first versus PS70 first), task sequence, and PS context, suggesting that it is unlikely that the order of PS context presentation affected switch costs or the LWPS effects in the current dataset. Before beginning the main experiment, participants completed one practice block with 50% switch trials and with catch trials occurring 30% of the time. Participants were instructed to maintain fixation on the central fixation dot throughout the experiment. Trials were coded as

being either task switch trials (following a different task on the previous trial) or task repeat trials (following the same task).

### **3.1.4 Behavioral Analysis**

We employed a 2 (task sequence: switch v. repeat) x 2 (PS: 30 v. 70) repeated-measures analysis of variance (rmANOVA) to assess effects on response times (RTs) and error rates. Significant interaction effects were followed up using dependent-samples t-tests. The first trial of each block and catch trials were excluded for both analyses (given that there was no target and no behavioral response on such trials), and trials with incorrect or no responses were excluded from RT analysis.

### **3.1.5 EEG Acquisition and Preprocessing**

Participants were seated approximately 60 cm from a 24-inch monitor in a dimly lit, electrically shielded, and sound-insulated room. EEG data were recorded from a 64-channel custom, extended-coverage cap (Woldorff et al., 2002) using active electrodes (actiCAP, Brain Products GmbH, Gilching, Germany). Electrodes were kept at impedances of <15 k $\Omega$  and referenced to the right mastoid during recording. The recording was obtained at a sampling rate of 500 Hz using a three-staged cascaded integrator-comb anti-aliasing filter with a corner frequency of 130 Hz (actiCHamp, Brain Vision LLC, Cary, NC, USA).

Data were pre-processed in MATLAB using the EEGLab toolbox (Delorme & Makeig, 2004). Offline EEG data were low-pass filtered at 30 Hz using EEGLAB's



Hamming-windowed sinc finite infinite response (FIR) filter (pop\_eegfiltnew), resampled to 250 Hz, high-pass filtered at .05 Hz also using an FIR filter, and subsequently re-referenced to the algebraic average of the left and right mastoids. Upon visual inspection of the data, noisy channels were interpolated using EEGLab's spherical spline interpolation (Perrin et al., 1989). A copy of the EEG data that was filtered more strongly with a high pass filter of 1 Hz was submitted to independent component analysis (ICA), and the resulting weights were then applied to the original 0.1 Hz dataset. High pass filtering of activity below 1-2 Hz to reduce low-frequency noise before the ICA calculations has been found to generally improve ICA performance and produce more clearly isolated components (Winkler et al., 2015; Klug & Gramann, 2020). Components reflecting eye movements and heartbeat were identified and removed via ICA.

In both the ERP and time frequency analyses, epochs were tagged for artifact rejection if activity between -500 to 1500 ms was greater than a participant-specific absolute threshold, which depended on variations in overall noise in the dataset. The smallest threshold for the artifact rejection was  $\pm 75 \mu\text{V}$  and the largest was  $\pm 95 \mu\text{V}$ .

### **3.1.6 ERP Analyses**

For the ERP analyses, data were epoched from -500 to 1500 ms time-locked to cue presentation and baseline corrected from -300 to 0 ms. Epochs were binned into four conditions reflecting the cells of the  $2 \times 2$  (task sequence  $\times$  switch proportion) design,

and averaged by condition to produce cue-locked ERPs. Only catch trials were included for the first round of analyses so that both early and late cue-related effects could be captured without potential contamination by target processing. Any early effects found among catch trials were then replicated in follow-up analyses that included both catch trials and non-catch trials that had targets occurring after the identified ERP differences in order to confirm the early effects with increased power. These additional late target trials were included in order to maximize trial count and increase power of any observed early effects.

Using the Fieldtrip toolbox (Oostenveld et al., 2010), we conducted dependent-samples two-tailed t-tests at each 4-ms time point between 0 ms to 700 ms (at the 250 Hz resample rate) and across all channels, corrected for multiple comparisons with nonparametric cluster-based Monte Carlo permutations (10,000 repetitions). Clusters were defined via the triangulation method, and samples with significant t values with less than 2 significant neighboring channels were discarded. To explore switch-related ERP activity, we tested the contrast between switch – repeat trials. To explore the potential interactions between task sequence and PS context, we calculated separate switch – repeat difference waves for PS30 and PS70 contexts, and compared the difference of differences (i.e.,  $PS30_{\text{switch-repeat}} - PS70_{\text{switch-repeat}}$ ). To further explore drivers of interaction effects, we conducted separate analyses of the effect of PS on switch trials (i.e.  $PS30_{\text{switch}} - PS70_{\text{switch}}$ ) and the effect of PS on repeat trials (i.e.  $PS30_{\text{repeat}} - PS70_{\text{repeat}}$ ).

Note that we did not constrain the interaction analysis based on channel locations and timepoints that were significant in the switch – repeat contrast. Rather, all timepoints from 0 ms to 700 ms across all channels were also tested for the interaction effect of task sequence and PS context. This approach allows us to uncover potential clusters where switch-related activity was moderated by PS context in the absence of a main effect of switching.

### **3.1.7 Time Frequency Analyses**

For time frequency analysis, the data were epoched from -1500 to 1500 ms time-locked to cue presentation, without baseline correction. Frequency decomposition was performed using Fieldtrip's multitaper method, where power was estimated using discrete prolate Slepian sequences in logarithmically spaced frequencies from 2 to 30 Hz. The window widths for the tapers were 2 cycles for 2-4 Hz, 3 cycles for 4-7 Hz, 5 cycles for 8-14 Hz (alpha band), 7 cycles for 15-20 Hz, and 10 cycles for 21-30 Hz. Multitaper smoothing was specified as  $5 \times \log_{10}$  of each frequency. The event-related power spectra, time-locked to the cue, for each participant were then binned and averaged according to the four conditions resulting from the  $2 \times 2$  interaction of the task sequence and PS factors. Condition-averaged cue-locked ERPs were subsequently subtracted from the power spectra in order to focus on the effects on the cue-induced oscillatory activity, or non-phase-locked power.

We conducted dependent-samples two-tailed t-tests at each 4-ms time point between 0 ms to 1000 ms and across all channels, averaging across 8-14 Hz (alpha-band power), correcting for multiple comparisons with nonparametric cluster-based Monte Carlo permutations (10,000 repetitions) implemented via the Fieldtrip toolbox. Clusters were defined using the triangulation method, and samples with significant t values with less than 2 significant neighboring channels were discarded.

We conducted these analyses timelocked to the trial-onsets as well as timelocked to the cues. For the trial-onset-locked analyses, we examined differences in alpha power between the PS30 and PS70 block-wide conditions. Trial onset-locked analyses were tailored to identify sustained modulations of attentional state (i.e., blockwise). Thus, we focused on the fixation period after trial-onset but before cue-onset to isolate potential baseline attentional differences between the block types unrelated to cue or stimulus processing. In other words, in the trial-onset-locked analyses, the fixation point functioned as a neutral cue (uninformative of the upcoming task identity) that simply signaled the start of a trial. For the cue-locked analyses, which were aimed at assessing oscillatory EEG reflections of phasic modulations of switch readiness triggered by the cues, we investigated the same contrasts as in the ERP analysis, namely switch-related effects (switch – repeat) and the interaction between task sequence and PS context ( $PS30_{\text{switch-repeat}} - PS70_{\text{switch-repeat}}$ ). All cue-locked analyses were done on catch trials

only. As with ERPs, the time-frequency interaction analyses were not constrained to timepoints and channels that were significant in the main effect of task sequence.

We also performed nonparametric cluster-based permutation analyses on theta (4-8 Hz) and beta (15-30 Hz) bands, but no significant main or interaction effects were observed, and thus the results are not reported in this manuscript.

## **3.2 Results**

### **3.2.1 Behavioral Results**

Descriptive and inferential statistics on the behavioral data are presented in Table 4 and Table 5, and summary data are displayed in Figure 9. The average RT across subjects on the task was 719 ms, and average error rate was 14%. Switch costs are calculated as switch – repeat trial RTs and error rates. The main effect of task sequence was significant for both RTs ( $p < .001$ ) and error rates ( $p < .001$ ) due to slower RTs and higher error rates on switch trials ( $M_{RT} = 736$  ms,  $M_{error} = .16$ ) compared to repeat trials ( $M_{RT} = 711$  ms,  $M_{error} = .10$ ), reflecting the classic switch cost. RTs were faster in general on PS30 trials ( $M = 717$  ms) compared to PS70 trials ( $M = 729$  ms), reflecting a main effect of PS ( $p = .024$ ). Most crucially, the interaction effect of task sequence  $\times$  PS was significant for both RTs ( $p = .002$ ) and error rates ( $p = .006$ ), due to greater switch costs ( $M_{RT} = 34$  ms,  $M_{error} = .064$ ) in the PS30 condition than in the PS70 condition ( $M_{RT} = 15$  ms,  $M_{error} = .039$ ), replicating the LWPS effect. Of note, this interaction was driven primarily by switch context affecting performance on task-repeat trials, which were reliably slower and less

accurate in the high-PS than in the low-PS context (RT:  $F(1,29) = 17.22, p < .001, \eta_p^2 = .37$ ; error rate:  $F(1,29) = 7.91, p = .009, \eta_p^2 = .21$ ), whereas switch-trial performance did not differ significantly between the two contexts (RT:  $F(1,29) = .15, p = .70$ ; error rate:  $F(1,29) = .21, p = .65$ ). In sum, these behavioral results replicate context-sensitive adjustments in switch processes as has been observed in the prior literature (e.g., Siqi-Liu & Egner, 2020), thus setting the stage for assessing the neural processes underpinning these adjustments in the EEG data.

**Table 4: Mean Response Times (ms) and Error (percentage) as a function of the proportion of switch trials (PS) and task sequence**

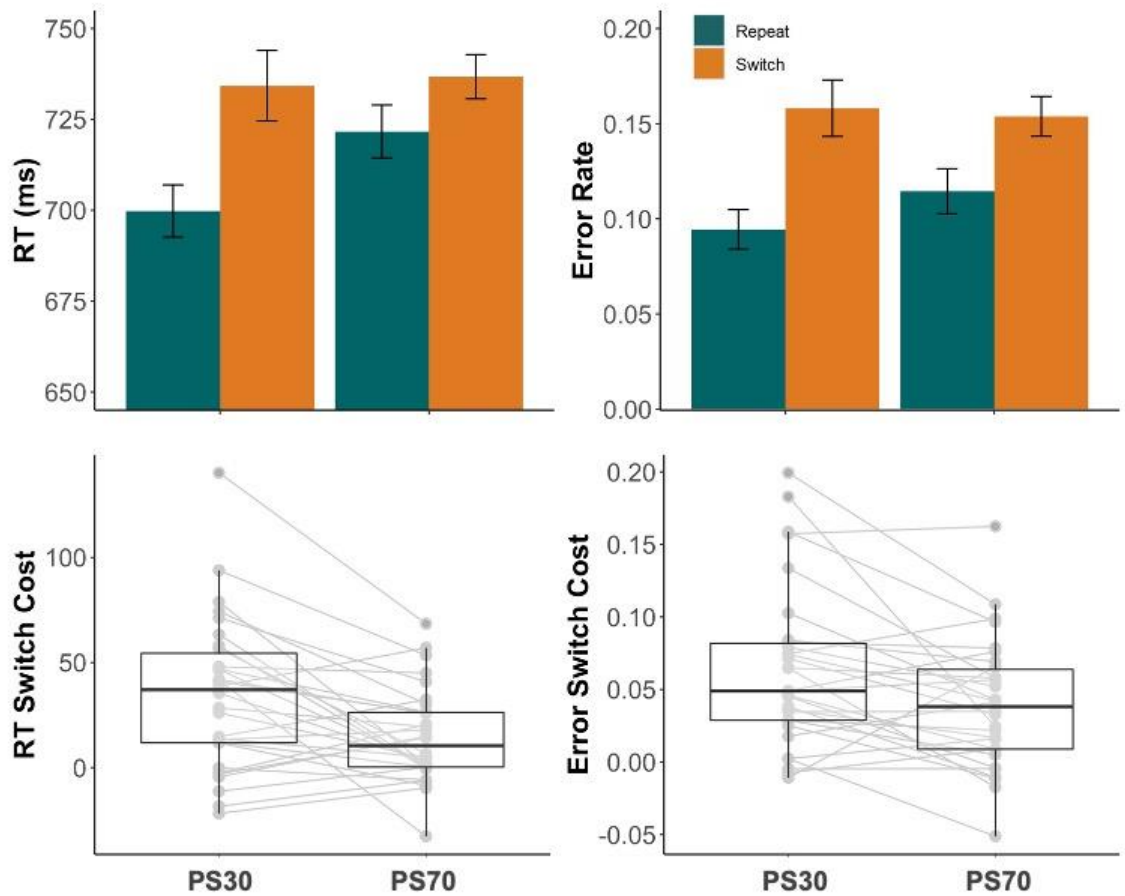
	<i>PS30</i>	<i>PS70</i>
<b>Switch</b>	734.3 / 15.8	736.7 / 15.4
<b>Repeat</b>	699.8 / 9.4	721.7 / 11.5

**Table 5: Inferential Statistics**

<b>RT Effects</b>	<b>df</b>	<b>F</b>	<b><math>\eta_p^2</math></b>	<b>p-value</b>
<b>Task Sequence (TS)</b>	1, 29	27.30***	.49	<.001
<b>Proportion Switch (PS)</b>	1, 29	5.66*	.16	.024
<b>TS x PS</b>	1, 29	12.04**	.29	.002

<b>Accuracy Effects</b>	<b>df</b>	<b>F</b>	<b><math>\eta_p^2</math></b>	<b>p-value</b>
<b>Task Sequence (TS)</b>	1, 29	39.01***	.57	<.001
<b>Proportion Switch (PS)</b>	1, 29	1.16	.39	.290
<b>TS x PS</b>	1, 29	8.86**	.23	.006



**Figure 9: Behavioral results.** Upper panels depict RT (left) and error rate (right) group means for repeat (teal) and switch (orange) trials with within-subject error bars indicating 95% confidence intervals ( $1.96 \times$  standard error). Lower panels depict switch costs for RTs (left) and error rates (right), which are calculated as switch – repeat. Box plots are overlaid above dots representing individual mean switch costs. Switch costs for both RT and error rates are higher in the PS30 condition compared to the PS70 condition.

### 3.2.2 ERP Results

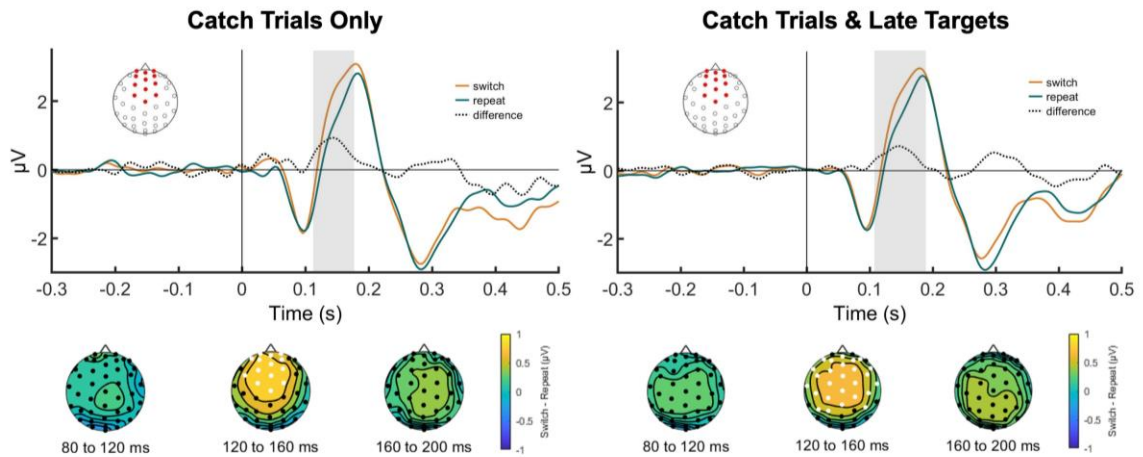
#### 3.2.2.1 Task Sequence Effects

Cluster-corrected dependent-sample t-tests on the switch versus repeat cue-evoked ERPs on the catch trials revealed one cluster with a significant positive-polarity

ERP deflection ( $p = 0.015$ ), occurring between approximately 112-172 ms, and one cluster with a significant negative-polarity deflection ( $p < .001$ ), occurring between 380-596 ms.

Visual examination of the positive cluster's topography revealed that this early switch-related effects was centered around frontal-central channels (Figure 10A), reflecting a larger positive-polarity waveform elicited on switch compared to repeat trials. To corroborate this early switch-related positivity with an analysis with higher trial counts, we conducted the same analysis after including trials with targets occurring later than 280 ms (i.e. well after significant activity was resolved). A similar positive cluster was observed in this larger trial-count analysis ( $p < .001$ ), again exhibiting frontal-central topography and with significant differences arising between approximately 104-196 ms (Figure 10B). These data document that differential neural processing on task switch compared to repeat trials is already apparent starting at ~100 ms after cue onset, and independently of physical cue changes (given that cue words changed on every trial), thus likely reflecting the rapid, cue-triggered initiation of task-set reconfiguration processes on switch trials.



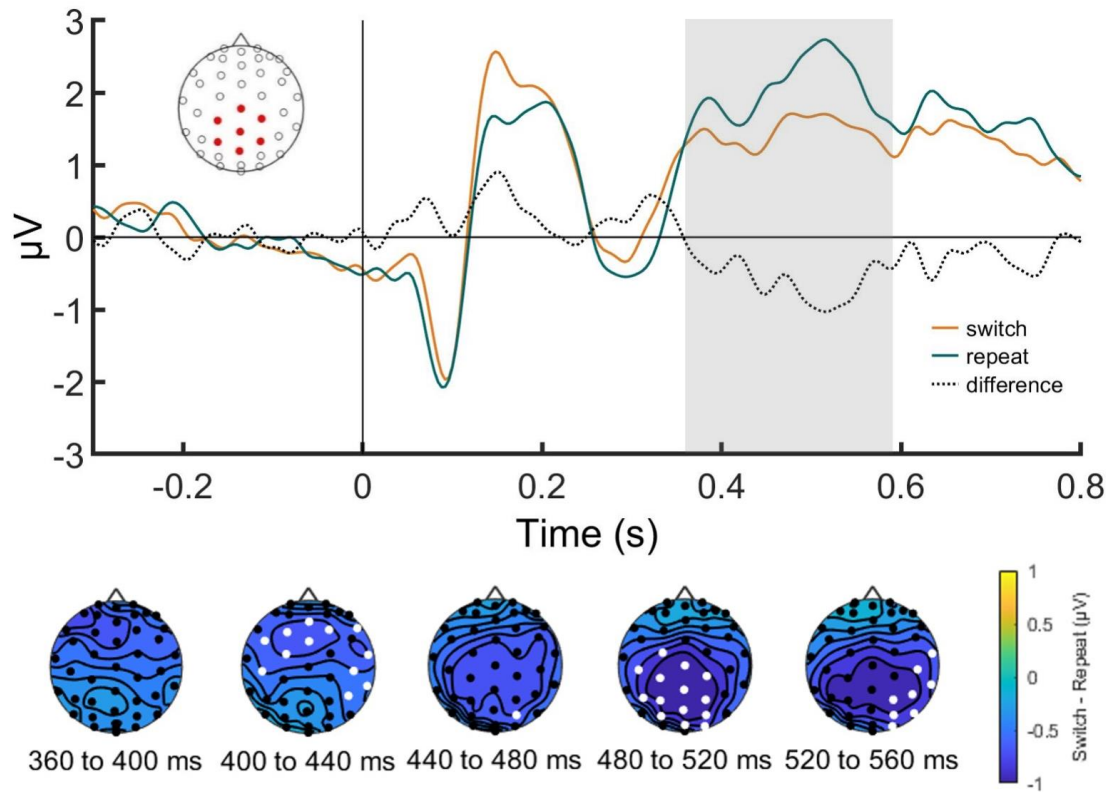


**Figure 10: Cue-locked early effects of task sequence. (Top Left) Switch, repeat, and switch minus repeat ERP activity over frontal-central electrodes (highlighted in red) for catch trials only, following cue onset. (Bottom Left) topography of a significant positive cluster ( $p = 0.015$ ) in frontal central electrodes revealed by a nonparametric cluster-based Monte Carlo permutation test of with 10,000 repetitions, unconstrained in time and electrode locations. Only electrodes that remain significant throughout the entirety of each time bin (e.g. 120 to 160 ms) are highlighted as an additional control for noise. Significant activity began  $\sim 112$  ms and lasted until  $\sim 172$  ms. (Top Right) Switch, repeat, and switch minus repeat ERP activity over frontal-central electrodes (highlighted in red) for catch trials and late target trials (post 240 ms). To maximize trial counts and increase power in this analysis, we also included non-catch trials in which the targets occurred after the significant difference was identified. (Bottom Right) The significant positive cluster ( $p < .001$ ) on this higher-trial-count analysis again exhibited similar topography and significant activity, lasting from  $\sim 104$  to  $\sim 196$  ms. Again, only channels that remained significant throughout the entirety of each time bin are highlighted in white on the topography plots.**

The longer-latency, negative cluster was characterized by lower amplitudes of ERP waveforms in switch compared to repeat trials, occurring at central-parietal electrodes (Figure 11). In line with the above interpretation for the early positive effect, this later effect might index late-stage task set reconfiguration processes in anticipation of the imminent target onset. However, a clean interpretation of this effect, observed in

catch trials, is complicated by the fact that it reaches significance around or even after the time window that the participant would expect the target to appear in non-catch trials. It may therefore be contaminated by participants terminating the process of task-set reconfiguration as they realized that no target was going to occur on that trial. No follow-up analyses that include non-catch trials with later targets were run since this negative cluster would overlap with the presentation of even later targets on those trials.

In sum, the ERP data provide evidence for both early and late differences in cue-driven processing between task switches and repetitions. We next sought to determine whether differential processing between these conditions was affected by the switch proportion context, akin to the interactions observed in the behavioral effects.



**Figure 11: Cue-locked longer-latency effects of task-sequence. Switch-minus-repeat activity in catch trials only, following cue onset. Top: ERPs over central parietal sites, corresponding to the negative cluster. Bottom: topography of negative cluster. Cluster was significant between ~380 and ~596 ms after correction for multiple comparisons ( $p < .001$ ) and generally located at central parietal channels. Only electrodes that remain significant throughout the entirety of each time bin are highlighted.**

### 3.2.2.2 Interaction Effect of Task Sequence and PS

Cluster-corrected dependent-sample t-tests on the switch-minus-repeat catch-trial difference waves in PS30 versus PS70 blocks ( $PS30_{\text{switch-repeat}} - PS70_{\text{switch-repeat}}$ ) revealed two significant negative clusters with similar parietal-occipital topography. These clusters were significant between approximately 176-280 ms ( $p = .005$ ) and 296-376 ms ( $p$

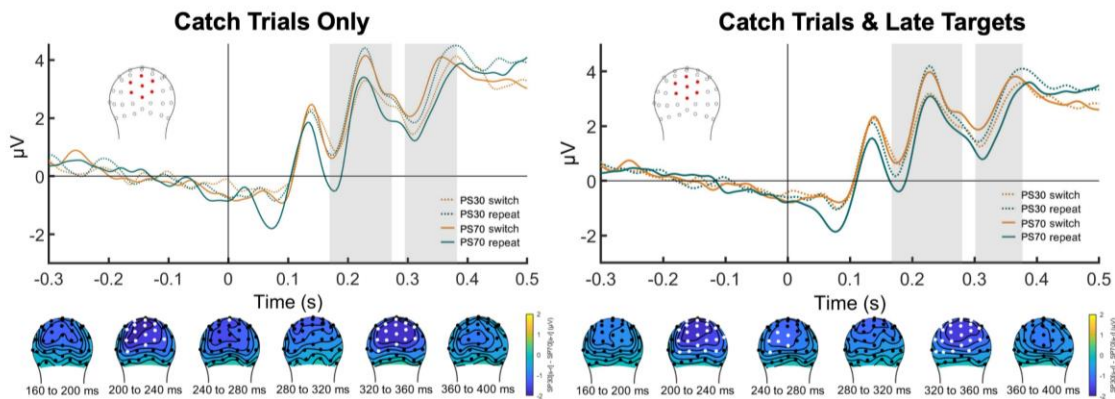
< .001), respectively. This interaction effect was driven by a switch-related negativity in the PS30 condition versus a switch-related positivity in the PS70 condition (Figure 12). In PS70 blocks, the switch-related waveform was greater in amplitude than the repeat-related waveform; on the other hand, in PS30 blocks, the switch-related waveform was smaller in amplitude than the repeat-related waveform. In other words, the more frequent trial type in each context elicited a larger positive response.

We conducted follow-up analyses to examine whether switch or repeat catch trials (or both) were driving this interaction. While the PS30 – PS70 contrast produced a significant negative cluster in repeat trials between 148-380 ms ( $p < .001$ ), the same contrast yielded no significant clusters in switch trials. This indicates that the interaction of task sequence and PS context was mostly driven by modulation of repeat-trial rather than switch-trial activity.

To corroborate these effects with higher trial counts, we conducted the same analysis after including trials with targets occurring later than 380 ms. Again, two negative clusters were found over parietal-occipital channels, displaying significant effects between 176-288 ms ( $p = .004$ ) and 300-384 ms ( $p = .007$ ), respectively (Figure 12).

These ERP results indicate that the switch-likelihood (PS context) reliably modulated cue-evoked processing starting from ~170 ms after cue onset, and that this contextual modulation affected primarily the way task repetitions were processed, thus paralleling the behavioral effects. Also note that while contextual effects only emerged

~170 ms post cue, task sequence effects were evident as early as ~100 ms post cue, suggesting that context-specific effects on processing only occurred after participants distinguished between switch task cues versus repeat task cues.



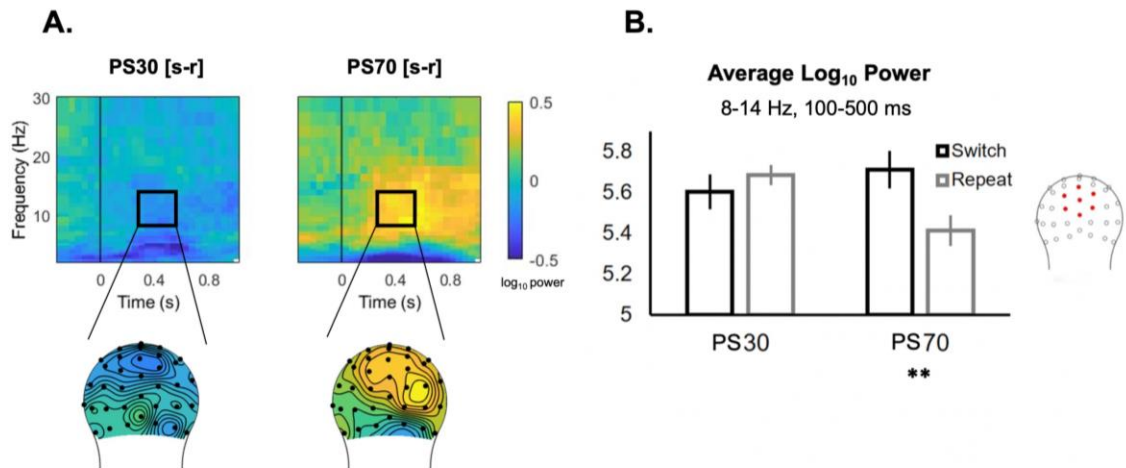
**Figure 12: Cue-locked interaction effect of block-wise PS and task sequence.**  
**(Top Left)** ERP traces of all four conditions reflecting the task sequence x PS interaction for catch trials only. Activity over parietal occipital electrodes (highlighted in red) plotted. The switch – repeat difference was positive in the PS70 condition, while it was negative in the PS30 condition. This difference is driven by block-wise frequency modulation on repeat trials (PS30 – PS70 contrast in repeat trials revealed positive cluster between 148 to 380 ms,  $p < .001$ ) but not on switch trials (no significant clusters found). **(Bottom Left)** Cluster-based analysis found two significant clusters both centered around parietal-occipital electrodes, with the first one being significant between 176 ms and 280 ms ( $p = .005$ ), and the second between 296 and 376 ms ( $p = .012$ ). **(Top Right)** ERP traces of all four conditions reflecting the task sequence x PS interaction for catch trials and the non-catch trials with targets appearing post-380 ms. **(Bottom Right)** Cluster-based analysis again found two significant clusters centered around parietal-occipital electrodes. The first one was significant between 176 ms and 288 ms ( $p = .004$ ), and the second between 300 and 384 ms ( $p = .007$ ).

### 3.2.2.3 Time Frequency Decomposition: Alpha (8-14 Hz) Power

We first analyzed alpha power time-locked to trial onset, examining activity during the pre-cue interval, to gauge sustained, across-block effects of PS context. The analyses of this trial onset-locked alpha power between the PS30 and PS70 conditions

yielded no significant effects. Thus, like the ERP results, we obtained no evidence for sustained tonic attentional state differences between the blocks where switches were frequent versus blocks where they were rare.

Next, we examined cue-locked alpha power. We observed no significant differences in alpha power between switch- and repeat- catch trials. However, the interaction effect between task sequence and PS ( $PS30_{\text{switch-repeat}} - PS70_{\text{switch-repeat}}$ ) yielded a significant cluster of decreased alpha-band power ( $p = .003$ ) over occipital-parietal channels, which was significant throughout the trial time course (Figure 13A). Follow-up specific contrasts found that this was due to significantly lower alpha power evoked in the repeat- compared to switch- trials in the PS70 condition (cluster  $p < .001$ ), but no differences between switch and repeat trials in the PS30 condition (Figure 13B). More specifically, this figure shows that there was a marked decrease in alpha for repeat trials in the PS70 blocks compared to the other three conditions, demonstrating the PS context by task-sequence interaction appeared to be driven by alpha suppression for the rare repeat trials in PS70 blocks. These alpha power differences thus mirror the ERP and behavioral data patterns, in that the PS context affected primarily activity on repeat trials rather than on switch trials.



**Figure 13: Cue-locked alpha effects.** Nonparametric analysis of the interaction effect between PS and Task Sequence revealed a positive cluster in the 8-14 Hz frequency band over parietal-occipital channels ( $p = .003$ ) which persisted throughout the trial time course. (A) Top: power spectrum of the difference between task switch and repeat trials in the PS30 condition (left) and in the PS70 condition (right). Bottom: topographic map of the power in the 8-14 Hz frequency band from 300-500 ms post-cue. (B) Average  $\log_{10}$  power between 8-14 Hz frequency, over occipital-parietal electrodes (highlighted in red) from 100-500 ms post-cue in all four conditions. The error bars represent the within-participant standard error. The interaction effect was driven by a greater difference in alpha power between switch and repeat trials in the PS70 condition. Follow up cluster-based analyses found significant positive switch related activity (\*\*cluster  $p < .001$ ) in the PS70 condition but no significant clusters in the PS30 condition.

### 3.3 Discussion

This study set out to identify ERP and oscillatory-EEG signatures underlying meta-flexibility. EEG was recorded during a cued task-switching paradigm with a block-wide manipulation of the proportion of task switches. Our behavioral results replicated previous findings of the LWPS effect (e.g., Dreisbach & Haider, 2006; Monsell & Mizon, 2006; Siqi-Liu & Egner, 2020) – of lower switch costs in high-PS compared to low-PS blocks, indicating adaptation of cognitive flexibility to situational demands. The ERP

results showed significant differences in neural processing between switch and repeat trials that emerged as early as ~100 ms post-cue, and significant modulation of switch-related processes by context by ~170 ms after cue onset. In addition, a modulation of task-cue processing by switch frequency was also identified in time-frequency decomposition of cue-locked alpha power, but not by any sustained, blockwise, oscillatory effects in place prior to cue onset. Both ERP and time-frequency results paralleled the patterns of RT effects, suggesting that they indexed neural processes that ramified to observed behavioral responses. In combination, these EEG signatures of the LWPS effect provide evidence that block-wise switch context phasically modulates task-set reconfiguration early on during the cue to stimulus interval, triggered by the task cue, rather than producing tonic cognitive state changes that are in place prior to cue onset.

In addition to the inverse relationship between switch costs and switch proportion, the behavioral results also replicate some previous observations that switch-cost adjustments were primarily driven by RT slowing on repeat trials, rather than RT acceleration on switch trials, and specifically in high-PS blocks (Siqi-Liu & Egner, 2020). This seemingly runs counter to the intuitive expectation for adjustments of cognitive flexibility to manifest as performance benefits on switch trials in frequent switch contexts, rather than performance costs on repeat trials. However, as suggested in Siqi-Liu & Egner (2020), “backward inhibition”, the finding that it is more difficult to return



to a task that was switched away from on the prior (N-2) trial (Mayr & Keele, 2003), may increase average switch trial RTs specifically in high switch frequency blocks, which contain more sequences with consecutive task alternations (e.g. 'ABABA') compared to low switch frequency blocks, which contain longer sequences of consecutive task repeats (e.g. "AAABB"). In line with this explanation, Bonnin et al. (2011) demonstrated that backward inhibition could limit switch-trial RT reductions in high switch frequency contexts, particularly in short response-stimulus interval (RSI) conditions. Though we cannot test for backward inhibition effects in the current dataset because participants only switched between two tasks (i.e. we cannot compare ABA versus ABC task sequences), the similar switch trial RTs in high and low PS contexts suggest that backward inhibition may have masked, or cancelled out existing context-sensitive switch-trial RT reductions in the high-PS context. Conversely, if no context-sensitive switch-trial RT reductions existed in the PS70 blocks, one would expect to observe slower switch-trial RTs compared to the PS30 blocks due to backward inhibition.

On the other hand, expectancy violations may have contributed to comparatively slower RTs on repeat trials in frequent switch contexts (King et al., 2012). Nonetheless, building up expectations for switch- versus repeat trials when they occur more frequently is an inherent and important part of making behavioral adaptations to different PS contexts.

ERP analyses of switching-related neural processing identified an early cue-locked switch-related positivity in frontal-central channels starting at ~100 ms following cue onset. This early difference would seem to indicate that task-set reconfiguration processes occurred quite rapidly after cue onset. Note that this observation was not confounded by cue repetitions in task repeat trials, since we included two cues for each task and used alternate cues even on repeat trials. This analysis further identified a slightly later switch-related negativity in posterior electrodes between ~400 ms to 600 ms. This switch related-negativity would seem to be inconsistent with previous findings of posterior switch-related *positivity* in the general P3b time window (Barceló & Cooper, 2018; Han et al., 2018; Jost et al., 2008; Karayanidis et al., 2010, 2011; Lavric et al., 2008; Nicholson et al., 2006; Wong et al., 2018). One potential explanation for this is our inclusion of different PS contexts, which was not manipulated in any of these prior studies.

Although there was no strict overlap between the significant timepoints of the main and interaction effects, the time window (~180 to 280 ms; ~300 ms to ~380 ms) and posterior topography of our interaction effects aligned more closely with the later switch-related negativity over posterior channels compared to the earlier activity over frontal channels in the switch-repeat contrast. The interaction effects also fell within the time window of the previously reported P3b-like switch-related positivity over centroparietal electrodes (see Karayanidis et al., 2010 for review). Importantly however,

a closer examination of the interaction effect revealed that it was driven by a switch-related negativity in PS30 blocks in contrast to a switch-related positivity in PS70 blocks. The switch-related positivity that emerged over posterior channels when task switches were frequent is thus largely consistent with previous findings of P3b-like switch-related positivity. Thus, the current results suggest that the switch-related positivity reported in previous studies may in fact be context-dependent, emerging only when switches are relatively common. This conjecture would require confirmation by future studies. Notably, these context-specific interaction effects occurred after basic switch-related processing (which occurred ~100 ms post-cue) and there was no main effect of PS context in our ERP analyses, contrary to the idea of a tonic change in flexibility between PS contexts. Notably, there was no main effect of PS context in our ERP analyses, and these context-specific interaction effects occurred after basic switch-related processing (occurring ~100 ms post-cue), thus speaking against the idea of a tonic change in flexibility between PS contexts.

Similar to the behavioral RT results, this interaction effect was driven primarily by repeat-trial amplitude differences across different PS contexts, with higher amplitudes associated with the PS70 condition. Since previous studies have shown that backward inhibition (relative to the task on the previous trial) in cued-task switching paradigms can affect ERP waveform amplitudes on the current trial (Sinai et al., 2007; Zhang et al., 2016), context-sensitive adjustments in the EEG data may also have been

harder to observe in switch than in repeat trials. In other words, repeat-trial ERPs may serve as cleaner representations of context effects because they are not affected by disproportionate backward inhibition in high switch proportion blocks. Relatedly, the lack of difference in switch-trial activity between the PS30 and PS70 conditions may also reflect the delayed or dampened initialization of task set reconfiguration processes in the PS70 condition due to backward inhibition. While a main effect of PS context would have provided evidence for “tonic” flexibility adjustments based on list-wide context, the lack of even numerical trends in this direction in our data set suggests that context-sensitive adjustments are expressed primarily after task-cue processing. The finding that contextual interaction effects only occurred after the ERPs differentiated between basic switch versus repeat trial types further supports this conclusion. The idea that the LWPS effect does not represent tonic changes of flexibility state is consistent with Siqi-Liu & Egner’s (2020) finding that behavioral adjustments occurred in a task-specific manner, rather than a block-wide one, such that switch costs were lower for tasks that were more frequently presented as switch compared to repeat trials (switch-biased), and vice versa, while no switch cost adjustments were observed in “unbiased” tasks (equally associated with switches and repeats) presented in high or low PS lists. Here, in addition, the high-temporal resolution of the ERPs supplies neural evidence for the time course and sequence of the influence of these cognitive factors on the underlying neural processes.

Time-frequency analyses of alpha power were also conducted to observe potential adjustments in attentional states. Consistent with the time course of our ERP effects, no alpha power differences between block-wide switch contexts were identified in the period preceding cue-onset, again further supporting the idea that flexibility adjustments did not occur in a block-wide, tonic manner, at least as reflected by ongoing neural activity. Power differences in other frequency bands such as theta and beta (Cooper et al., 2015, 2017, 2019; McKewen et al., 2020) have previously been shown to underly proactive control recruitment, though no significant effects at those frequency bands were identified in the current study. It is important to note however, that neither lack of pre-cue alpha effects nor the lack of effects at theta and beta frequency bands in the current study constitutes evidence for a strong conclusion that lasting differences in attentional or control states across blocks as a function of PS context do not exist.

While there was no main effect of either switch/repeat trial type or list-wide PS context on the cue-locked alpha activity, there was a significant interaction effect between trial type and context on the alpha, again centered over parietal-occipital electrodes. This effect was driven by relative alpha-suppression in task-repeat trials compared to task-switch trials in the PS70 condition, with no differences in alpha power being observed between switch and repeat trials in the PS30 condition. Again, paralleling the behavioral and ERP results, the context-sensitive adaptations were driven by changes in repeat-trial rather than switch-trial processing.

Since alpha power decrease (or suppression) is generally thought to correspond with increased recruitment of attentional resources (e.g., Foster et al., 2017), our time-frequency results could reflect the increased difficulty of repeat-trials in contexts where they are rare compared to when they are common. That these adjustments only occur in repeat trials may indicate that participants employ a default strategy of expecting and preparing for task repetitions when they occur most of the time (Duthoo et al., 2012), and changing this default repetition expectation may be effortful. That is, when the repetition expectation is violated in PS70 blocks, participants may have to abandon this default strategy, which leads to less efficient processing of task repetitions in high-PS compared to low-PS contexts. Alternatively, as discussed above, it is possible that parallel effects on switch trials are masked by backward inhibition effects.

In sum, the lack of any main effects of list-wide context in either ERP or trial-onset-locked time-frequency measures in the current study suggests that contextual modulations occur during cue-processing, wherein task cues can rapidly trigger different processing strategies depending on whether they are associated with frequent task switches or repetitions (Braem & Egner, 2018). An intriguing possibility is that block-wide differences in flexibility state may be “activity-silent” in terms of ERP and time-frequency power, analogous to the concept of activity-silent working memory representations (reviewed in Stokes, 2015) that are only revealed only in response to external inputs (Rose et al., 2016; Wolff et al., 2017). That is, there may be sustained state

differences in the brain as a function of PS context, but such state differences may not be observable in EEG and ERP activity patterns until they are triggered by the task cue. Indeed, our finding that the PS context influenced switch-vs-repeat cue processing as early as 176 ms following cue onset would seem to suggest that PS context may induce some latent preparedness for task repetitions or switches, such that the optimal neural response is rapidly triggered when the task cues occur. Future studies will be needed to corroborate this conjecture.

Chapter 2 (Siqi-Liu & Egner, 2020) proposed that such post-cue adjustments of cognitive flexibility could potentially be initiated via decreased task-set activation of frequently switched-to tasks. That is, task-sets frequently associated with switches may be maintained at lower levels of activation, which promotes greater ease of switching at the cost of reduced repetition benefits. The idea that greater flexibility is mediated by weaker task-set activation is supported by neuroimaging evidence from Qiao et al., (2017), who found that neural task-set representations were less stably encoded in frontal-parietal cortex activity patterns on switch compared to repeat trials. Reduced repetition advantage in high-PS contexts is evident in the current study's finding that behavioral, ERP, and time frequency effects were all largely driven by repeat-trial adaptations, rather than switch-trial adaptations. Furthermore, the observed alpha-suppression in response to rare repeat trials in high-PS blocks supports the idea that

repeat trials were perceived as more surprising or effortful (despite behavioral switch costs) when participants have learned to expect frequent task switches.

Importantly, the lack of sustained neural activity evidence for tonic flexibility adjustments to list-wide PS context in the current study does not necessarily contradict findings of list-wide adjustments of conflict-control in the congruency literature, which are often interpreted as evidence for proactive control in contrast to stimulus-driven, reactive control (Gonthier, Braver, & Bugg, 2016). For example, using a picture-word Stroop task, Bugg & Chanani (2011) found list-wide adjustments of conflict-control after controlling for stimulus-specific associations. Importantly, in contrast to task-switching paradigms, conflict paradigms like the Stroop task never involve task-set shifts (e.g., participants are always instructed to name the picture and ignore the word) and so may not be as suited for identifying tonic control strategies that function across different task-sets. The current study's findings suggest that list-wide context may only influence switch-related neural activity after task-cue onset, highlighting the important role that task sets play in determining information processing strategies. People may learn to link appropriate control settings to task sets and rapidly retrieve those control settings when cued in a bottom-up manner in a process that represents an integration of proactive and reactive control mechanisms (see also Siqi-Liu & Egner, 2020).



### **3.3.1 Conclusions**

The present study identifies ERP and time frequency signatures of the LWPS effect and provides evidence that neural processing is modulated by block-wide switch proportion context during the cue-target interval. Behavior, ERP, and time-frequency measures consistently indicated that modulations of repeat-trial processing drove context-sensitive flexibility adaptations. The lack of pre-cue effects or main effects of switch context suggest that “tonic” adjustments in attentional or flexibility states across the temporal context of the block are not what drives the LWPS effect. Rather, contextual modulation seems to occur during cue-processing and is potentially underpinned by weaker activation of tasks when participants must switch between tasks more frequently compared to contexts where task repeats are more common.

Behavioral and EEG effects in LWPS paradigms jointly point to the conclusion that meta-flexibility depends heavily on associative learning between task-sets and switch-frequencies. Chapter 4 investigates whether this type of control learning also behaves similarly to associative learning: i.e., can it be improved by memory consolidation?

## **4. Minimal effects of consolidation on learned switch-readiness**

There are some commonalities between this type of “control learning,” i.e., learning to associate switch-readiness with items, tasks, or contexts, and classic reward-learning, i.e., learning to associate rewards and punishment with different items or

contexts. These include the interplay between incremental and episodic learning mechanisms: while control learning has been successfully modeled using an incremental reinforcement learner that updates control parameters (e.g., Botvinick et al., 2001; Jiang et al., 2014, 2015), it could also be driven by episodic memory (e.g., Blais et al., 2007; Verguts & Notebaert, 2009), and control representations could be learned through episodic memory of a single experience (Whitehead et al., 2020). Second, control and reward learning also both involve forming associations: For example, like reward (Wimmer & Shohamy, 2012), learned control-demand associations can transfer across a small set of paired associates (Bejjani et al., 2018), suggesting generalization of cognitive strategies in linked contexts. Given the apparent parallels in learning mechanisms between control and reward learning, the underlying motivation for the present study was to probe whether a known enhancer of reward learning would also strengthen control learning; namely, consolidation, or the process by which certain memory traces stabilize over time (cf. Wimmer et al., 2018).

We were thus interested in whether consolidation could also strengthen learned associations between tasks, items, and switch-readiness that are thought to underpin the LWPS effect (Siqi-Liu & Egner, 2020). Our expectation was that consolidation would strengthen these associations. As noted above, similar to the reward and memory-based decision-making literature (e.g., Murty et al., 2016; Wimmer et al., 2018; Wimmer & Büchel, 2016), control learning has been shown to involve episodic memory traces and

long-term memory associations (Brosowsky & Crump, 2018; Spinelli et al., 2019; Whitehead et al., 2020). Thus, online reactivation of learned control associations via post-encoding consolidation could strengthen already existing control representations.

In the current, preregistered study, we implemented a cued-task switching paradigm from Siqu-Liu & Egner (2020) to examine the impact of consolidation on learned control associations. Here, based on a verbal pre-stimulus cue, participants are asked to perform either a digit or a letter categorization task. The proportion of switch trials per block is manipulated to create some blocks with a high frequency of switch trials (mostly switch or MS blocks) and other blocks with a low frequency of switch trials (mostly repeat or MR blocks). In order to isolate the effects of performing specific task sets more or less often as switch trials from effects of performing more or less switches on specific stimuli, MS and MR blocks were created using two sets of stimuli, a biased set and an unbiased (or transfer) set (Siqu-Liu & Egner (2020); see also Bugg & Chanani, 2011). The biased stimuli have a high switch frequency in MS blocks and a low switch frequency in MR blocks, whereas the unbiased stimuli are presented equally often on task repeat and task switch trials in both contexts. Thus, adaptation in switch costs for biased items can reflect item-level and task-level effects (since each task occurs more frequently on switch than repeat trials in MS vs. MR blocks), whereas adaptation in switch costs in unbiased items reflects task-level effects only.

Crucially, in the present study, we implemented this paradigm in two sessions that were separated by a ~24-hour delay period, during which episodic memory representations of context-appropriate control settings could be affected by consolidation. We implemented this longitudinal design through MTurk, where other multi-day studies have successfully yielded response rates of ~ 38% -75% (Boynton & Richman, 2014; Daly & Nataraajan, 2015). We investigated the effect of consolidation on control learning by comparing context-sensitive switch cost adjustments (the LWPS effect) between session 1 and 2. Further, the biased and unbiased stimuli set manipulation allowed us to tease apart how memory consolidation may differentially affect switch-readiness learning depending on whether it is tied to specific items (i.e., switch-biased stimuli) or not (unbiased stimuli).

The nature of this design also allowed us to examine whether consolidation affects different levels of learning that may occur between session 1 and session 2. Participants could acquire general knowledge about the statistical structure of the paradigm (“meta-learning”, i.e., learning to learn), including, but not limited to, how only a subset of stimuli are predictive of block-wide switch frequency and how predictive contexts within sessions switch between MS and MR blocks. Accordingly, even if the identity of the biased/unbiased stimuli and the order of MS/MR blocks were to change across sessions, participants may nevertheless be able to benefit from their previous exposure to the paradigm and re-learn these associations more rapidly in

session 2. Learning could also occur on the level of item-specific statistics, where knowledge that specific stimuli, such as the letter “A” or the digit “1”, are predictive of the list-wide (block-wise) switch frequency would result in more effective switch-readiness adaptations in session 2. We examined these different levels of learning across Experiments 1 and 2. Experiment 1 changed the assignment of stimuli to the biased versus unbiased conditions across sessions to examine potential effects of meta-learning of general task structure. On the other hand, stimuli assignments remained constant across sessions in Experiment 2 so that participants could rely on stimulus-specific associations.

In sum, we aimed to investigate 1) how consolidation affects learning of switch demands as expressed in the LWPS effect and, if it does, 2) whether such effects can occur at the meta level or only with respect to learning that specific stimuli are predictive of switch frequency (in which case they would only occur in Experiment 2). We expected to replicate previous findings of reduced switch costs for stimuli that were associated with more frequent switches as opposed to more frequent repeats in both sessions in both Experiments 1 and 2 (Siqi-Liu and Egner, 2020). If long-term memory consolidation benefits learning of switch-readiness, these switch cost adjustments would be larger in session 2, when participants revisited the task after a ~24-hour delay. Conversely, if long-term memory consolidation did not affect switch-readiness learning, we should find the LWPS effect of similar magnitude in sessions 1 and 2. Note that we

did not entertain the possibility of the LWPS effect decreasing from session 1 to session 2, as participants would be expected to show a within-session LWPS effect in session 2 even if they had completely forgotten about their prior exposure to the task on the previous day.

## **4.1 Experiment 1**

### **4.1.1 Method**

#### **4.1.1.1 Sample Size**

Experiment 4 within Siqi-Liu & Egner (2020) deployed a similar experimental design as the current study and recruited 82 participants. In that experiment, the overall trial type by proportion switch interaction had a  $\eta_p^2$  of 0.13. Because we maintained the same trial count within each session of the current experiments, this effect size provided an appropriate estimate for a power analysis. With a power of 0.80 and an alpha of 0.05, we would have had to recruit 56 participants for the overall two-way interaction. We therefore aimed to recruit ~82 participants with data for both sessions, matching the sample size within Siqi-Liu and Egner (2020) and exceeding the sample size recommendations for greater power in subsequent analyses.

#### **4.1.1.2 Participants**

127 Amazon Mechanical Turk (MTurk) workers consented to participate in the first session for \$0.13/minute and were invited, via email, to participate in the second session only if they achieved greater than 65% accuracy in the first session. Participant

data across two sessions were matched according to participants' unique MTurk IDs. 23 workers were excluded for poor accuracy on the first session, with 3 workers who were mistakenly invited to participate in the second session. Of the 107 workers invited for the second session, 95 participated and earned the \$1 bonus for completion. We excluded an additional 5 workers for low accuracy in the second session, 2 workers for old age (over 60-years-old), and 1 worker for having been presented with the exact same biased/unbiased stimuli across sessions by chance. This resulted in a final sample size of 102 MTurk workers for session 1 data (mean age =  $35.04 \pm 8.05$ ) and 84 for sessions 1 and 2 data (mean age =  $35.27 \pm 8.06$ ). The attrition rate between sessions 1 and 2 was 11.5% (12/104). Note that old age was not included as an exclusion factor in the preregistration; however, we excluded older participants because memory performance declines with age, and we did not want to handicap our ability to detect benefits of consolidation.

#### **4.1.1.3 Stimuli**

Task stimuli included the digits 2-9 and the letters A, E, I, U, G, K, M, and R. On each trial, one digit and one letter were presented side-by-side, one to the left and one to the right of the center of the screen (Figure 14).

#### **4.1.1.4 Procedure**

The experimental procedure consisted of a two-session cued task-switching paradigm that manipulated list-wide and item-level task-switch likelihood within participants. As shown in **Figure 14**, on each trial, participants were cued to perform

either a letter classification task, indicating whether a given letter stimulus was a vowel or consonant, or a digit classification task, indicating whether a given digit was odd or even. A 2:1 cue-to-task mapping was employed to avoid any exact cue repetitions over successive trials (Mayr & Kliegl, 2003). The cues Letter and Alphabet were used for the letter classification task, and the cues Digit and Number were used for the digit classification task, and the cue word always changed from one trial to the next. Responses were given via the d and k keys on a QWERTY keyboard, and response mappings were counterbalanced across sessions and task rules. For example, for the digit task, some participants had to press 'd' if the number was odd and 'k' if the number was even while other participants pressed 'k' for odd and 'd' for even. These assignments also changed within-participants across sessions. Each trial began with a blank screen for 1010 milliseconds, followed by a fixation cross for 450 milliseconds, a task cue of 150 milliseconds, another blank interval for 40 milliseconds, and then the task stimulus (one letter and one digit) for 1200 milliseconds. Correct responses were succeeded by a blank screen, while incorrect responses were succeeded by the word "ERROR", both displayed for 500 milliseconds. To become familiar with the task demands, participants first performed a 61-trial practice block with an equal likelihood of task-repeat and switch trials and no predictive relationship between any stimuli and switch-likelihood.



The main experiment involved three primary manipulations. First, the study involved a list-wide (or block-wise) switch likelihood manipulation: four blocks were comprised mostly of task-switch trials (70% proportion switch), while the other four were comprised mostly of task-repeat trials (30% proportion switch), creating contexts of different demands with respect to cognitive stability vs. flexibility. These eight blocks consisted of 61 trials each, ensuring participants encountered every stimulus item as both a repeat and switch trial at least once within each block. Moreover, each task was presented an equal number of times across blocks via a pseudo-random trial sequence, whereby each categorization task was presented 9 times and 21 times as switch and repeat trials respectively within the 30% switch block, and vice versa for the 70% switch block. All four blocks of each switch probability were presented consecutively, but whether participants were presented first with the 70% or 30% switch probability (i.e., “list-wide PS order”) was randomized across participants.

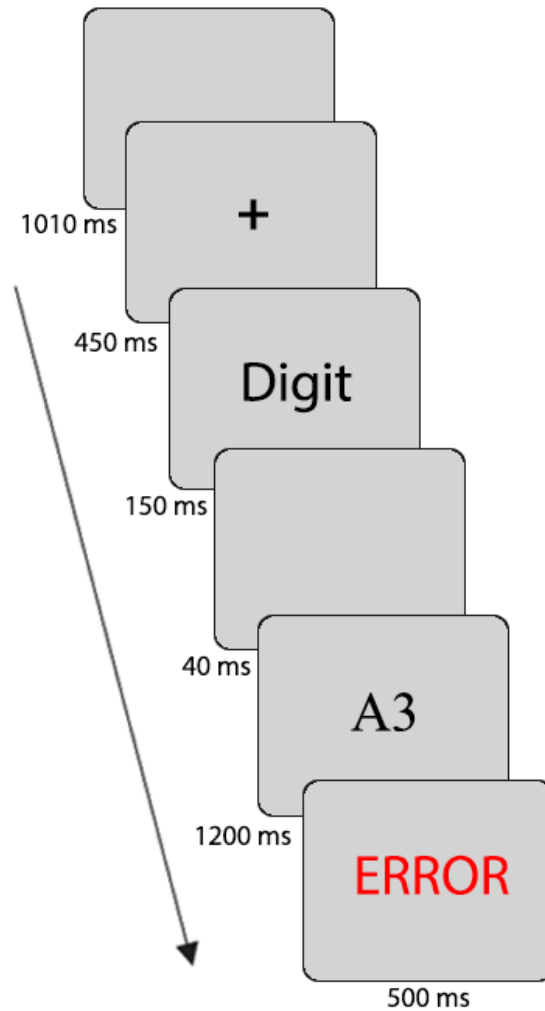
Second, we divided the stimuli used in each task into a biased set (4 digits and 4 letters), which drove the overall list-wide switch proportion by being predictive of task-switches (when presented in the PS-70 blocks) and task-repetitions (when presented in the PS-30 blocks), and an unbiased set (4 digits and 4 letters), which was associated with an equal number of task repetitions and switches in every block. Our pseudorandom stimulus sequence ensured that, within 30% switch blocks, the eight biased items were presented four times as repeat trials and once as switch trials, while the eight unbiased

items were presented once each as repeat and switch trials, except for two stimuli that were presented twice as each trial type. Thus, while the overall switch likelihood was 30% (i.e., 18:42 switch:repeat trials), the biased stimuli were associated with switch trials 20% of the time (8:32) and the unbiased stimuli was equally associated with switch and repeat trials (10:10). The corresponding manipulation was applied for the 70% switch blocks. The irrelevant feature in each trial was randomly selected from all 8 possible stimuli in that set (e.g., the letter stimulus was randomly selected from the 8 possible letters if the current trial involves the digit task). The fact that each trial contains both a letter and a digit, and that no specific letters or digits occur more frequently in one task than the other, also prevents any predictive associations to be formed between particular stimuli and task cues.

Note that in this paradigm, none of the specific stimuli, including the biased stimuli, were reliably associated with a high or low switch-readiness throughout the experiment; rather, the biased stimuli were predictive of the proportion of switches in the current block: in PS-70 blocks, the biased items occurred more often as switch trials. In the PS-30 blocks, the same biased items occurred more often as repeat trials instead. Participants would thus have to relearn any “item-level” association for the biased items when switching between PS contexts. The unbiased stimuli, on the other hand, are always equally associated with switch and repeat trials, regardless of the list-wide PS. This means that any LWPS effect observed in unbiased items is not due to item-specific

associations. Rather, LWPS in unbiased items reflect learning that certain contexts (here task sets), rather than specific items, are associated with higher/lower switch frequencies.

Third, the study involved a multi-session learning manipulation: participants who performed in accordance with predetermined accuracy requirements in session one were invited to complete the task again at a second session approximately 24 hours following the first session, thus allowing us to test for effects of memory consolidation on task performance. Participants had three-hour windows within which to complete each session (time allotted per assignment), and their ability to accept the MTurk task (expiration time) lasted six hours for both sessions. The experimental task in session 2 was exactly the same as in session 1, except that – importantly – the assignment of stimuli as biased (i.e., predictive of task repetitions or switches) or unbiased (non-predictive) was altered between sessions. This means that any effects of consolidation on task performance between sessions 1 and 2 in Experiment 1 would have to be mediated by meta-learning, that is, by knowledge about the task having a predictive structure rather than by learning of the specific stimulus-switch associations.



**Figure 14: Experiment design.** Participants performed a cued task-switching paradigm consisting of a letter classification task (vowels/consonants) following the cues Letter/Alphabet and a digit classification task (odd/even) following the cues Digit/Number. Participants received feedback when they responded incorrectly and were shown a blank screen when they responded correctly. Block-wise proportion switch (PS: mostly task-switch (MS)/mostly task-repeat (MR)) and trial type (task-switch/task-repeat) were manipulated and analyzed in addition to the session (1/2) and randomized list-wide PS order (MS first or MR first) factors.

#### 4.1.1.5 Data Analysis

All analyses only included participants who finished both sessions 1 and 2. We excluded all practice trials and the first trials of each block. We analyzed accuracy rates for all subsequent trials and reaction time (RT) data for correct trials that did not follow incorrect trials and were within 1.5 times the interquartile RT range of the remaining sample. To establish whether the proportion switch manipulation successfully induced adaptation in switch-readiness, we ran a  $2 \times 2 \times 2$  repeated-measures ANOVA for biased (PS-80/20) and unbiased (PS-50) items, respectively, with list-wide proportion switch (PS: mostly switch (MS)/mostly repeat (MR)), trial type (task-switch/task-repeat), and session (1/2) as within-participants factors. Participants who had fewer than ten trials per cell due to filtering of RT were eliminated from that analysis.

In addition to the above frequentist statistics, we report inclusion Bayes Factors, calculated using the R packages `bayestestR` and `BayesFactor`, for the repeated-measures ANOVAs. Inclusion Bayes Factors (BF) answer the question, “Are the observed data more probable under models with a specific predictor than they are under models without that specific predictor?” For instance, an inclusion BF of 2 indicates that a model with that predictor is two times more likely than a model without that predictor. We calculate the inclusion BFs for matched models<sup>4</sup>: model comparison is thus restricted to

---

<sup>4</sup> See [https://easystats.github.io/bayestestR/reference/bayesfactor\\_inclusion.html](https://easystats.github.io/bayestestR/reference/bayesfactor_inclusion.html) or Mathot, S. (2017). Bayes like a Baws: Interpreting Bayesian Repeated Measures in JASP [Blog post]. Retrieved from <https://www.cogsci.nl/blog/interpreting-bayesian-repeated-measures-in-jasp> for more details.

models without interactions that include the term of interest and for interactions specifically, averaging is restricted to models that include the main effect terms that comprise the interaction term. We chose matched models inclusion BFs to accurately account for the evidence of the interaction terms.

All materials are available online at <https://osf.io/2ueca/>. We preregistered the basic study idea at <https://osf.io/zunq6> (i.e., how does consolidation impact learned switch-readiness?) and subsequently manipulated whether biased and unbiased stimuli changed (Experiment 1) or remained constant (Experiment 2) across sessions. This allowed us to test whether consolidation would result in meta-learning benefits (Experiment 1) or depended on the particular long-term stimulus associations formed (Experiment 2).

#### **4.1.2 Results**

The analyses of interest concerned RT and accuracy rates for the biased and unbiased items across sessions. A trial type x PS interaction, whereby switch costs are reduced for high compared to low PS blocks (the LWPS effect), would indicate the presence of task- and/or item-specific learned switch-readiness for the biased items, and only task-level learning of switch-readiness for the unbiased items. The session factor provided a test of the alternative consolidation hypotheses outlined in the Introduction, that is, whether the LWPS effect would differ in magnitude for session 2 compared to session 1. Notably, we did not make strong predictions about whether these hypotheses

were specific to biased vs. unbiased stimuli, but if we observed consolidation benefits only for biased stimuli, this would suggest a greater impact of long-term memory on item-specific level associations.

#### 4.1.2.1 Biased, PS-20/PS-80 Items

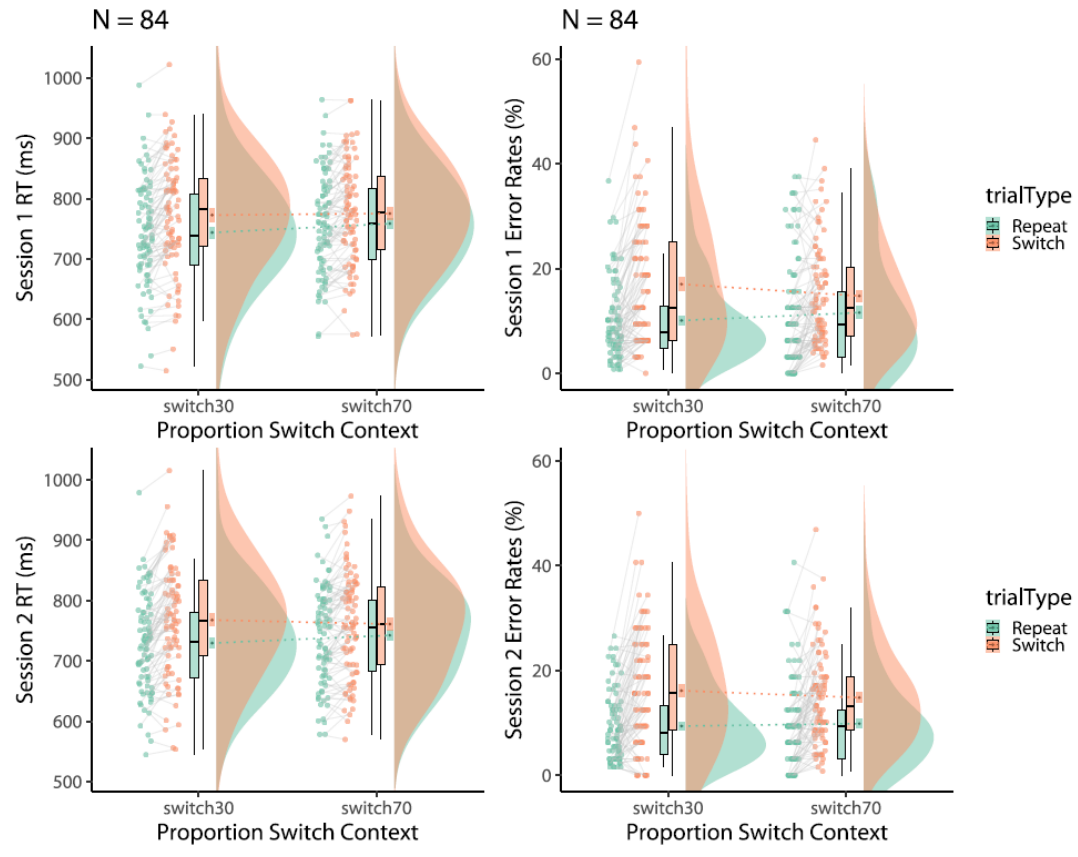
##### 4.1.2.1.1 Reaction Time

Participants showed strong switch costs ( $M = 28$  ms) (trial type:  $F(1,83) = 110.48$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.57$ , inclusion BF =  $5.65 \times 10^{10}$ ) that were modulated by proportion switch context ( $F(1,83) = 25.64$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.24$ , inclusion BF = 1.84). Switch costs were smaller in the PS-70 ( $M = 18$  ms) than PS-30 ( $M = 33$  ms) contexts. Notably, although participants responded faster in session 2 than 1 (session:  $F(1,83) = 4.71$ ,  $p = 0.033$ ,  $\eta_p^2 = 0.05$ , inclusion BF = 58.08), the size of the LWPS effect, i.e., modulation of switch costs across PS contexts, did not vary by session (Figure 15; session x proportion switch x trial type:  $F(1,83) = 1.63$ ,  $p = 0.205$ ,  $\eta_p^2 = 0.02$ , inclusion BF = 0.26; all other effects,  $F < 2.19$ ).

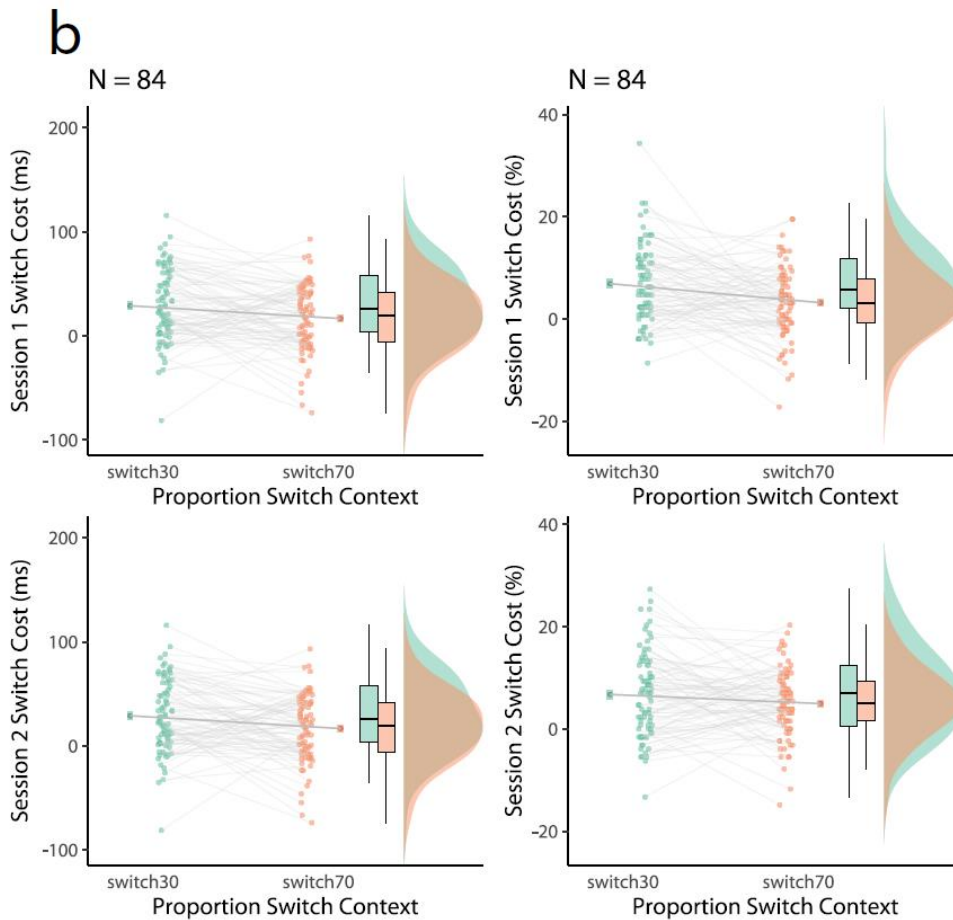
##### 4.1.2.1.2 Error Rates

We observed a strong main effect of trial type ( $F(1,83) = 136.89$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.62$ , inclusion BF =  $4.44 \times 10^{19}$ ), and these switch costs ( $M = 5.8\%$ ) were modulated by proportion switch context ( $F(1,83) = 11.29$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.12$ , inclusion BF = 3.16). Participants showed reduced switch costs during the PS-70 ( $M = 4.1\%$ ) compared to PS-30 ( $M = 6.8\%$ ) contexts. This effect did not vary by session (Figure 15; session x proportion switch x trial type:  $F(1,83) = 2.43$ ,  $p = 0.123$ ,  $\eta_p^2 = 0.03$ , inclusion BF = 0.24; all other effects,  $F < 1.27$ ).

a







**Figure 15: Behavioral results from Experiment 1 for biased (PS-20/80) items. A:** RT (ms, left panel) and error rates (% , right panel) are shown as a function of proportion switch context (PS-30/PS-70), trial type (task-repeat/task-switch) and session (1/2). On the left of each plot, light grey lines connect mean data points for each participant across trial types (task-repeat (left) connected to task-switch (right); positive slopes indicate switch costs). The boxplots (left: task-repeat, right: task-switch) show a box for the first through third quartiles of the data, a thick black line in the middle of the box to represents the median (of participant mean data), and thin black lines extending from the boxes to indicate the range of non-outlier data. Next to the box plots are the aggregate means ( $\pm$ SEM) for the trial type conditions (task-switch shown in orange, task-repeat in green), which are connected by a dotted line across PS contexts. Finally, next to the means are violin plots that represent the respective distributions of the trial type data. Session 1 data are shown on the first row, while session 2 data are shown on the second row. **B:** Switch costs for RT (ms, left panel) and error rates (% , right panel) are displayed as a function of proportion switch context (PS-30/PS-70) and session (1: top/2: bottom). On the left of each plot, mean

switch costs for individual participants are connected across proportion switch contexts (PS-30 (left) connected to PS-70 (right); negative slopes indicate an LWPS effect). We again show boxplots and violin plots to represent the respective distributions, with the PS-30 context shown in green on the left and top and the PS-70 context shown in orange on the right and bottom.

#### 4.1.2.2 Unbiased, PS-50 Items

Although we did not observe a greater LWPS effect from session 1 to 2 for biased items, it is possible that consolidation benefits associations with only unbiased stimuli, which have task-level, but not item-specific, associations. We thus again examined whether session moderated the LWPS.

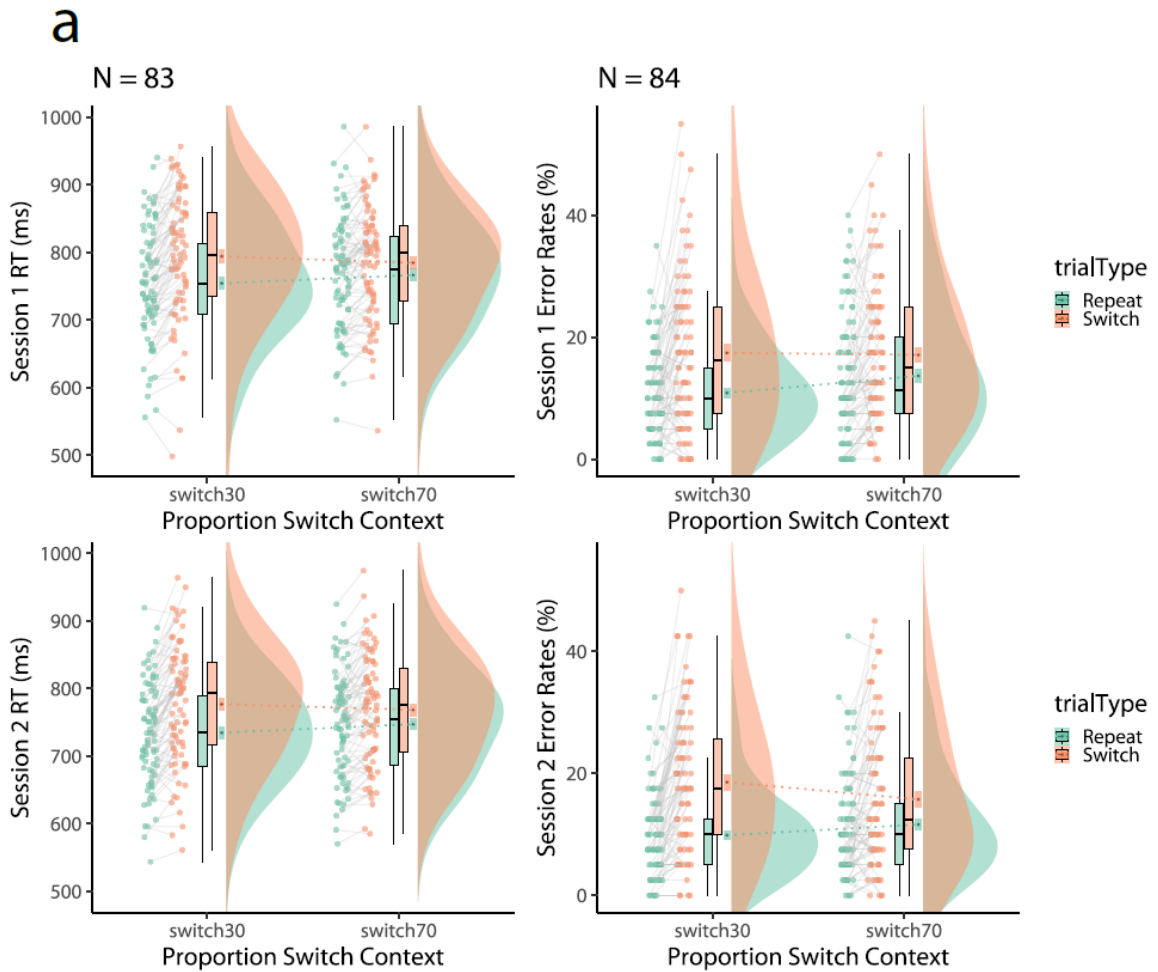
##### 4.1.2.2.1 Reaction Time

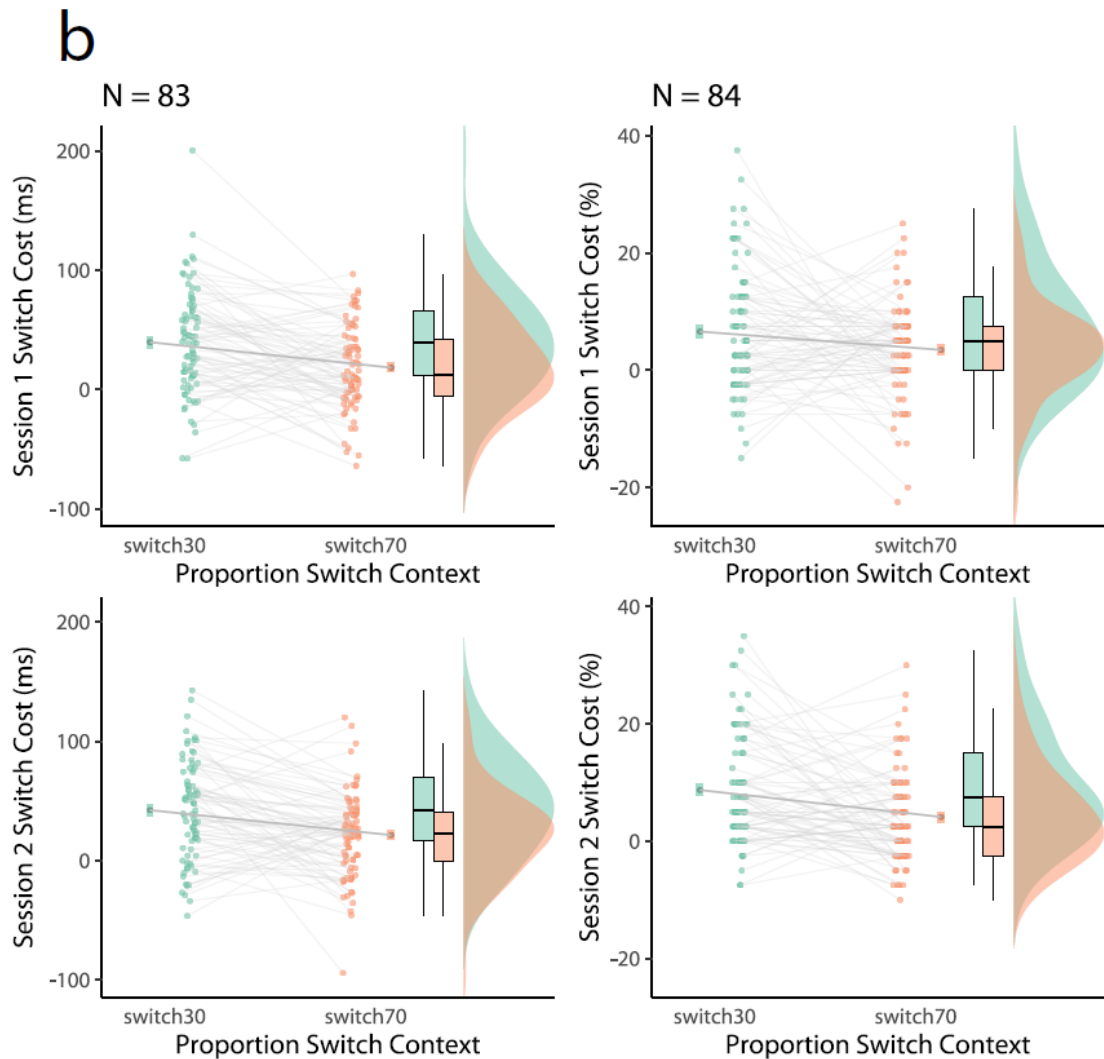
Participants showed strong switch costs ( $M = 30$  ms) ( $F(1,82) = 106.05, p < 0.001, \eta_p^2 = 0.56$ , inclusion BF =  $5.72 \times 10^{14}$ ) that were modulated by the PS context ( $F(1,82) = 37.08, p < 0.001, \eta_p^2 = 0.31$ , inclusion BF = 13.31). As expected, switch costs were smaller in the PS-70 ( $M = 20$  ms) than PS-30 ( $M = 40$  ms) contexts. Although participants grew faster in their responses from session 1 to 2 (session:  $F(1,82) = 9.83, p = 0.002, \eta_p^2 = 0.11$ , inclusion BF =  $6.36 \times 10^4$ ), the size of the LWPS effect did not differ between sessions (Figure 16; session  $\times$  trial type  $\times$  PS:  $F(1,82) = 0.00, p = 0.947, \eta_p^2 < 0.001$ , inclusion BF = 0.14; all other effects,  $F < 0.56$ ).

##### 4.1.2.2.2 Error Rates

We again observed a strong main effect of trial type ( $M = 5.7\%$ ;  $F(1,83) = 106.45, p < 0.001, \eta_p^2 = 0.56$ , inclusion BF =  $2.57 \times 10^{19}$ ) that was modulated by PS context ( $F(1,83) =$

14.62,  $p < 0.001$ ,  $\eta_p^2 = 0.15$ , inclusion BF = 31.4). Participants showed smaller switch costs in the PS-70 ( $M = 3.8\%$ ) than PS-30 ( $M = 7.6\%$ ) contexts. Finally, there was little evidence for a change in the size of the LWPS effect between sessions 1 and 2 (Figure 3; session  $\times$  trial type  $\times$  PS:  $F(1,83) = 0.76$ ,  $p = 0.387$ ,  $\eta_p^2 = 0.01$ , inclusion BF = 0.21). All other effects were likewise not significant (session  $\times$  PS:  $F(1,83) = 3.04$ ,  $p = 0.085$ ,  $\eta_p^2 = 0.04$ , inclusion BF = 0.43; other  $F$ s  $< 2.34$ ).





**Figure 16: Behavioral results from Experiment 1 for unbiased (PS-50) items. A:** RT (ms, left panel) and error rates (% , right panel) are shown as a function of proportion switch context (PS-30/PS-70), trial type (task-repeat/task-switch) and session (1/2). On the left of each plot, light grey lines connect mean data points for each participant across trial types (task-repeat (left) connected to task-switch (right); positive slopes indicate switch costs). The boxplots (left: task-repeat, right: task-switch) show a box for the first through third quartiles of the data, a thick black line in the middle of the box to represents the median (of participant mean data), and thin black lines extending from the boxes to indicate the range of non-outlier data. Next to the box plots are the aggregate means ( $\pm$ SEM) for the trial type conditions (task-switch shown in orange, task-repeat in green), which are connected by a dotted line across PS contexts. Finally, next to the means are violin plots that represent the respective

distributions of the trial type data. Session 1 data are shown on the first row, while session 2 data are shown on the second row. B: Switch costs for RT (ms, left panel) and error rates (% , right panel) are displayed as a function of proportion switch context (PS-30/PS-70) and session (1: top/2: bottom). On the left of each plot, mean switch costs for individual participants are connected across proportion switch contexts (PS-30 (left) connected to PS-70 (right); negative slopes indicate an LWPS effect). We again show boxplots and violin plots to represent the respective distributions, with the PS-30 context shown in green on the left and top and the PS-70 context shown in orange on the right and bottom.

#### 4.1.2.3 Discussion

In sum, across both RT and accuracy, and for both item-level and task-level biased stimuli, we observed an LWPS effect that was not modulated by session. In other words, while participants learned to associate switch-readiness with specific stimuli and task sets in each of the two sessions, they did not benefit from a 24-hour consolidation period between session 1 and 2; the LWPS effect was significant and of similar magnitude in sessions 1 and 2. Experiment 2 sought to test whether this lack of a consolidation effect was due to the fact that the predictive items changed between sessions 1 and 2 in Experiment 1.

## 4.2 Experiment 2

The results of Experiment 1 speak against the possibility of participants benefiting from meta-learning of the statistical structure of the task. While robust LWPS effects were observed in both sessions, these effects were not larger in magnitude for session 2 than session 1. However, the possibility remains that consolidation would benefit participants learning that specific stimuli are predictive of the list-wide switch

frequencies (rather than learning the meta demand structure), if the biased stimuli were kept constant across sessions. To address this possibility, we held the biased and unbiased stimuli constant across sessions in Experiment 2 and reexamined the task- and item-specific LWPS effects.

If consolidation were to benefit participants learning that particular stimuli were predictive of switch proportion, we would expect to observe a greater LWPS effect in session 2 than in session 1. If stimulus-specific associations were not impacted by consolidation, as with Experiment 1, we would expect the LWPS to vary little between sessions.

## **4.2.1 Method**

### **4.2.1.1 Participants**

335 MTurk workers consented to participate in the first session for \$0.13/minute. 229 workers were excluded for poor accuracy on the first session<sup>5</sup>. Of the 106 workers invited for the second session, 82 participated and earned the \$1 bonus for completion. We excluded an additional 5 workers for bad accuracy on the second session and 2 workers for old age (greater than 60-years-old). This resulted in a final sample size of 104 MTurk workers for session 1 (mean age =  $37.79 \pm 10.29$ ) and 75 for sessions 1 and 2

---

<sup>5</sup> We recognize that this exclusion rate is larger than normal for MTurk, which typically ranges between 3-37% (Chandler et al., 2014). This large exclusion rate may have been related to the influx of new workers due to COVID-19 (Arechar & Rand, 2020; see also Moss et al., 2020). These workers may thus have been unfamiliar with MTurk demands and norms, particularly the few workers who did the HIT later in the day and on a weekend (Arechar et al., 2017).

(mean age =  $37.62 \pm 10.13$ ). Thus, the attrition rate between sessions 1 and 2 was 22.64% (24/106).

#### 4.2.1.2 Procedure

Experiment 2 served as a conceptual replication of Experiment 1, with one key design change. As before, half of the letters/digits were assigned to the unbiased stimulus condition while the other half were assigned to the biased stimulus condition. Importantly, however, biased and unbiased stimuli were now held constant across sessions for each participant. In other words, the specific set of letters and digits that were predictive of the block-wide switch frequencies in session 1 were also predictive in session 2. The letters/digits in the biased condition for half of the participants were assigned to the unbiased condition for the other half. This design allowed us to test whether consolidation benefits would only be observed when long-term stimulus-switch (and –repeat) associations remained the same.

### 4.2.2 Results

#### 4.2.2.1 Biased, PS-20/PS-80 Items

##### 4.2.2.1.1 Reaction Time

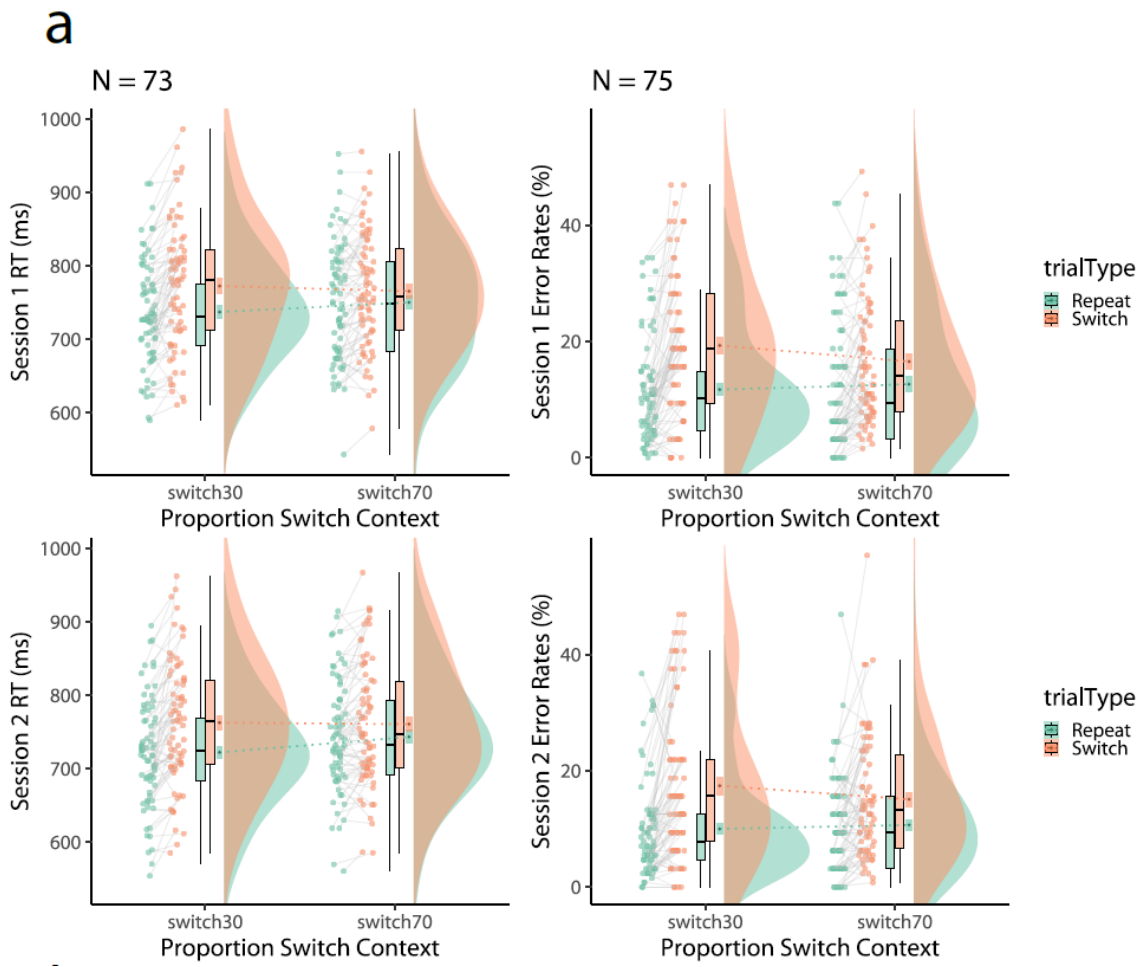
We observed a strong main effect of trial type ( $F(1,72) = 57.04, p < 0.001, \eta_p^2 = 0.44$ , inclusion BF =  $3.20 \times 10^{10}$ ), and these switch costs ( $M = 30$  ms) were modulated by proportion switch context ( $F(1,72) = 33.29, p < 0.001, \eta_p^2 = 0.32$ , inclusion BF = 9.25). Participants had smaller switch costs in the PS-70 ( $M = 15$  ms) than PS-30 ( $M = 37$  ms) contexts. However, the size of the LWPS effect did not differ between session 1 and 2

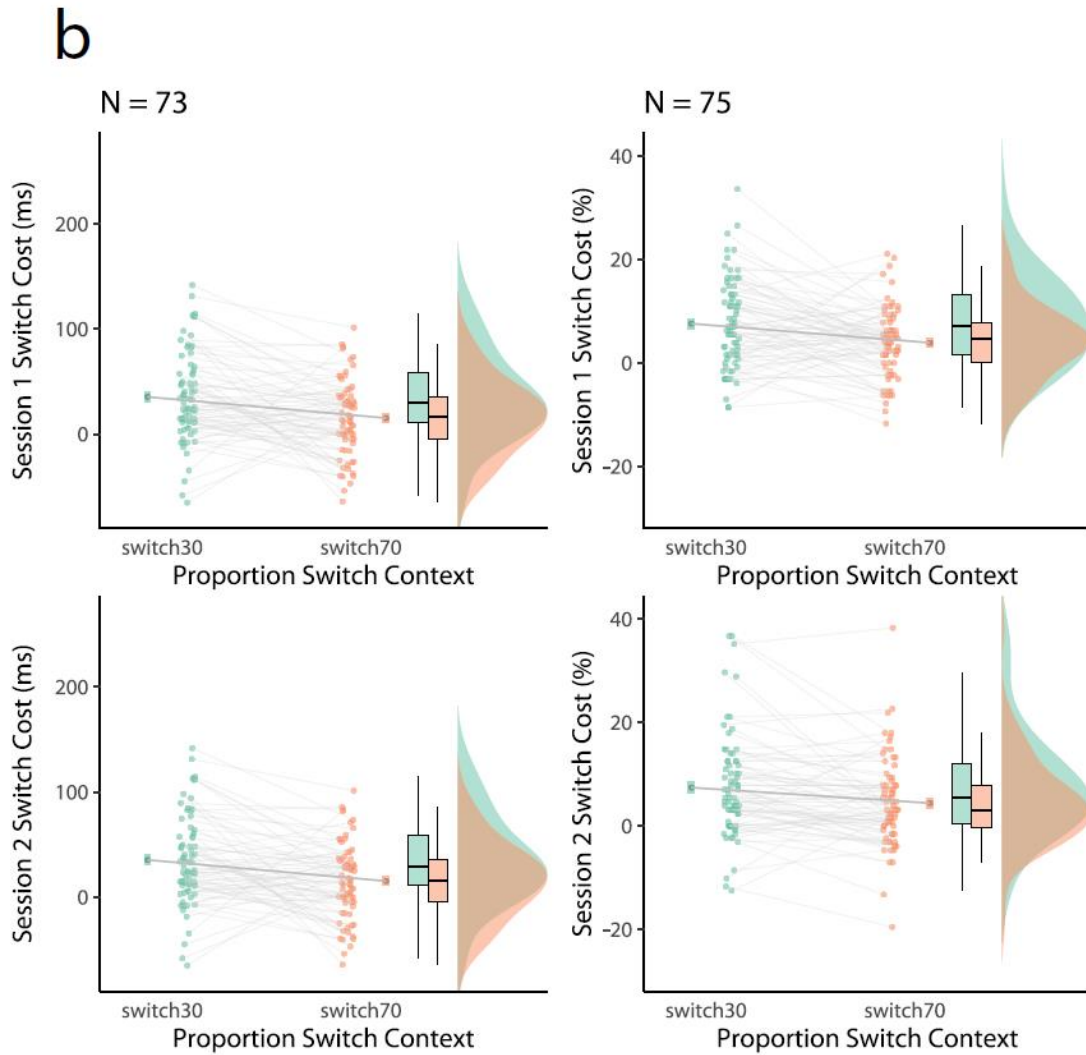
(Figure 17; session x trial type x proportion switch:  $F(1,72) = 0.17, p = 0.681, \eta_p^2 = 0,$  inclusion BF = 0.12) All other effects were likewise not significant (proportion switch:  $F(1,72) = 2.93, p = 0.091, \eta_p^2 = 0.04,$  inclusion BF = 0.46; other  $F_s < 2.44$ ).

#### 4.2.2.1.2 Error Rates

Participants showed strong switch costs ( $M = 5.3\%$ ) (trial type:  $F(1,74) = 79.14, p < 0.001, \eta_p^2 = 0.52,$  inclusion BF =  $1.59 \times 10^{15}$ ) that were modulated by proportion switch context ( $F(1,74) = 21.85, p < 0.001, \eta_p^2 = 0.23,$  inclusion BF = 3.29). Switch costs were smaller in the PS-70 ( $M = 4.2\%$ ) than PS-30 ( $M = 7.5\%$ ) contexts. Participants did not show larger LWPS effects in session 2 than session 1 (Figure 17; session x trial type x PS:  $F(1,74) = 0.22, p = 0.639, \eta_p^2 = 0,$  inclusion BF = 0.17). All other effects were likewise not significant (session:  $F(1,74) = 3.43, p = 0.068, \eta_p^2 = 0.04,$  inclusion BF = 3.1; other  $F_s < 1.66$ ).







**Figure 17: Behavioral results from Experiment 2 for biased (PS-20/80) items. A:** RT (ms, left panel) and error rates (% , right panel) are shown as a function of proportion switch context (PS-30/PS-70), trial type (task-repeat/task-switch) and session (1: top/2: bottom). **B:** Switch costs for RT (ms, left panel) and error rates (% , right panel) are displayed as a function of proportion switch context (PS-30/PS-70) and session (1: top/2: bottom).

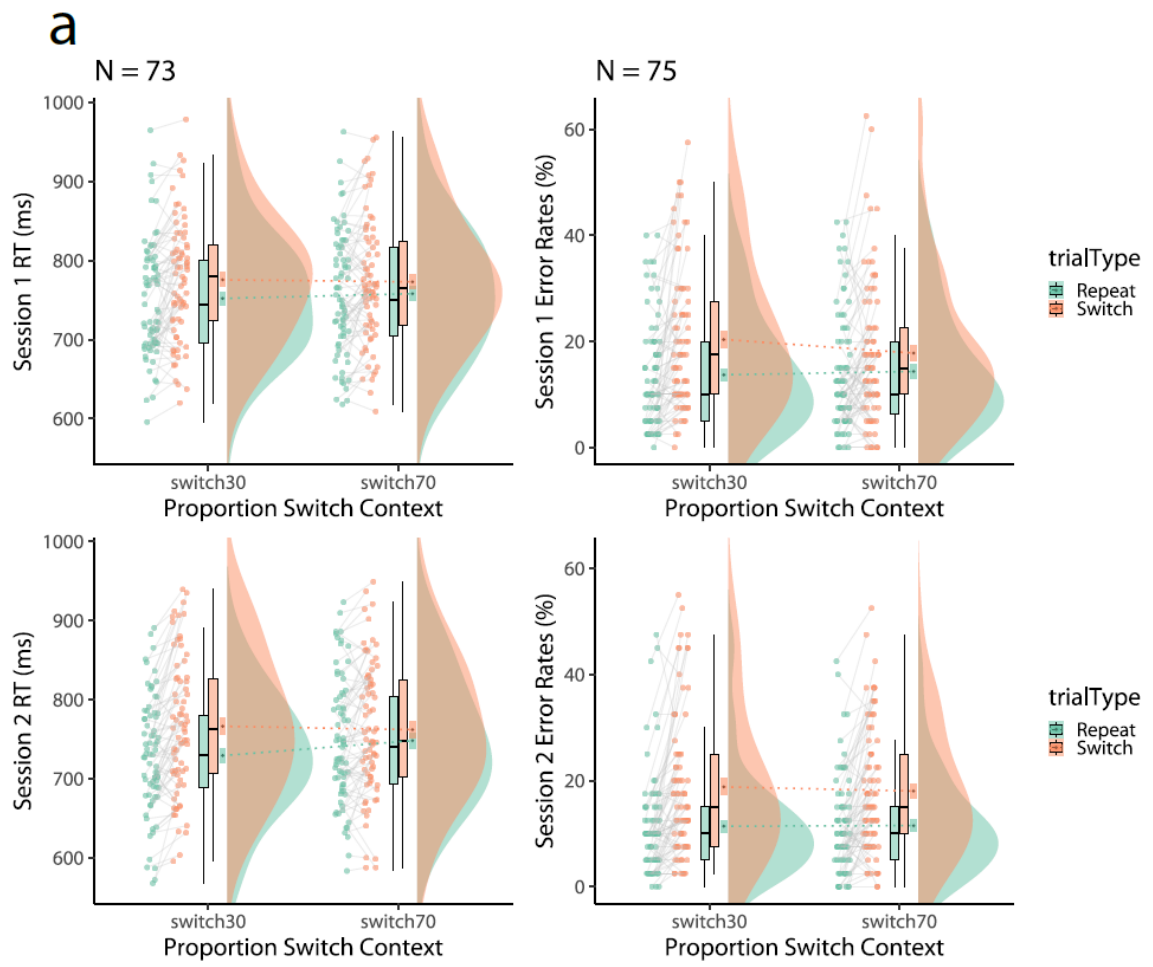
#### 4.2.2.2 Unbiased, PS-50 Items

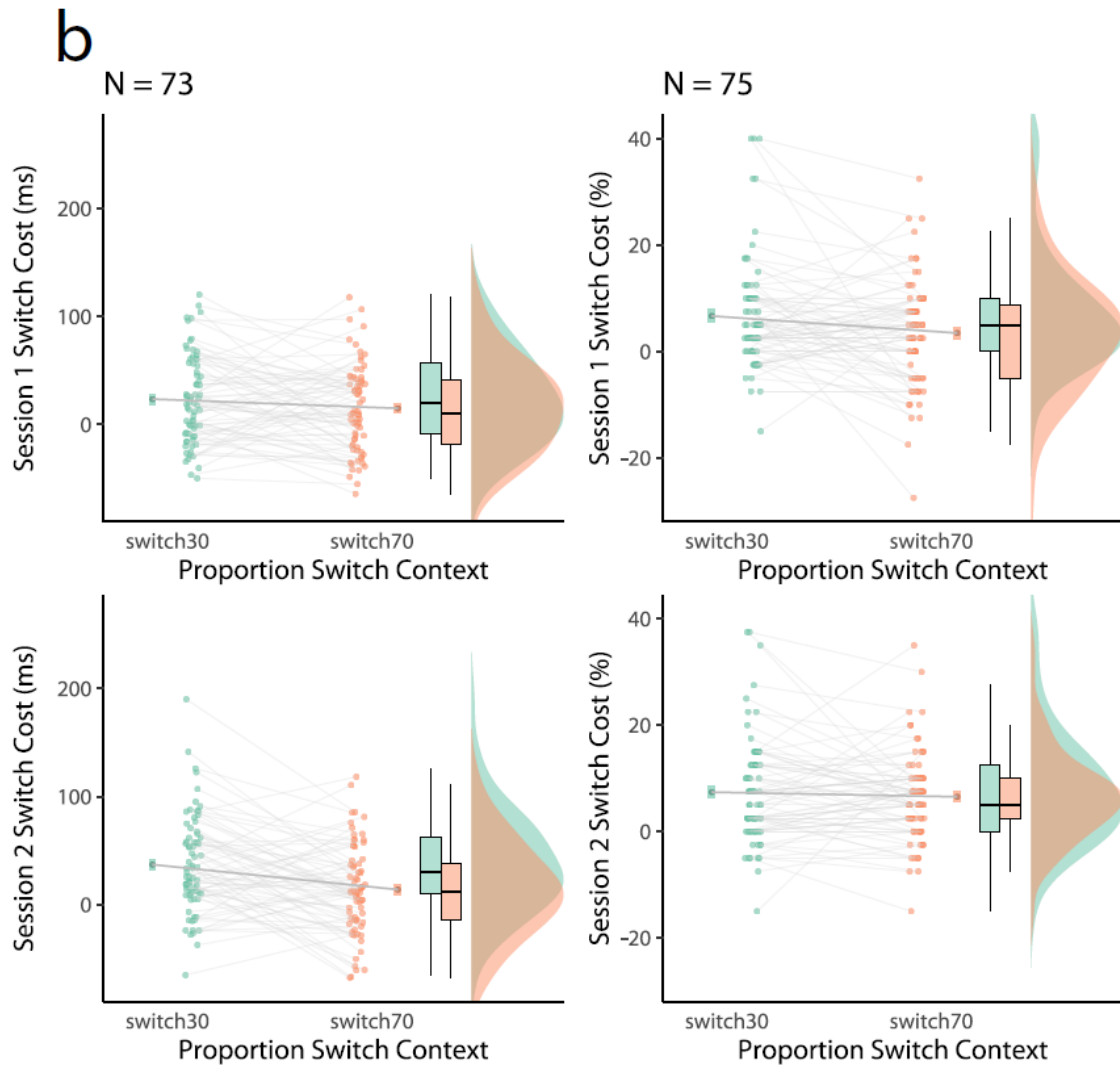
##### 4.2.2.2.1 Reaction Time

We observed a strong main effect of trial type ( $F(1, 72) = 40.47, p < 0.001, \eta_p^2 = 0.36$ , inclusion BF =  $2.04 \times 10^7$ ), and these switch costs ( $M = 24$  ms) were modulated by proportion switch context ( $F(1,72) = 15.21, p < 0.001, \eta_p^2 = 0.17$ , inclusion BF = 1.54). Participants had smaller switch costs in the PS-70 ( $M = 17$  ms) than PS-30 ( $M = 31$  ms) contexts. However, although participants responded faster in session 2 ( $M = 750$  ms) than session 1 ( $M = 765$  ms) (session:  $F(1, 72) = 5.27, p = 0.025, \eta_p^2 = 0.07$ , inclusion BF = 87.39), the size of the LWPS effect did not differ between session 1 and session 2 (Figure 18; session x trial type x proportion switch:  $F(1, 72) = 3.51, p = 0.065, \eta_p^2 = 0.05$ , inclusion BF = 0.34; all other effects,  $F < 2.52$ ).

##### 4.2.2.2.2 Error Rates

Participants showed strong switch costs ( $M = 6.0\%$ ) (trial type:  $F(1,74) = 68.91, p < 0.001, \eta_p^2 = 0.48$ , inclusion BF =  $1.16 \times 10^{14}$ ) that were modulated by proportion switch context ( $F(1,74) = 5.88, p = 0.018, \eta_p^2 = 0.07$ , inclusion BF = 0.36). Switch costs were smaller in the PS-70 ( $M = 5.0\%$ ) than PS-30 ( $M = 7.0\%$ ) contexts. Again, although participants were also slightly more accurate on session 2 than session 1 (session:  $F(1,73) = 2.84, p = 0.096, \eta_p^2 = 0.04$ , inclusion BF = 1.31), they did not show any difference in LWPS effects between session 1 and session 2 (Figure 18; session x trial type x PS:  $F(1,74) = 1.65, p = 0.203, \eta_p^2 = 0.02$ , inclusion BF = 0.29; all other effects,  $F < 2.44$ ).





**Figure 18: Behavioral results from Experiment 2 for unbiased (PS-50) items. A:** RT (ms, left panel) and error rates (% , right panel) are shown as a function of proportion switch context (PS-30/PS-70), trial type (task-repeat/task-switch) and session (1: top/2: bottom). **B:** Switch costs for RT (ms, left panel) and error rates (% , right panel) are displayed as a function of proportion switch context (PS-30/PS-70) and session (1: top/2: bottom).

#### 4.2.2.3 Discussion

In sum, similar to Experiment 1, in Experiment 2 we observed an LWPS effect for

biased and unbiased items that was not modulated by session. Thus, we did not find evidence for benefits of consolidation even when the item-control associations were kept constant between the two sessions.

### **4.3 Between-Experiment Analysis**

To assess the robustness of the LWPS effect with greater statistical power, we combined the data from both experiments and added “Experiment” as a between-participant factor. While the data patterns in the experiments looked similar, this analysis would also reveal any potential differences in the effects of consolidation between the two experiments. All participants who were excluded in Experiments 1 and 2 for having fewer than ten trials per cell were again excluded. Here, we highlight the main LWPS effect and its potential interactions with session and Experiment.

In an initial analysis, we sought to maximize sample size, thus including participants who did not return to complete session 2 (for a total of  $N = 206$ ). We therefore analyzed session 1 data, pooling across both experiments and adding Experiment as an additional factor. We observed a strong LWPS effect (trial type  $\times$  proportion switch: RT,  $F(1,194) = 37.99$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.16$ , inclusion BF =  $1.88 \times 10^3$ ; Error rate,  $F(1,202) = 41.42$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.17$ , inclusion BF =  $7.39 \times 10^4$ ) that did not vary between biased and unbiased items (stim bias  $\times$  trial type  $\times$  proportion switch: RT,  $F < 0.27$ ; Error rate,  $F < 0.70$ ). The size of the LWPS effect did not differ between experiments (experiment  $\times$  trial type  $\times$  proportion switch: RT,  $F < 1.20$ ; Error rate,  $F < 0.04$ ).

Next, we included only participants who contributed data to both sessions, pooled data across experiments, and added Experiment and session as additional factors into the ANOVA. We again found a robust LWPS effect (trial type x proportion switch: RT,  $F(1,152) = 83.15$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.35$ , inclusion BF =  $1.53 \times 10^5$ ; Error rate,  $F(1,157) = 42.34$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.21$ , inclusion BF =  $2.09 \times 10^4$ ) that did not interact significantly either with session (session x trial type x proportion switch: RT,  $F(1,152) = 2.88$ ,  $p = 0.092$ ,  $\eta_p^2 = 0.02$ , inclusion BF = 0.12; Error rate,  $F < 1.15$ ) or with Experiment (RT,  $F < 0.43$ ; Error rate,  $F < 0.61$ ). In sum, the between-experiment analyses corroborated a robust LWPS effect for both biased and unbiased stimuli that was not modulated by a one-day consolidation period or by whether biased items were held constant between sessions.

#### **4.4 Exploratory Analyses**

Finally, we took advantage of the fact that our design included different PS context orders across participants. Some participants experienced the high PS context first whereas others experienced the low PS context first. This allowed us to probe whether the LWPS effect can be observed *between* participants. A between-participant LWPS effect has been reported in in one previous study that did not use a design with unbiased items (Monsell & Mizon, Experiment 4), and in a hybrid free- and forced-choice setting (Fröber & Dreisbach, 2017), as well as in studies of the proportion congruency effect (Spinelli et al., 2019; Bejjani & Egner, 2021). We did this by analyzing only the first PS context (the first four blocks) of each session, that is, before participants

experience a PS context shift, and treating PS context order as a between-participants factor. This meant that about half of the participants contributed their first four blocks to the mostly switch (PS-70) context, while the other half contributed to the mostly repeat (PS-30) context. Because block order was not counterbalanced across session, we analyzed each session separately. We expected that the effect of trial type (i.e., switch cost) would interact with the between-participants PS factor: participants in a low proportion switch context would show greater switch costs than participants in a high proportion switch context.

Accordingly, we analyzed only the first PS context (the first four blocks) of each session and treated block order as a between-participants PS factor. We observed strong evidence for between-participants LWPS effects, both in session 1 (trial type  $\times$  PS: RT,  $F(1,152) = 33.49$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.18$ , inclusion BF =  $8.34 \times 10^7$ ; Error rate,  $F(1,155) = 22.11$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.13$ , inclusion BF =  $1.86 \times 10^5$ ) and in session 2 (trial type  $\times$  PS: RT,  $F(1,154) = 25.59$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.14$ , inclusion BF =  $2.97 \times 10^7$ ; Error rate,  $F(1,155) = 12.14$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.07$ , inclusion BF =  $1.57 \times 10^3$ ). These between-participants LWPS effects were not modulated by Experiment (RT, S1:  $F < 0.83$ , S2:  $F < 1.07$ ; Error rate, S1:  $F < 1.00$ , S2:  $F < 0.16$ ).

Finally, because the design involved a one-day consolidation manipulation, this allowed us to examine test-retest reliability (cf. Enkavi et al., 2019) of repeat and switch trials within each PS context for stimuli across sessions. Since stimulus bias had little



impact in the exploratory and across experiment analyses above, we collapsed across unbiased and biased stimuli. Performance on repeat trials within the PS-30 (RT:  $r(157) = 0.76, p < 0.001$ ; Error rate:  $r(157) = 0.56, p < 0.001$ ) and PS-70 (RT:  $r(157) = 0.72, p < 0.001$ ; Error rate:  $r(157) = 0.47, p < 0.001$ ) contexts as well as on switch trials within the PS-30 (RT:  $r(157) = 0.72, p < 0.001$ ; Error rate:  $r(157) = 0.53, p < 0.001$ ) and PS-70 (RT:  $r(157) = 0.69, p < 0.001$ ; Error rate:  $r(157) = 0.53, p < 0.001$ ) contexts was significantly correlated from session 1 to session 2. Accuracy showed weaker reliability largely because most participants were at ceiling performance; thus, participants whose performance varied dramatically across sessions drove the strength of the correlations. Nonetheless, RT showed acceptable test-retest reliability.

Despite their poor psychometric properties (Draheim et al., 2019), switch costs are often used in individual difference research; we therefore ran a similar analysis on the test-retest reliability of switch costs across session 1 and 2, collapsing across biased and unbiased stimuli. Switch costs within the PS-30 (RT:  $r(157) = 0.47, p < 0.001$ , Error rate:  $r(157) = 0.27, p < 0.001$ ) and PS-70 (RT:  $r(157) = 0.47, p < 0.001$ , Error rate:  $r(157) = 0.23, p = 0.003$ ) contexts showed poor test-retest reliability from session 1 to session 2. Of note, difference scores usually have a lower retest reliability due to inherent variability (Haaf & Rouder, 2019), but we show acceptable test-retest reliability across sessions for RTs in each cell of the trial type x switch proportion interaction. Therefore, if this experiment were repeated, we are fairly confident that these condition-specific means

would be similar, thus reproducing the switch cost adjustments seen in the present study.

#### **4.5 General Discussion**

In this study, we examined whether long-term memory consolidation strengthened the learned association between control settings and contexts (stimuli or task sets). To the best of our knowledge, this is the first control learning study to incorporate a consolidation manipulation. Participants performed a cued task-switching paradigm, alternating between digit and letter classification tasks. Blocks contained either mostly switch or mostly repeat trials. Biased stimuli were either more frequently associated with task-switch or task-repeat trials, while unbiased stimuli were not predictive of task-switch or -repeat trials, but were associated with the block- and task-level switch context created by the biased items. After about 24 hours, participants who performed well on the first session were invited for a second session, where they performed the task-switching paradigm again. This allowed us to test whether the delay period strengthened either item-specific and task-level associations (biased items) or nonpredictive task-level associations (unbiased items). Moreover, in Experiment 1, biased and unbiased items were randomized across sessions while in Experiment 2, they were held constant. With this experimental manipulation, we could thus examine whether any benefits of consolidation impacted “meta-learning” (learning to learn, e.g.,

the task structure in Experiment 1) or depended on specific stimulus associations (Experiment 2).

Across both Experiments, we replicated Siqi-Liu and Egner (2020), observing a robust within-participant list-wide proportion switch (LWPS) effect whereby participants showed reduced switch costs when contexts were associated with more task-switching. This was true for both switch-biased stimuli, demonstrating item-based learning of switch-readiness, and unbiased stimuli, where the effects may stem from associating either the list-wide context or each task set with the varying proportions of switching. Since Siqi-Liu and Egner (2020) showed that list-wide effects do not occur for unbiased “transfer” tasks, we interpret the latter finding as reflecting task set based associations with switch likelihood.

Further, we replicated Monsell & Mizon’s (2006, Exp 4) finding of between-participant LWPS effect in a more high-powered sample and using unbiased items when we analyzed the first four experimental blocks alone and treated list-wide proportion switch context as a between-participants factor. This finding had also previously been documented for the list-wide proportion congruent effect (Bejjani & Egner, in press; Spinelli et al., 2019) and in a hybrid free- and forced- choice setting (Fröber and Dreisbach, 2017, supplemental materials).

In addition to providing a high-powered replication of the effects observed in Siqi-Liu and Egner (2020) and Monsell & Mizon (2006, Exp 4), the present study reports

two novel findings. First, we tested the prediction that consolidation would strengthen learned context-control associations, which should have resulted in a larger LWPS effect in session 2 than session 1. However, we found little impact of session on the size of the LWPS effect, regardless of either Experiment or biased and unbiased stimuli. This indicates that robust switch-readiness learning occurred in both sessions, but consolidation of learning in session 1 did not benefit switch-readiness learning in session 2.

Second, the current study design allowed us to examine test-retest reliability in the LWPS paradigm across sessions, which had not previously been assessed. We observed acceptable levels of test-retest reliability for response time, but not difference score measures in this protocol, which is encouraging for the potential suitability of using control learning tasks as individual difference measures in future research.

Arguably, one limitation to our experimental design was that we did not include a group of participants who completed session 2 immediately after session 1 as a control group for immediate practice effects that occur in absence of overnight memory consolidation. To address this limitation, we investigated the effect of practice in this paradigm in two additional analyses. First, the ideal approach would be to probe for a 3-way interaction between block, trial type, and PS, assessing whether the LWPS effect (represented by the 2-way interaction between trial type and PS) becomes larger with practice (as a function of block). This approach was not possible for the current data sets,

as the PS factor was confounded with the block factor, with participants either completing all the PS-30 blocks first or all the PS-70 blocks first, rather than having alternating blocks of the two conditions. As an alternative approach, we combined data from the two experiments in the current dataset and looked at interaction between block and trial type as well as session. We plotted this effect separately for participants who completed PS-30 first (MR-first) versus those who completed PS-70 first (MS-first) in Appendix B, **Figure 23**. In both sessions, switch costs were similar for blocks belonging to the same PS context and did not show a linear increase or decrease within PS context across individual blocks, which would have indicated practice effects (see statistics within Appendix B, **Table 8**).

Second, given the limitations of the above analytic approach, we also analyzed data collected for Siqu-Liu, Egner, & Woldorff (in prep.), which employs a nearly identical paradigm consisting of the same stimuli and frequencies as in the current study (though without an unbiased stimulus set). Crucially, in Siqu-Liu, Egner, & Woldorff (in prep.), PS-30 and PS-70 blocks were presented in an alternating manner (e.g., PS-30 then PS-70 then PS-30 then PS-70), unlike in the current study; we could thus test for how practice affects within-participant LWPS using the 3-way interaction test described above. We found that the LWPS effect did not differ in the first versus second half of trials participants completed (each half consisting of 488 trials) in RTs or error rate, and neither did switch costs. These results from a near-identical paradigm suggest that

switch cost adaptation occurs early in the experiment and does not increase significantly due to practice over time. We thus have strong grounds to believe that the LWPS effect in the current study would not change if participants had completed session 1 and 2 in a single, longer session.

Together, these findings suggest that the acquisition of stimulus-switch-readiness associations is stable within participants, and can be learned relatively quickly (cf. Bejjani & Egner, 2021), but is not benefited by long-term memory consolidation (or “spacing” of learning). In the following, we discuss different possible reasons for why learned switch-readiness may not have benefited from the consolidation period in our study.

#### **4.5.1 Benefits of consolidation may occur for reward-learning, but not control learning**

One explanation for the lack of consolidation benefits in the current study, even though such benefits are well established in the reward-learning literature, is that control learning may differ from reward-learning in important ways.

First, a key notable distinction with typical reward learning paradigms is that the learning of switch-readiness takes place at a more abstract level. That is, rather than associating a particular stimulus with a specific motor response or a concrete reward value, participants in our protocol (and other typical control learning tasks) learn to link contextual cues with an abstract (generalizable) internal cognitive control setting (switch-readiness). Second, in addition to differing in terms of the learning target

(switch-readiness vs. stimulus-choice associations), control and reward learning also involve distinct experimental procedures. Importantly, in reward-learning tasks, there is typically a single “correct” choice (response) to a given stimulus for obtaining the maximum reward, and participants usually receive immediate performance feedback on whether they have made the correct choice. By contrast, in control learning tasks, there is no explicit “correct” level of control to apply in a certain context, and performance feedback (if given) relates to whether the response rather than the level of control applied was correct (for additional discussion, see Bejjani, Tan, and Egner, 2020). Learning about, and being reinforced by, a quantifiable reward amount may differ from learning about a less concrete amount of mental effort to be exerted, especially in terms of the final outcome observed. Given this more complex nature of control learning, it is possible that more delay time or additional retrieval/practice would be needed to strengthen long-term stimulus-control associations than is needed for stimulus-reward associations (cf. Wimmer et al., 2018; Miendlarzewska et al., 2018).

Moreover, the lack of explicit correct choices and feedback has two consequences that may have rendered consolidation effects unobservable in the current study. First, unlike reward tasks, we cannot analyze learning curves for the proportion of “correct” choices over time. To understand how participants learn the stimulus-switching contingencies on a trial-by-trial level, we would have to train models capable of capturing the latent dynamics of learning, such as a Gaussian Process model (McDonald

et al., 2019). Second, we also don't have measures of implicit or explicit value, focusing instead on reaction time and error rate. Thus, while studies have shown benefits of consolidation on reward (Braun et al., 2018; Patil et al., 2017; Wimmer et al., 2018) and cognition generally (Biderman et al., 2020; Dudai, 2012; Tambini & Davachi, 2019), one possibility is that consolidation quickened the predictive learning or altered the value of the stimuli, which our current design cannot fully measure. Overall, an important direction for future research could therefore be to assess consolidation effects in control learning experiments that focus on choice data (e.g., Kool et al., 2010), in order to more closely approximate reward learning studies.

Finally, it should be noted that neither control learning, nor reward learning are mediated by single learning mechanisms. Behavioral signatures of control learning can be successfully modeled using an incremental reinforcement learner that nudges control parameters (e.g., switch-readiness) up or down based on ongoing variations in control demand, but without encoding specifics about the stimulus or response features associated with each trial (e.g., Botvinick et al., 2001; Jiang et al., 2014, 2015). On the other hand, control learning can also be driven by episodic memory mechanisms (e.g., Blais et al., 2007; Verguts & Notebaert, 2009; Whitehead et al., 2020), such that perceptual and motor features in memory, as well as concurrent internal states, including our level of switch-readiness, are bound into an episodic "event file" (Hommel et al., 2001). This



interplay between incremental and episodic learning mechanisms is also key to reward learning (Biderman et al., 2020; Gershman & Daw, 2017; Shohamy & Daw, 2015).

Different types of memory consolidation could have different effects on these two types of learning mechanisms. Consolidation can either occur offline, via sleep-dependent replay, or online, via reactivation of memories (Antony et al., 2017). That is, memories could be strengthened through repeated online testing of material (i.e., “testing effect”) or through greater offline processing of learned associations during sleep. A recent study suggests that sleep could have diverse effects on learning. For instance, sleep-dependent memory benefits were not observed for a cross-modal information-integration paradigm (implicit, practice-based learning), but were present for paired-associate learning (episodic learning), and the consolidation delay facilitated subsequent category learning (Ashton et al., 2018). Since the current study employs both substantive practice within sessions and a 24-hour delay (during which presumably participants slept), we cannot dissociate the effects of offline versus online consolidation. It is also well-established that sleep benefits (motor) skill learning (Kuriyama, Stickgold, & Walker, 2004; Walker et al., 2002; Walker et al., 2003); however, the lack of benefits observed in the present data sets implies that control learning relies on different mechanisms than the acquisition of procedural skills. We also cannot draw any conclusions about the effects of sleep on consolidation because we neither manipulated, nor collected data on sleep quality. Nonetheless, teasing apart the relationship between

offline versus online consolidation and episodic versus incremental learning is an important target for future studies.

#### **4.5.2 Benefits of consolidation may depend on the type of control learning**

In addition to differences between control and reward learning paradigms, it is also possible that the lack of a consolidation benefit observed here is related to the specific control learning task employed in the current study. While there is evidence that people can learn control-demand associations both through incremental learning and episodic mechanisms (i.e., retrieving specific previous memory traces), it is presently not clear under what circumstances either of these mechanisms exerts a greater influence. A handful of studies have found evidence for one-shot learning of control in the context of trial-unique events (Brosowsky & Crump, 2018; Spinelli et al., 2019; Whitehead et al., 2020). However, the vast majority of control learning studies (including the present one) use small stimulus sets that do not allow researchers to dissociate between potentially simultaneous incremental and episodic learning mechanisms. It is therefore possible that the present protocol involves primarily incremental, non-episodic learning, and that this form of learning is less susceptible to the effects of a one-day consolidation period than episodic memory. This possibility could be tested in the future by combining a “one-shot” control learning task (e.g., Whitehead et al., 2020) with a consolidation manipulation.

Second, the present protocol differs from some other control learning tasks in that it does not involve fixed stimulus-control associations. Rather, in the present study, stimulus-control associations had to be relearned between task phases and days, whereby in Experiment 1 only the overall statistical structure of the task could be learned and consolidated. In Experiment 2, a subset of stimuli was consistently predictive of control demands, but those demands varied over time. It is therefore possible that control learning protocols with non-variable stimulus-control associations would benefit from consolidation. To test this, future studies could use similar multiday paradigms where stimuli-specific associations do not have to be relearned, such as in an item-specific switch proportion (ISSP) paradigm (Chiu & Egner, 2017). In the ISSP task, block-wide PS is held constant, and a subset of stimuli occur more often as task-switches while others occur more often as task-repeats. Thus, if consolidation benefited item-specific learning, researchers would observe stronger effects for these task-switch or task-repeat biased stimuli over time. Additionally, proportion congruency (PC) paradigms, such as Bejjani and colleagues (2020), where stimuli and contexts are associated more frequently with congruent or incongruent trials instead of task-switches or task-repeats, could be incorporated into a multiday design to investigate whether consolidation benefits other types of stimulus-control learning. In contrast to the ISSP task, PC paradigms also often involve stimulus-response associations, which allows

researchers to examine whether these direct mappings – more similar to explicit reward associations – benefit from consolidation.

Finally, another plausible factor in mediating consolidation effects might be participants' awareness of the switch frequency manipulations. The extent to which awareness plays a role in control learning in general remains debated. For example, Farooqui and Manly (2015) found that switch costs were most reduced when participants could use subliminal cues to (subconsciously) anticipate upcoming control-demand. However, when Bejjani and colleagues (2020) ran a preregistered direct replication of this study, they found that participants could only modulate their switch costs when they consciously perceived and knew what the task-relevant predictive cues meant. Whether conscious perception and explicit knowledge of task structure impact control learning may also vary in the type of proportion manipulation used in the current study. In many such studies, participants do not become aware of the proportion manipulations (e.g., Bejjani et al., 2018; King et al., 2012), or only explicitly notice the manipulation at its extremes (e.g., with 95% task-switch or task-repeat trials) (cf. Blais et al., 2012). Accordingly, theories of control learning tend to assume that this learning is largely implicit (Abrahamse et al., 2016; Egner, 2014), and though we did not measure participants' awareness, we assume that the learning of switch proportions in the present study was likely implicit as well. The possibility thus remains that we could have observed differences due to consolidation if participants had been made aware of

the stimulus-level manipulation. The latter might be accomplished more easily in tasks where the stimulus-control associations remain constant over time, like in the ISSP protocol discussed above (Chiu & Egner, 2017). Directly assessing the impact of contingency awareness on consolidation effects represents an interesting topic for future study.

### **4.5.3 Conclusions**

The present study examined whether consolidation strengthens learned stimulus-switching associations by having participants perform a list-wide proportion switch paradigm across two sessions separated by a 24-hour delay. We observed a robust within-participant and between-participant LWPS effect for both item-level and task-set based control associations, and acceptable test-retest reliability in response time measures. However, learned switch-readiness did not benefit from a 24-hour consolidation period, thus pointing to important potential difference in the mechanisms underlying control versus reward learning.

In conjunction, Chapters 2, 3, and 4 present a coherent picture of stable list-wide flexibility learning effects that are 1) ambivalent to stimulus associations, 2) specifically tied to task-level learning, and 3) behave differently than reward learning. Chapter 5 further investigates boundary conditions within LWPS paradigms where we may expect cross-stimuli or cross-task transfer to occur.

## **5. Task sets define boundaries of learned cognitive flexibility in list-wide switch proportion manipulations**

In the current of series of experiment, we conducted additional tests of the hypothesis that flexibility learning is task-specific in the LWPS paradigms.

First, to further determine the extent that flexibility learning is agnostic to stimulus-level associations, in Experiment 1, we employed trial-unique stimuli to investigate whether the LWPS adjustments of switch cost would occur for completely novel stimuli that have never been associated with task switches or repeats. This represents a stronger test of the stimulus-independence of the LWPS effect than that supplied in Siqi-Liu & Egner (2020). We hypothesized that we would replicate the LWPS effect even when stimulus-level control associations are completely non-existent.

Second, though Siqi-Liu & Egner (2020, Exp 3a-b) suggested that flexibility learning is task-specific, it remains unclear whether people's switch-readiness in this study was linked to the physical task cues, which were also biased, or to the abstract task-set representations, i.e., the rules that define stimulus-response associations (Kiesel et al., 2010; Monsell, 2003; Rogers & Monsell, 1995). In Experiment 2, we asked participants to switch between two biased tasks that were each associated with two cues. One of the cues for each task was biased, i.e., it occurred more frequently on switch trials in high PS contexts and more frequently on repeat trials in low PS contexts, and the other cue was presented equally often on switch and repeat trials, regardless of PS context. Assuming that flexibility learning is tied to the actual task sets rather than the

task cues, we expected to see switch cost adjustments to PS for both biased and unbiased cues.

Third, Siqi-Liu & Egner (2020, Exp 3a-b) found that flexibility learning did not transfer across tasks that used distinct sets of stimuli (categorizing numbers vs. letters vs. colors) or stimulus features that were spatially segregated. It is therefore in principle still possible that a flexibility transfer effect to the unbiased task could be observed for tasks operating on overlapping features of the same, integrated stimulus. In Experiment 3, we investigated whether flexibility transfer could occur when the transfer task was performed on the same, integrated stimuli as the biased tasks (using age/gender/emotion dimensions of face stimuli).

### **5.1 Experiment 1**

In Siqi-Liu & Egner (2020, Exp 4), an unbiased stimulus set was used to probe whether flexibility adjustments depended on stimulus-level switch/repeat associations. Even though stimuli in the unbiased set were presented equally often as switch and repeat trials, some form of biased stimulus-switch associations may in theory still form if participants did not fully process the stimulus on some trials (e.g., missing the trial completely or making a perceptual error). Moreover, in all experiments of the prior study by Siqi-Liu & Egner (2020) as well as in the LWPS literature at large (Bonnin et al., 2011; Dreisbach & Haider, 2006; Dreisbach et al., 2002; Duthoo et al., 2012; Mayr, 2006; Monsell & Mizon, 2006; Siqi-Liu & Egner, 2020), there were always recurring stimuli, at

least some of which were associated with a switch- or repeat-bias. It is therefore unknown whether an LWPS effect can be observed in the complete absence of any stimulus repetitions, and thus of any stimulus-switch associations.

In the current experiment, we used trial-unique stimuli to conduct a more rigorous test of the hypothesis that stimulus-level associations are not necessary for flexibility adjustments. Since each stimulus only appeared once, there was no opportunity for stimulus-switch associations to form. We used the same basic design as Siqi-Liu & Egner (2020, Exp 2) and manipulated LWPS context while participants switched between two object categorization tasks. Each trial consisted of a unique object stimulus. We hypothesized that participants would nevertheless exhibit lower switch costs in the high compared to low PS contexts, providing more decisive evidence that flexibility adjustments do not depend on any associations between specific stimuli and control states.

## **5.1.1 Method**

### **5.1.1.1 Participants**

We conducted a power analysis (G\*Power 3.1; Faul et al., 2009) using the effect size for the LWPS effect in response times (RT) found in Siqi-Liu & Egner (2020, Exp 2, partial eta squared ( $\eta_p^2$ ) = 0.28), which employed a similar cued-switching paradigm involving two tasks. This revealed that a sample size of 23 participants would be necessary for 80% statistical power to detect the effect at  $p < 0.05$ . We therefore targeted a sample size of 30 participants from the Duke University Department of Psychology and



Neuroscience Subject Pool who surpassed an accuracy inclusion criterion of > 65%.

There were 20 female participants and 10 male participants. Average age was 19.2 with a standard deviation (*SD*) of 1.03 years. All participants gave informed consent and received payment or course credit in accordance with a protocol approved by the Duke University Institutional Review Board.

### **5.1.1.2 Stimuli**

Experimental stimuli consisted of 427 pictures of objects gathered using Google Image Search (taken from Wen & Egner, 2022, Exp 3 & 4). Objects were either manmade (e.g., a car) or natural (e.g., a tree) and either smaller than a shoe box (e.g., a key) or larger than a shoebox (e.g., a house), resulting in four stimulus categories. There were 107 each of small-manmade, large-manmade, large-natural, and 106 small-natural objects.

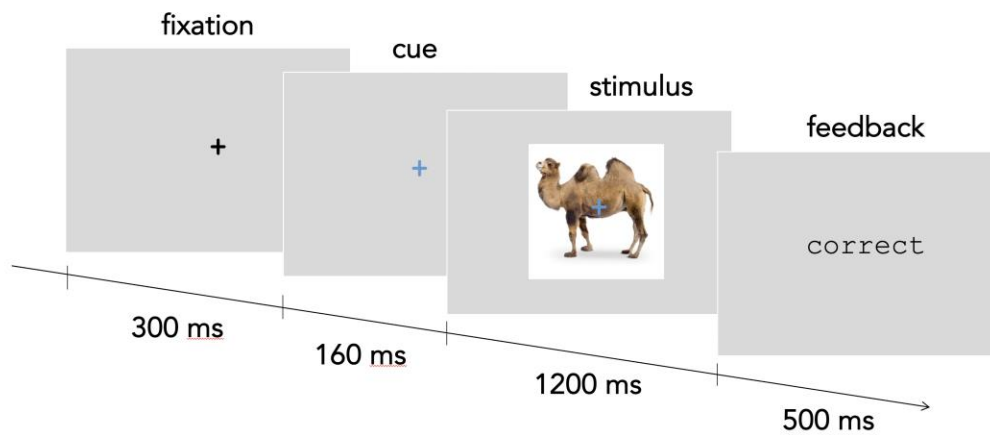
### **5.1.1.3 Procedures**

Each trial began with a central fixation cross displayed in black on a grey background for 300 ms. The fixation cross then turned green or blue to cue the upcoming task. If the fixation turned green, participants had to perform the *origin* task and categorize the upcoming object as natural or manmade. If the fixation turned blue, participants had to perform the *size* task and categorize the object as larger or smaller than a shoebox. After 160 ms, the object stimulus was displayed centrally behind the fixation cross. The object was randomly selected on each trial without replacement from the 427 stimuli. The fixation and stimulus stayed on screen for 1200 ms, during which participants used left or right arrow key presses to indicate their response to the task.

Responses during the stimulus presentation duration were logged, and feedback (“correct”, “incorrect”) was presented after stimulus offset for 500 ms.

Participants completed 1 practice block and 6 experimental blocks of 61 trials each (427 trials total). The practice block consisted of 50% switch trials and 50% repeat trials. Half of the experimental blocks had a low (30%) proportion of switches (PS) and half had a high (70%) PS. Participants were randomly assigned to complete all low or all high PS blocks first. A reminder of the correct response mappings was displayed between the blocks. Both tasks were presented an equal number of times in each block. All trial sequences were pseudo-randomly generated by a JavaScript algorithm.

We report how we determined our sample size, data exclusions, manipulations, and experimental measures, and follow JARS (Kazak, 2018). Data, task and analysis code for all experiments are available at the project’s Open Science Framework page (<https://osf.io/ueqa3/>). Data were analyzed using R, version 3.6.1 (R Core Team, 2019). The study’s design and analysis were not preregistered.



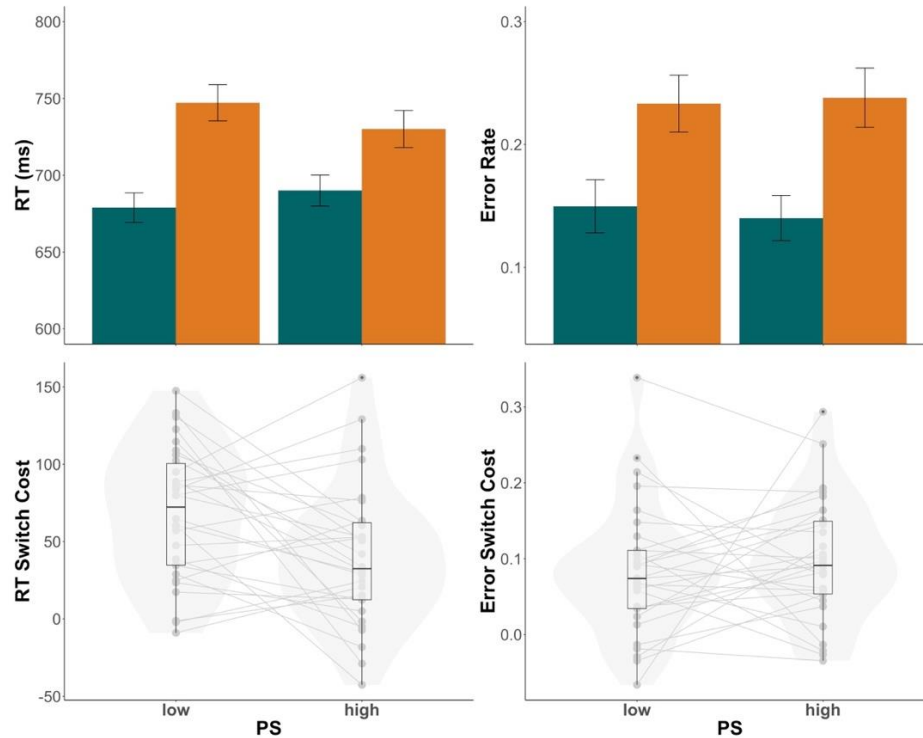
**Figure 19: Example stimulus and trial timing of Experiment 1 protocol**

### 5.1.2 Results

To assess performance accuracy, we analyzed data from all trials after excluding practice blocks and the first trial of each experimental block. For RT analyses, we additionally excluded incorrect trials and post-error trials. After applying these exclusion criteria, trials with response times outside 1.5 times the interquartile range were filtered out.

We conducted a 2 (PS: low v. high)  $\times$  2 (trial type: repeat v. switch) mixed ANOVA on RTs and accuracy. The data are summarized in **Figure 20**. Responses were slower ( $F(1,29) = 73.93$ ,  $p < .001$ ,  $pes = .72$ ) and more error-prone ( $F(1,29) = 62.82$ ,  $p < .001$ ,  $pes = .68$ ) on switch ( $M_{rt} = 729.9.1$  ms;  $M_{error} = 0.24$ ) compared to repeat trials ( $M_{rt} = 679.2$  ms;  $M_{error} = 0.15$ ). Consistent with previous reports of the LWPS effect, trial type significantly interacted with PS in response times ( $F(1,29) = 7.72$ ,  $p = .009$ ,  $pes = .21$ ) but not error rates. This interaction reflected smaller RT switch costs in high PS blocks

( $M_{\text{switchcost}} = 40.0$  ms) compared to low PS blocks ( $M_{\text{switchcost}} = 68.3$  ms), representing the LWPS effect.



**Figure 20: Experiment 1 results. Top panels depict group mean RTs (top left) and error rates (top right) in low (30%) versus high (70%) proportion switch (PS) conditions. Bottom panels depict switch costs for RTs (bottom left) and error rates (bottom right), which were calculated as switch - repeat. Individual dots represent individual mean switch costs. Violin plots are overlaid to visualize the distribution of individual means. In the box plots, central lines depict group medians, box edges show the interquartile range (IQR), and the length of the whiskers mark 1.5 x IQR.**

### 5.1.3 Discussion

To probe whether flexibility adjustments to varying PS could be obtained in the absence of any possible stimulus-task or stimulus-switch associations, Experiment 1 consisted of a standard LWPS design but with trial-unique stimuli. The fact that we nevertheless obtained a typical LWPS effect provides further evidence that the

contextual adaptation of switch-readiness can occur independently of any stimulus-level learning of control demands. Rather, flexibility learning seemed to transfer to completely novel stimuli upon first exposure.

These findings are consistent with results from the conflict control literature with experiments that utilize conceptually similar designs of manipulating the proportion of high conflict trials in a list to measure conflict adaptation effects (“list-wide proportion congruency effects”; for review see (Bugg & Crump, 2012). Using nonrepeating stimuli in a picture-word interference paradigm, Spinelli et al. (2019) observed proportion congruent effects in the absence of stimulus-level associations, similar to our observations of flexibility adaptations with trial-unique task-switching stimuli. In sum, the Experiment 1 results indicate that list-wide adaptation of flexibility does not rely on any kind of stimulus-level contingency learning.

## **5.2 Experiment 2**

Just as target stimuli could be associated with switch- or repeat biases, task-cues are stimulus features of a trial that could also be targets for flexibility learning. For instance, Bejjani et al. (2018) found that task-irrelevant scenes that were either more frequently associated with congruent or incongruent trials cued control adaptations to upcoming Stroop-task stimuli, such that smaller conflict was found in Stroop trials preceded by scenes that were mostly paired with incongruent stimuli. Since trial features preceding a target stimulus can thus cue control processes, it is possible that

previous observations of task-specific flexibility learning in the LWPS paradigm could be the result of learned flexibility associations with physical features of the task-cue, rather than the abstract task-set (i.e., the set of rules that define stimulus-response associations). Although Siqu-Liu & Egnér (2020, Exp 3a-b) included separate biased and unbiased tasks, within the biased task-sets, the relevant task cues were also more frequently associated with task switches in high PS contexts and with task repeats in the low PS contexts. In other words, participants could have also retrieved specific flexibility states in response to the physical task cues (e.g., the word “letter”), rather than retrieving them in response to the meaning of the cue, or the cued task set.

To determine whether flexibility learning depends on switch/repeat-associations with specific task cues, in the current experiment, we isolated the effects of cue-bias and investigated whether Siqu-Liu & Egnér’s (2020, Exp 3) pattern of results could be replicated with biased and unbiased task cues rather than biased and unbiased tasks. Based on the assumption that learned switch-readiness is linked to task sets rather than their (arbitrary) cues, we predicted that flexibility state *would* transfer to unbiased task cues, even though it did not transfer to unbiased task sets (Siqu-Liu & Egnér, 2020). This would indicate that flexibility learning not tied the specific visual features that cue the task and provide further evidence that flexibility learning is specifically tied to the task rules themselves.

## **5.2.1 Method**

### **5.2.1.1 Participants**

We conducted a power analysis (G\*Power 3.1; Faul et al., 2009) based on the effect size for the three-way interaction between trial type x PS x task bias in RT data found in Siqu-Liu & Egner (2020, Exp 3a,  $\eta_p^2 = 0.09$ ). This revealed that a sample size of 82 participants would be necessary for 80% statistical power to detect the effect at  $p < 0.05$ . Based on these power analysis results, our final sample for data analysis consisted of 83 participants from the Duke University Department of Psychology and Neuroscience Subject Pool who surpassed an accuracy inclusion criterion of  $> 65\%$ . There were 53 female participants, 27 male participants, and 3 participants who did not provide gender information. Average age of participants was 18.9 with an *SD* of 0.98 years. All participants gave informed consent and received payment or course credit in accordance with a protocol approved by the Duke University Institutional Review Board.

### **5.2.1.2 Stimuli**

Participants switched between letter (“is the letter a consonant or vowel?”) and digit (is the number odd or even?) classification tasks. The letter task was cued by the words “letter” and “alphabet” and the digit task was cued by the words “digit” and “number.” For each participant, one of the cues for each task was randomly selected to be a biased cue and one was selected as an unbiased transfer cue.

On each trial, one of eight letters (A, E, I, U, G, K, M, R) and one of eight digits (2, 3, 4, 5, 6, 7, 8, 9) were randomly selected and presented simultaneously to the left and right of a fixation cross. The location of letter and digit stimuli was randomized on a trial-by-trial basis.

### **5.2.1.3 Procedures**

Trial timing parameters followed Siqu-Liu & Egnér (2020, Exp 3a). Each trial began with a blank screen lasting 1010 ms, followed by a fixation cross displayed for 450 ms. The cue was then displayed for 160 ms, followed by another blank screen lasting 40 ms, and a stimulus display of 1300 ms during which participants used a 'd' or 'k' key press to indicate whether the stimulus was a consonant/vowel or odd/even based on instructions received at the beginning of the task. Responses during the stimulus presentation duration were logged, and feedback ("correct", "incorrect") was presented after stimulus offset for 500 ms.

All participants completed 1 practice block and 14 experimental blocks of 33 trials each. The practice block consisted of 50% switch trials and 50% repeat trials. Half of the experimental blocks had a low (30%) proportion of switches (PS) and half had a high (70%) PS. Participants were randomly assigned to complete either all the low or all the high PS blocks first. In between blocks, participants could take a short self-paced break, during which a reminder of the response mappings was displayed.



In the low PS blocks, each task was presented 16 times total, 11 times as repeat trials and five times as switch trials. Of the 11 times each task appeared as a repeat trial, it was cued by the biased cue seven times and the unbiased cue four times. Out of the five times each task appeared as a switch trial, it was cued by the biased cue one time and the unbiased cue four times. Thus, the unbiased cue was associated with switch and repeat trials an equal number of times, while the biased cue appeared more often in association with repeat more than with switch trials in the low PS blocks. The same trial counts were used for the high PS blocks, only with the count for switch and repeat trials reversed. Thus, the unbiased cue was equally often associated with switch and repeat trials whereas the biased cue appeared more often in association with switch than with repeat trials in high PS blocks. These trial sequences were pseudo-randomly generated.

### 5.2.2 Results

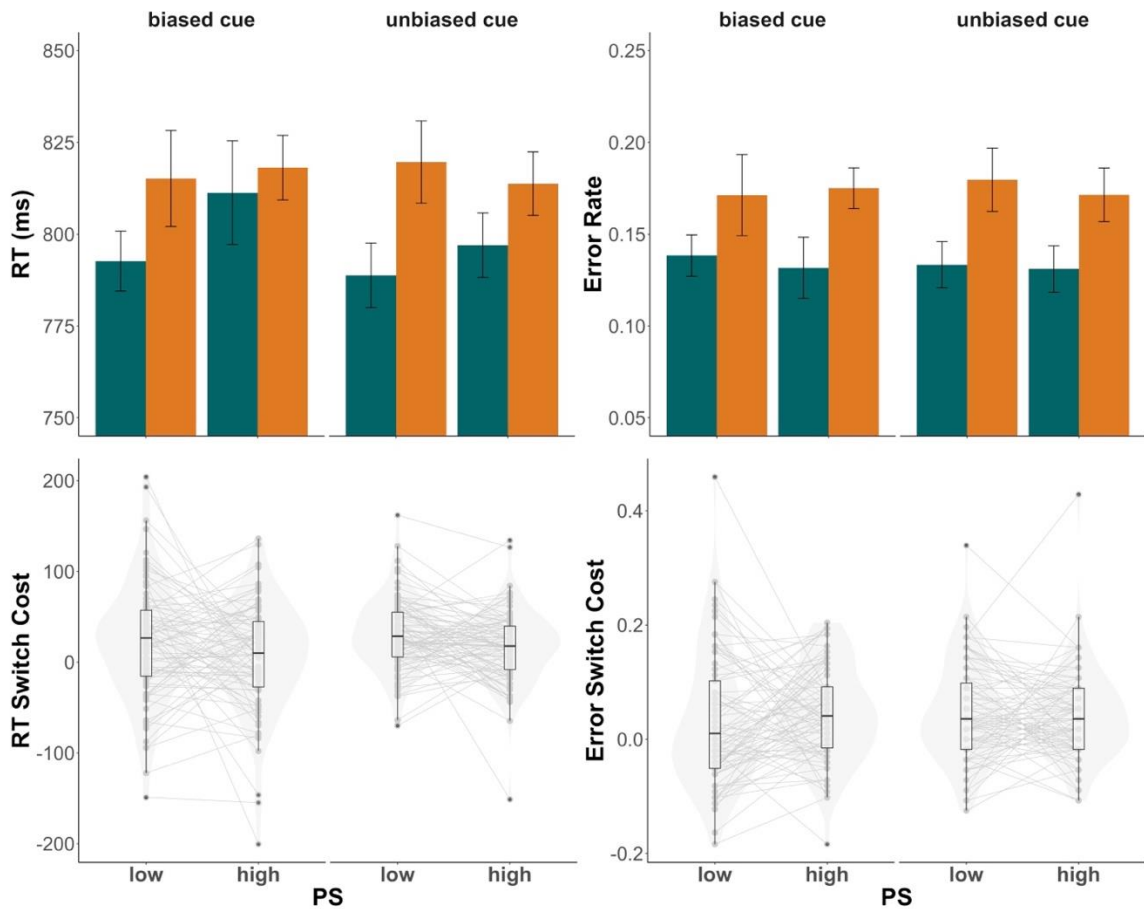
The same exclusion criteria for accuracy and RT analyses as Experiment 1 were applied. We conducted a 2 (PS: high v. low)  $\times$  2 (trial type: switch v. repeat)  $\times$  2 (cue bias: biased v. transfer) mixed ANOVA on response times and accuracy. The data are summarized in **Figure 21**.

Responses were slower ( $F(1,82) = 32.39, p < .001, \eta^2 = .28$ ) and more error-prone ( $F(1,82) = 49.09, p < .001, \eta^2 = .37$ ) on switch ( $M_{rt} = 812.1$  ms;  $M_{error} = 0.18$ ) compared to repeat trials ( $M_{rt} = 789.1$  ms;  $M_{error} = 0.13$ ). Trial type significantly interacted with PS in response times ( $F(1,82) = 7.69, p = .007, \eta^2 = .10$ ) but not error rates. Consistent with

previous reports of the LWPS effect, this interaction reflected smaller RT switch costs when PS was high ( $M_{\text{switchcost}} = 11.5$ ) blocks compared to when PS was low ( $M_{\text{switchcost}} = 27.5$ ). Notably, there was no three-way interaction of trial type  $\times$  PS  $\times$  cue bias ( $F(1,82) = 0.04$ ,  $p = .84$ ,  $\eta^2 < .001$ ), indicating no difference in the LWPS effect between biased and unbiased cues.

### 5.2.3 Discussion

In sum, the results from Experiment 2 replicated the LWPS effect and suggest that flexibility adjustments in the LWPS protocol are not moderated by switch-frequency learning tied to the task cue. Though flexibility adjustments did not transfer to the unbiased task set in Siqi-Liu & Egner (2020), we demonstrate here that they do transfer to unbiased task cues associated with a biased task set (see **Figure 21**). These results suggest that switch frequency learning is specifically tied to task-sets associated with frequent switching or repeating, as speculated in Siqi-Liu & Egner (2020), rather than to the arbitrary physical stimulus features of the task cue.



**Figure 21: Experiment 2 results. Top panels depict group mean RTs (top left) and error rates (top right) in low (30%) versus high (70%) proportion switch (PS) conditions. Bottom panels depict switch costs for RTs (bottom left) and error rates (bottom right), which were calculated as switch - repeat. Individual dots represent individual mean switch costs. The left side of each panel show trials with biased cues while the right side shows trials with unbiased cues. Violin plots are overlaid to visualize the distribution of individual means. In the box plots, central lines depict group medians, box edges show the interquartile range (IQR), and the length of the whiskers mark 1.5 x IQR.**

### 5.3 Experiment 3

#### Experiment 3

The results of Experiment 2 indicate that flexibility learning is specifically tied to task-set representations rather than visual features of the task cues. Together with a lack of transfer to unbiased task sets found in Siqu-Liu & Egner (2020, Exp 3a-b), this provides further evidence for the conclusion that flexibility learning is task-specific in the list-wide PS protocol. However, the lack of transfer to the unbiased task set observed in Siqu-Liu & Egner (2020, Exp 3a-b) may have had to do with the fact that the letter, digit, and color tasks used in that study were each performed on rather distinct and spatially separated stimuli or stimulus features, which may have encouraged task-specific flexibility learning.

There is some prior evidence of control transfer based on shared stimulus features. Bustamante et al. (2021) demonstrated that transfer occurred when novel stimuli consisted of composite features that were previously associated with incentives for control exertion. Specifically, participants initially learned that one set of features in a color-word Stroop task (e.g., the color **blue** and the word RED) predicted that the color naming task would be rewarded, and that another set of features predicted that the word reading task would be rewarded. Then, the experimenters introduced new stimuli (e.g., **RED**) that combined features that were previously associated with reward for the color-naming task but together predicted that the word-reading task would be rewarded. They found that participants overexerted control (chose to complete color naming more) for new stimuli containing features that previously predicted high value

of control. Thus, feature-control state associations transferred to new, integrated stimuli. Similarly, task-set control state associations may be more easily transfer to other tasks when those tasks share integrated stimulus features. Transfer of control states has also been demonstrated across stimuli linked through associative learning (Bejjani et al., 2018) and across linked spatial locations (Weidler & Bugg, 2016).

In the current experiment, we therefore investigated whether task sets would still form the boundary of flexibility learning if the tasks involved categorizing different integrated feature dimensions of the same stimuli. For this purpose, we compared flexibility transfer between two scenarios in a between-groups design. One group performed three face feature categorization tasks, such that transfer was assessed between two biased face categorization tasks to one unbiased face categorization task (high level of stimulus overlap). The other group performed two facial categorization tasks and one object categorization task, and transfer was tested from the two biased face categorization tasks to the object categorization task (i.e., no stimulus overlap).

### **5.3.1 Method**

#### **5.3.1.1 Participants**

Power analysis (G\*Power 3.1; Faul et al., 2009) based on effect sizes from Siqi-Liu & Egner (2020, Exp 3a,  $\eta_p^2 = 0.09$ ) found that a sample size of 82 was necessary to detect a three-way interaction between trial type x PS x task bias at  $p < 0.05$  with 80% power. As such, we aimed to include at least 82 participants in each group. Our final sample for

data analysis consisted of 170 participants from Amazon Mechanical Turk who performed above the accuracy criteria of >65%. The overlap and non-overlap groups each consisted of 85 randomly assigned participants. There were 77 female participants, 81 male participants, and 2 participants who responded with “Other.” Average age of participants was 39.9 with an *SD* of 12.19 years. All participants gave informed consent and received payment or course credit in accordance with a protocol approved by the Duke University Institutional Review Board.

#### **5.3.1.2 Stimuli**

For the overlap group, stimuli consisted of 104 pictures of faces of unique identities (Ebner et al, 2010). Faces differed along three dimensions: age (younger/older than 30), gender (male/female), and emotion (happy/sad).

For the non-overlap group, stimuli consisted of 80 pictures of faces of unique identities (Ebner et al., 2010) and 40 pictures of objects that were either manmade (e.g., house) or natural (e.g., tree) (Wen & Egner, 2022).

#### **5.3.1.3 Procedures**

For each participant, two tasks were designated as biased tasks and one as the unbiased transfer task. In the overlap group, the two biased tasks and transfer task were randomly selected from three face categorization tasks. These were a “gender” task, where participants decided whether the face stimulus was male or female, an “age” task, which involved categorizing the face as young (roughly under 30) or old (roughly over

30), and an “emotion” task, which involved deciding whether the expression on the face shown is happy or sad. On each trial, one face stimulus was randomly selected with replacement from the set of 104 faces.

In the non-overlap group, the biased tasks were gender and age tasks, performed on pictures of faces. On each trial consisting of an age or gender task, a face stimulus was randomly selected with replacement from the set of 80 faces. The transfer task was performed on a separate stimulus set of pictures of objects and consisted of an “origin” task where participants responded whether each object stimulus was manmade or natural. On each trial of the origin task, an object stimulus was randomly selected with replacement from the set of 40 object stimuli.

Within each experimental block, all participants randomly switched between the three tasks based on a preceding experimental cue. Each trial began with a fixation cross, displayed for 300 ms, followed by a cue word (“gender”, “age”, “emotion” or “origin”) that preceded the stimulus for 200 ms and indicated which task the participant had to perform on the upcoming stimulus. The overlap group only received gender, age, and emotion cues while the non-overlap group only received gender, age, and origin cues. The face/object stimulus was then displayed for 1200 ms, during which the task cue stayed on screen superimposed on top of the face/object. During the stimulus display, participant used a left or right arrow key press to indicate whether the stimulus was young/old, happy/sad, young/old, or manmade/natural based on instructions for

response mappings received at the beginning of the task. After stimulus offset, feedback (“correct,” “incorrect,” or “too slow”) was displayed for 300 ms.

All participants completed 1 practice block and 16 experimental blocks of 31 trials each. Stimuli that were used in the practice block did not appear in the experimental blocks. Half of the experimental blocks had a low (30%) PS trial sequence and half had a high (70%) PS rate. Participants were randomly assigned to complete either all the low or all the high PS blocks first. In between blocks, participants could take a short self-paced break, during which a reminder of the response mappings was displayed.

To achieve an overall switch: repeat trial ratio of 9:21 (30%) in the low PS blocks, the two biased tasks were each presented twice as switch trials and eight times as repeat trials, and the transfer task was presented five times as switch and five times as repeat trials. For a 21:9 switch: repeat trial ratio (70%) in the high PS blocks, the number of switch versus repeat trials was reversed for the biased tasks, whereas the unbiased task was still presented five times as switch and five times as repeat trials. Thus, the transfer task appeared equally often as switch and repeat trials in both the high and low PS blocks. These trial sequences were pseudo-randomly generated.

### **5.3.2 Results**

The same trial exclusion and filtering criteria were applied as in experiment 1. We conducted a 2 (group: overlap v. non-overlap)  $\times$  2 (PS: high v. low)  $\times$  2 (trial type:



switch v. repeat) x 2 (task bias: biased v. transfer) mixed ANOVA on response times and accuracy. The data are summarized in **Figure 22**.

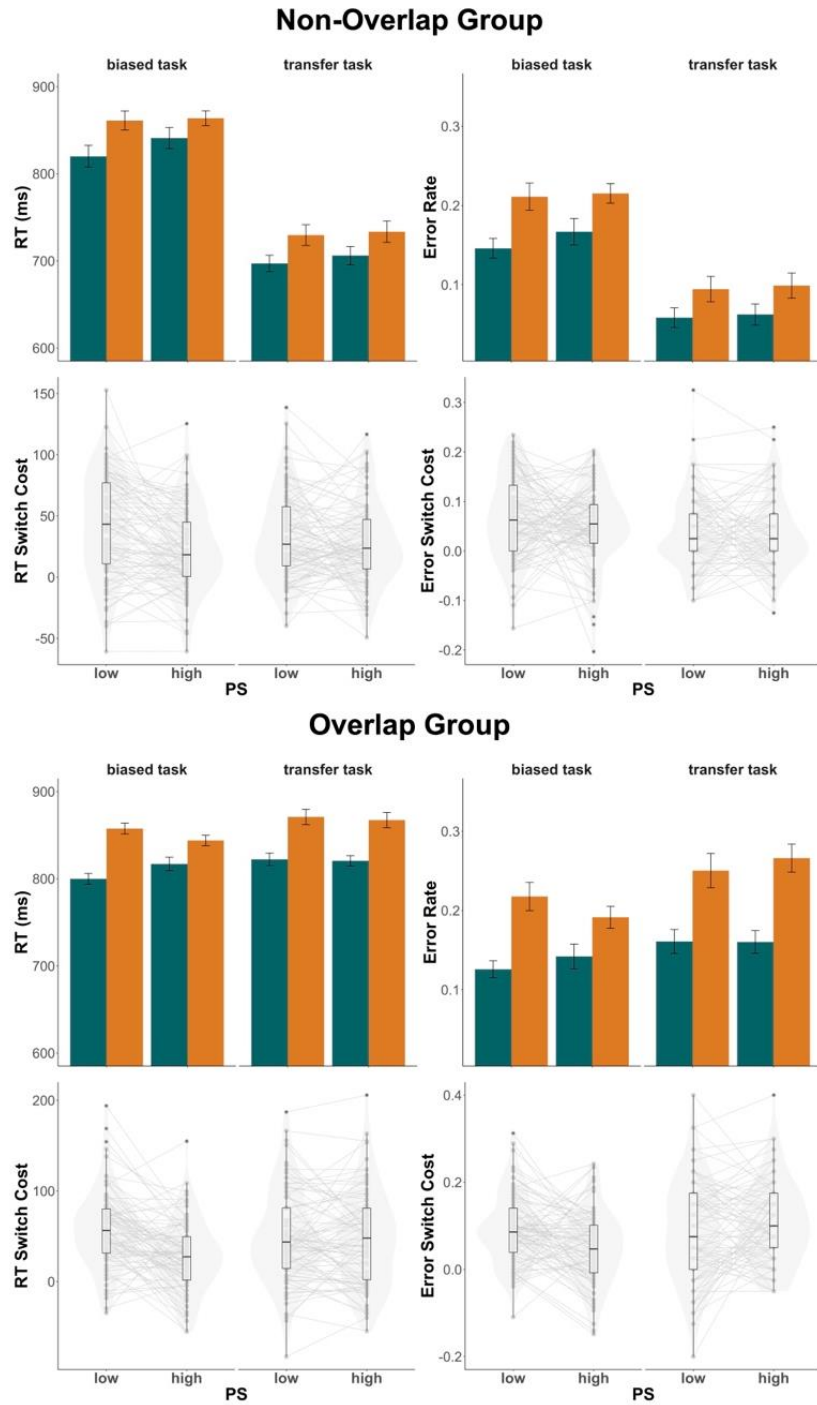
Responses were slower ( $F(1,168) = 9.38, p = .003, \eta^2 = 0.053$ ) and more error-prone ( $F(1,168) = 23.61, p < .001, \eta^2 = .12$ ) in the overlap group ( $M_{rt} = 822.8$  ms;  $M_{error} = 0.18$ ) compared to the non-overlap group ( $M_{rt} = 788.6$  ms;  $M_{error} = 0.15$ ). There were significant switch costs in response times ( $F(1,168) = 258.6, p < .001, \eta^2 = .61$ ) and error rates ( $F(1,168) = 294.2, p < .001, \eta^2 = .64$ ), wherein switch trials were associated with longer response times and higher error rates ( $M_{rt} = 824.9$  ms;  $M_{error} = 0.20$ ) compared to repeat trials ( $M_{rt} = 787.4$  ms;  $M_{error} = 0.13$ ). On average, transfer tasks were associated with faster RTs and lower error rates ( $M_{rt} = 769.6$  ms;  $M_{error} = 0.15$ ) than biased tasks ( $M_{rt} = 823.9$ ;  $M_{error} = 0.17$ ) (RT:  $F(1,168) = 127.2, p < .001, \eta^2 = .43$ ; Err:  $F(1,168) = 24.33, p < .001, \eta^2 = .13$ ). No other main effects reached significance.

We found an interaction effect of group x task bias in both RTs ( $F(1,168) = 206.4, p < .001, \eta^2 = .55$ ) and error rates ( $F(1,168) = 119.6, p < .001, \eta^2 = .42$ ). In the overlap group, the transfer task was associated with slower RTs and higher error rates ( $M_{rt} = 836.9$  ms;  $M_{error} = 0.21$ ) than the biased tasks ( $M_{rt} = 816.2$  ms;  $M_{error} = 0.16$ ). This relationship was reversed in the non-overlap group, where the transfer task was instead associated with faster RTs and lower error rates ( $M_{rt} = 711.9$  ms;  $M_{error} = 0.08$ ) compared to the biased tasks ( $M_{rt} = 831.8$  ms;  $M_{error} = 0.18$ ). This interaction effects may reflect differences in task difficulty: the origin task (which was the transfer task for the non-

overlap group) appeared to be somewhat easier than the age/gender/emotion tasks (transfer tasks for the overlap group). The face tasks may be more difficult in part because of increased cross-task interference due to stimulus overlap between the different face tasks.

Switch costs were significantly higher in the overlap group ( $M_{rt} = 45.0$  ms;  $M_{error} = 0.08$ ) compared to the non-overlap group ( $M_{rt} = 30.9$  ms;  $M_{error} = 0.04$ ) in both RTs ( $F(1,168) = 8.9$ ,  $p = .003$ ,  $pes = .05$ ) and error rates ( $F(1,168) = 24.4$ ,  $p < .001$ ,  $pes = .13$ ). We also replicated the LWPS effect, finding a significant interaction effect of trial type  $\times$  PS in RTs ( $F(1,168) = 34.0$ ,  $p < .001$ ,  $pes = .17$ ) but not error rates. This was driven by smaller switch costs in the high PS condition ( $M_{rt} = 30.9$  ms) compared to the low PS condition ( $M_{rt} = 45.0$  ms).

Crucially, there was a three-way interaction of trial type  $\times$  proportion switch  $\times$  task bias in RTs ( $F(1,168) = 441.5$ ,  $p < .001$ ,  $pes = .12$ ) and error rates ( $F(1,168) = 10.5$ ,  $p = .001$ ,  $pes = .06$ ), indicating that the LWPS effect was modulated by task bias (see **Figure 22**). Furthermore, there was no four-way interaction with group in RTs or error rates, suggesting that the LWPS effect was modulated by task bias regardless of whether the transfer task contained stimulus overlap with the biased tasks.



**Figure 22: Experiment 3 results. Top panels of each group condition depict group mean RTs (top left) and error rates (top right) in low (30%) versus high (70%) proportion switch (PS) conditions. Bottom panels depict switch costs for RTs (bottom**

left) and error rates (bottom right), which were calculated as switch - repeat. Individual dots represent individual mean switch costs. The left side of each panel show trials with the biased tasks while the right-side shows trials with the transfer task. Violin plots are overlaid to visualize the distribution of individual means. In the box plots, central lines depict group medians, box edges show the interquartile range (IQR), and the length of the whiskers mark 1.5 x IQR.

To further investigate how task bias affected LWPS effects in the two groups, we ran separate 2 (trial type) x 2 (PS) ANOVAs for each condition resulting from crossing task bias x group.

In the overlap group, the trial type x PS interaction was significant for the biased tasks in RTs ( $F(1,84) = 33.4, p < .001, \eta^2 = .28$ ) and error rates ( $F(1,84) = 13.4, p < .001, \eta^2 = .14$ ) but was nonsignificant for the transfer task in both RTs ( $F(1,84) = .13, p = .72, \eta^2 = .001$ ) and error rates ( $F(1,84) = 1.7, p = .20, \eta^2 = .02$ ).

Similarly, in the non-overlap group, the trial type x PS interaction was significant for biased tasks in RTs ( $F(1,84) = 21.0, p < .001, \eta^2 = .20$ ), while the interaction representing the LWPS effect was nonsignificant for the transfer task ( $F(1,84) = 1.5, p = .22, \eta^2 = .02$ ). The LWPS effect did not reach significance in any of the error rate ANOVAs.

### 5.3.3 Discussion

The results of Experiment 3 replicated the LWPS effect of reduced switch costs when participants switched more frequently; however, as in Siqu-Liu and Egner (2020), flexibility learning did not transfer to unbiased tasks presented within the same blocks, even when the transfer task shared overlapping stimulus features with the biased tasks.

This indicates that task sets form the boundaries of switch-readiness learning in the LWPS paradigm, regardless of whether they are applied to distinct or identical stimuli.

## ***General Discussion***

Across three cued task-switching experiments that utilized LWPS manipulations, the current study replicated findings from Siqi-Liu & Egnér (2020) and provided additional evidence that flexibility learning generalizes over novel stimuli and physical task-cues. Furthermore, flexibility learning seems to be tied to abstract task-sets rather than physical task cues, and this robust task-specific learning occurred regardless of whether there is stimulus overlap between tasks.

Experiment 1 replicated prior findings of list-wide flexibility adjustments (Bonnin et al., 2011; Dreisbach & Haider, 2006; Dreisbach et al., 2002; Duthoo et al., 2012; Mayr, 2006; Monsell & Mizon, 2006; Siqi-Liu & Egnér, 2020) but using trial-unique stimuli, thus providing the most stringent evidence that flexibility adaptations do not rely on learning of any kind of stimulus-level bias. To our knowledge, this is the first time that the LWPS effect has been found with completely trial-unique stimuli.

Experiment 2 probed whether flexibility adaptations depended on control associations with abstract task-sets (i.e., the rules that define stimulus-response associations) as suggested by Siqi-Liu & Egnér (2020, Exp 3), rather than associations with the arbitrary physical cue stimuli, or the visual features that indicated to participants which task they should perform – in a similar manner that task-irrelevant contextual features have been

shown to trigger conflict adaptation effects (Bejjani et al, 2018). We observed flexibility adaptations to PS context in trials using biased task-cues, which occurred more often on switch/repeat trials in the high/low PS blocks, but also in trials using unbiased transfer cues, which appeared equally often on switch and repeat trials, regardless of the PS context.

In conjunction, Experiment 1 and 2 suggest that flexibility adaptations in cued task switching do not depend on control associations with specific physical features of the trial, whether these be features of the task cue or the target stimulus. Rather, flexibility seems to be specifically tied to the categorization rules that are applied to task stimuli. These results support the conclusions from Siqi-Liu and Egnér (2020) that flexibility learning in the LWPS protocol is task-specific, and that task-, but not stimulus-level associations are necessary for flexibility adaptations.

Experiment 3 further tested whether task-specific flexibility learning would still occur even when tasks are performed on the same stimuli, with integrated, overlapping features. Across two groups of participants, we compared cross-task transfer of the LWPS effect from two biased face tasks to a third unbiased transfer task, in one group utilizing the same face stimuli, and in the other group using an unbiased object categorization task with a completely non-overlapping stimulus set. In both groups, we found LWPS effects in the two biased tasks, but not in the unbiased transfer task, suggesting that stimulus-feature overlap, or the lack thereof, did not have any effect on

task-specific flexibility learning. It is important to note that the lack of a statistically significant LWPS effect in the transfer tasks does not necessarily mean that context-specific switch cost adaptations did not occur, since nonsignificant findings can also be explained by a lack of power (despite our sample size calculations) or chance. However, Experiment 3 results generally suggest that task-specific flexibility learning is highly robust and is not affected by stimulus-set overlap in between task-sets.

Siqi-Liu & Egnér (2020) speculated that task-specific switch readiness adjustments may occur via rapid bottom-up priming of context-appropriate control settings (King et al., 2012), which aid preparatory task-set reconfiguration processes for those frequently switched-to tasks. Evidence that task-sets may specifically define the boundaries of control strategies have also been identified in Grant et al. (2020), which found that congruency sequence effects (CSE) were eliminated when sensory modalities (for stimuli and distractors) changed, but only when different sensory modalities defined different task sets. That is, when the distractor modality predicted the target modality, participants used the distractor information to orient themselves to the modality in which the target stimuli would appear, facilitating the formation of modality-specific task sets. These modality-specific task-sets acted as boundaries for the CSE – when task-sets (rather than simply target or distractor modality) switched, participants abandoned control expectations formed in the previous trial, thus eliminating the CSE. Similarly, in task-switching paradigms, task-sets may help

participants orient to the relevant stimulus dimension for producing a correct response, providing a strong “context” to which flexibility learning could be bound.

In conclusion, the current study the findings of task-specific flexibility learning from Siqu-Liu & Egner (2020) under novel circumstances, providing additional evidence that task sets form robust boundaries for flexibility learning in cued task-switching paradigms. Further research is required to determine the conditions of cross-task flexibility transfer in other paradigms, such as ones that involve un-cued task switching or hierarchical task-set organization.



## 6. Conclusions

Across behavioral and EEG studies, we show that flexibility adaptations do not occur at a tonic, block-wide level (Chapter 2 & 3), nor are they based in switch associations with stimuli, task-cues, or feature overlap between stimulus sets (Chapter 2, Chapter 5). Rather, task-sets establish robust boundaries for flexibility-learning, and task-level switch-associations seem to be a necessary condition for switch cost adaptations in cued-task switching paradigms. Additionally, these learned switch-associations with tasks are rapidly acquired, robust across sessions, and do not seem to benefit from memory consolidation (Chapter 3).

### ***6.1 Potential mechanisms for task-specific switch readiness learning***

We speculate that task-specific switch readiness learning could be accomplished via two mechanisms: 1) biased task-cues priming faster task-set reconfiguration and 2) weakened task-set representations of frequently switched to tasks, such that these tasks suffer less from associative interference. Thus, meta-flexibility is facilitated by top-down control processes (faster task-set reconfiguration) working in tandem with bottom-up priming (from task cues and stimulus associations).

Much previous work in control learning has often focused on applying experimental controls to eliminate effects of feature integration and isolate “pure” top-down control mechanisms (see Schmidt & Liefoghe, 2016). An implicit assumption of these approaches is that qualitatively different processes underly the acquisition of

stimulus-response associations and control recruitment, and that top-down and bottom-up processes have orthogonal effects on performance. However, task-specific flexibility learning demonstrates one case where, rather than being simply additive, bottom-up feature-based learning seem to adaptively prime faster top-down control recruitment, such as task-set reconfiguration, on a trial-by-trial basis. This interaction between anticipatory control and feature-based learning mechanisms may be crucial to understanding adaptive behavior.

Newer event-file approaches to control learning (Frings et al., 2020) are particularly helpful for rethinking the relationship between top-down and bottom-up processes. Accumulating evidence suggests that control learning is also integrated with stimulus-response associations and is based in the same associative learning mechanisms (Abrahamse et al., 2016; Braem et al., 2011; Braem & Egner, 2018; Tobias Egner, 2014). Rather than being generalizable, cognitive flexibility learning is specific to task-sets and stimuli (Chiu & Egner, 2017; Leboe et al., 2008; Siqi-Liu & Egner, 2020), much like the learning of concrete stimulus-response associations. Braem & Egner (2018) propose a “learning perspective” on cognitive flexibility, which conceptualizes cognitive demands (such as proportion switch context) as an abstract stimulus (or event) feature that is bound to co-occurring temporal contexts, task-sets, and other stimulus-response features, much like how traditional feature-integration accounts construe concrete co-

occurring stimulus-response features (e.g., “A,” left-button press, etc.) as being bound together.

This approach combines the advantages of both memory- and control- based accounts: it allows for higher order control functions to play a role in task-performance while also providing a specific mechanism of control-learning to occur (i.e., via feature-integration).

## ***6.2 Task-sets as an event-file boundary for flexibility learning in cued-switch paradigms***

Task-sets may be unique because they assert influence on response strategies in a deeper way than providing an episodic context. Studies of response-repetition effects support the idea that task-sets provide hierarchical organization for stimulus-response associations. For example, Koch et al. (2018) found that varying the episodic overlap between task cues (via a sensory modality change) reduced response-repetition benefits in repeat trials but did not reduce response-repetition costs in switch trials, demonstrating that a task-switch is more than just a switch in episodic context.

Why should task-sets rather than blocks of trials/temporal episodes be the feature which defines the boundaries for switch-readiness learning? Ideomotor theory suggests that in order to perform an action, the agent must first anticipate the perceptual effects that the action would produce; only then can they retrieve and execute the motor program necessary to produce the desired effect (see Shin et al., 2010 for a modern review), thus action is thought to be central to the formation of bound representations.

Task-sets define how percepts should be transformed to actions, by specifying relevant stimulus dimensions and response rules, and what results those actions would produce (correct or incorrect responses). Given the importance of action representations to event files, it may be intuitive that task-sets, or the rule defining action mappings should serve as boundaries for the binding of concrete event characteristics and internal control settings.

In contrast, a block-wide temporal context is often irrelevant to task instructions and does not specify different relevant stimulus or response features. Since response rules do not change depending on block in typical task-switching paradigms, the temporal context of specific blocks never informs participants how they should act. This abstract temporal context may not be incorporated into an (extended) event file along with task-set and stimulus representations, because it is not subjectively perceived as salient, or made implicitly salient through action. Even when participants are given explicit knowledge about list-wide context, this knowledge is not necessarily helpful for switch-readiness adaptations. Employing explicit cues informing participants of the list-wide PS conditions at the beginning of each block, Dreisbach & Haider (2006) found comparable LWPS effects as other similar paradigms. The LWPS effect only became larger via slower repeat trial responses when the authors made the list-wide PS manipulation salient through repeated cuing at the beginning of each trial. This

demonstrates that list-wide context does not usually serve as a useful cue for adjustments in control-strategy.

In sum, though few, existing studies suggest that the temporal context constituted by a block of trials may be a poor boundary for control binding, because block context is both abstract (not associated with concrete stimulus features) and task irrelevant. Rather than blocks of trials (temporally extended episodes), task-sets may be prioritized as an anchoring point for control strategies because they determine processing strategies, including attentional sets and response modes (Egner, 2014; Grant et al., 2020; Schumacher & Hazeltine, 2016).

### ***6.3 Limitations and Future Directions***

An important limitation of the current set of studies is that we only examined flexibility learning in the context of cued task-switching and cannot rule out the possibility that cross-task flexibility transfer could occur under other experimental conditions. Specifically, cross-task flexibility transfer was recently demonstrated in a probabilistic version of the Wisconsin Card Sorting Task (Wen et al., 2023). There, participants completed the card sorting task in either high (frequent uninstructed rule switches) or low volatility (infrequent rule switches) environments before transitioning to a medium-volatility transfer phase. Using reinforcement learning modeling, the authors found that participants exhibited higher learning rates, or faster adaptation to rule changes, in transfer phases that followed exposure to the high-volatility

environment. These learning rate adaptations occurred even when novel tasks and stimuli were used in the transfer phase. Wen et al.'s (2023) paradigm may be particularly effective at inducing general changes in flexibility because participants had to voluntarily switch between categorization rules to discover rule switches based on feedback amidst uncertainty. Optimal performance may thus depend on changes in general exploratory behavior because of the lack of explicit task cuing.

In contrast, cued or forced task-switching does not involve any uncertainty regarding the correct rule to implement on each trial, and thus may encourage participants to prioritize maintaining only frequently switched-to task-sets in working memory. In particular, Dreisbach and Fröber (2019) argue that some cases of flexibility modulations may depend on concurrent task activation, rather than the adjustment of a general 'updating threshold.' Lowering the updating threshold may increase flexibility in general; on the other hand, adaptations based on concurrent task activation may manifest as increased flexibility between only specific tasks held in working memory. While paradigms without explicit task-cuing such as Wen et al (2023) may encourage adjustments to the general updating threshold, cued-switching paradigms may encourage task-specific flexibility adaptations via concurrent task activation.

Investigating different paradigm designs with instructional or contextual factors that encourage or discourage task-specific learning is one potentially fruitful direction for future studies of cross task flexibility transfer. For example, the boundaries of

flexibility learning may be blurred in paradigms that involve uncertainty in task-rules, or that encourage voluntary switches via feedback rather than forcing switches with cues, such as in Wen et al. (2023).

## Appendix A

Table 6: Excluded Trial Counts and Final Number Remaining in Smallest and Largest Cells

	Total Count	# of Error or Post-Err Trials	# of RT Outliers	Final Count	Smallest Cell	Largest Cell
<b>Exp 1 (N = 40)</b>						
<b>2 (task sequence) x 3 (switch frequency) x 2 (CSI)</b>						
Overall <sup>6</sup>	21600	3102	132	18366	876	2281
Per Subject	540	77.55	3.3	459.15	21.9	57.025
% per subj		0.14	0.006	0.85		
<b>Exp 2 (N = 80)</b>						
<b>2 (task sequence) x 2 (switch frequency)</b>						
Overall	9840	1750	68	8022	1140	2950
Per subj	123	21.88	0.85	100.28	14.25	36.88
% per subj		0.18	0.007	0.82		
<b>Exp 3a (N = 41)</b>						
<b>2 (task sequence) x 2 (switch frequency) x 2 (task bias)</b>						
Overall	19680	3242	173	16265	992	4483
Per subj	480	79.07	4.22	396.71	24.20	109.34
% per subj		0.16	0.008	0.83		
<b>Exp 3b (N = 61)</b>						
<b>2 (task sequence) x 2 (switch frequency) x 2 (task bias)</b>						
Overall	41724	7484	261	33979	1664	9961
Num per subj	684	122.69	4.28	557.03	27.28	163.30
% per subj		0.18	0.006	0.81		
<b>Exp 4 (N = 82)</b>						
<b>2 (task sequence) x 2 (switch frequency) x 2 (stimulus bias)</b>						
Overall	39360	7020	165	32175	2029	8859
Num per subj	480	85.61	2.01	392.38	24.74	108.04
% per subj		0.18	0.004	0.82		

<sup>6</sup> Overall refers to the total number of trials in the full dataset. Average per subject trial counts is calculated by dividing overall trial count by number of subjects in data set. Average percentage of trials per subject excluded due to error or RT outliers refers to the total count divided by the average per subject total.

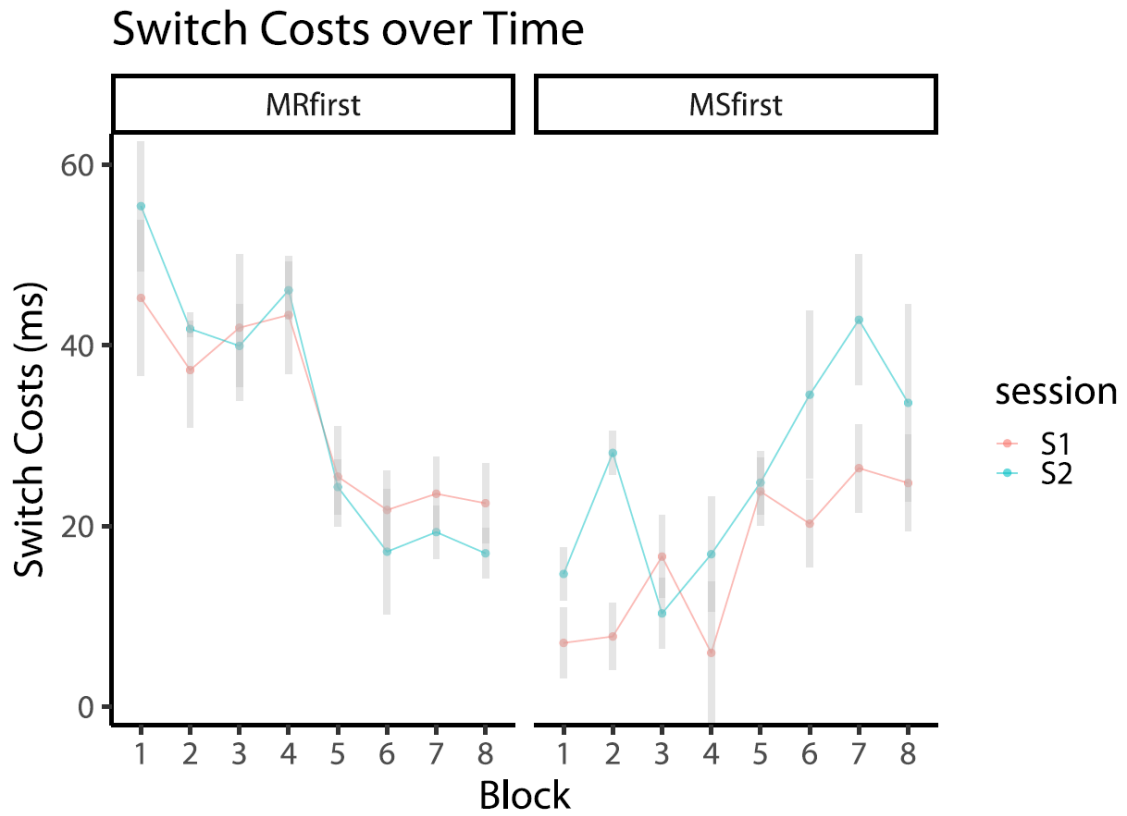


Table 7: Frequencies of Task Presentations as Switch and Repeat *per Different Switch Frequency Blocks*

	30% Switch			70% Switch			50% Switch		
	<i>Sw</i>	<i>Rep</i>	<i>Total</i>	<i>Sw</i>	<i>Rep</i>	<i>Total</i>	<i>Sw</i>	<i>Rep</i>	<i>Total</i>
<b>Exp 1<sup>7</sup></b>									
<i>Task A</i>	4	11	15	11	4	15	7	8	15
<i>Task B</i>	5	10	15	10	5	15	8	7	15
<i>Total</i>	9	21	30	21	9	30	15	15	30
<b>Exp 2</b>									
<i>Task A</i>	3	7	10	7	3	10			
<i>Task B</i>	3	7	10	7	3	10			
<i>Task C</i>	3	7	10	7	3	10			
<i>Total</i>	9	21	30	21	9	30			
<b>Exp 3a</b>									
<i>Biased 1</i>	2	8	10	8	2	10			
<i>Biased 2</i>	2	8	10	8	2	10			
<i>Transfer</i>	5	5	10	5	5	10			
<i>Total</i>	9	21	30	21	9	30			
<b>Exp 3b</b>									
<i>Biased 1</i>	5	10	15	10	5	15			
<i>Biased 2</i>	6	9	15	9	6	15			
<i>Transfer</i>	4	4	8	4	4	8			
<i>Total</i>	15	23	38	23	15	38			
<b>Exp 4</b>									
<i>Task A</i>	9	21	30	21	9	30			
<i>Task B</i>	9	21	30	21	9	30			
<i>Total</i>	18	42	60	42	18	60			

<sup>7</sup> The identity of each task (digit, number, color tasks) is randomized across participants to be task A/B/C or biased task 1/biased task 2/transfer task. Numbers do not include the random starting task (first trial in each block) or the practice block, which were removed from analyses. There were 18 blocks in E1, 8 blocks in E2, 16 blocks in E3a, 18 blocks in E3b, and 8 blocks in E4.

## Appendix B



**Figure 23: Across experiment switch cost results. Mean switch costs ( $\pm$  SEM) are shown as a function of Block (1-8), Block Order (Mostly Repeat or Mostly Switch context experienced first), and session (1/2). For inferential statistics, see Appendix B, Table 8.**

**Table 8: Across Experiments ANOVA with Block Number (1-8) instead of PS factor for Reaction Time**

<b>Effect</b>	<b>Df1</b>	<b>Df2</b>	<b>F</b>	<b><math>\eta_p^2</math></b>	<b>p value</b>	<b>BF</b>
Experiment	1	147	0.24	0.002	0.627	0.22
<b>Trial Type</b>	<b>1</b>	<b>147</b>	<b>182.87</b>	<b>0.55</b>	<b>&lt;0.001</b>	<b>4.04 x 10<sup>109</sup></b>
Experiment x Trial Type	1	147	0.84	0.006	0.360	0.12
<b>Stim Bias</b>	<b>1</b>	<b>147</b>	<b>65.95</b>	<b>0.31</b>	<b>&lt;0.001</b>	<b>3.62 x 10<sup>13</sup></b>
Experiment x Stim Bias	1	147	2.63	0.018	0.107	0.14
<b>Block</b>	<b>5.23</b>	<b>768.15</b>	<b>106.12</b>	<b>0.42</b>	<b>&lt;0.001</b>	<b>2.08 x 10<sup>214</sup></b>
<b>Experiment x Block</b>	<b>5.23</b>	<b>768.15</b>	<b>3.15</b>	<b>0.021</b>	<b>0.007</b>	<b>3.46</b>
<b>Session</b>	<b>1</b>	<b>147</b>	<b>11.34</b>	<b>0.072</b>	<b>0.001</b>	<b>2.74 x 10<sup>26</sup></b>
Experiment x Session	1	147	0.18	0.001	0.671	0.10
Trial Type x Stim Bias	1	147	0.27	0.002	0.605	0.04
<b>Experiment x Trial Type x Stim Bias</b>	<b>1</b>	<b>147</b>	<b>5.54</b>	<b>0.036</b>	<b>0.019</b>	<b>0.18</b>
Trial Type x Block	6.04	888.10	0.78	0.005	0.588	<0.001
Experiment x Trial Type x Block	6.04	888.10	1.28	0.009	0.265	<0.001
Stim Bias x Block	6.29	925.12	1.46	0.010	0.187	<0.001
<b>Experiment x Stim Bias x Block</b>	<b>6.29</b>	<b>925.12</b>	<b>1.80</b>	<b>0.012</b>	<b>0.093</b>	<b>&lt;0.001</b>
<b>Trial Type x Session</b>	<b>1</b>	<b>147</b>	<b>4.17</b>	<b>0.028</b>	<b>0.043</b>	<b>0.30</b>
Experiment x Trial Type x Session	1	147	0.01	<0.001	0.915	0.05
<b>Stim Bias x Session</b>	<b>1</b>	<b>147</b>	<b>2.91</b>	<b>0.019</b>	<b>0.090</b>	<b>0.13</b>
Experiment x Stim Bias x Session	1	147	0.48	0.003	0.491	0.06
<b>Block x Session</b>	<b>5.63</b>	<b>828.29</b>	<b>6.98</b>	<b>0.045</b>	<b>&lt;0.001</b>	<b>1.16 x 10<sup>8</sup></b>
Experiment x Block x Session	5.63	828.29	1.46	0.010	0.192	0.001
Trial Type x Stim Bias x Block	6.56	964.55	0.93	0.006	0.478	<0.001
Experiment x Trial Type x Stim Bias x Block	6.56	964.55	0.97	0.007	0.452	<0.001
Trial Type x Stim Bias x Session	1	147	0.11	0.008	0.736	0.05
Experiment x Trial Type x Stim Bias x Session	1	147	1.92	0.013	0.168	0.10
Trial Type x Block x Session	6.48	952.50	0.79	0.005	0.584	<0.001
Experiment x Trial Type x Block x Session	6.48	952.50	1.60	0.011	0.137	0.002
<b>Stim Bias x Block x Session</b>	<b>6.34</b>	<b>931.36</b>	<b>1.97</b>	<b>0.013</b>	<b>0.064</b>	<b>&lt;0.001</b>

Experiment x Stim Bias x Block x Session	6.34	931.36	0.81	0.005	0.567	<0.001
<b>Trial Type x Stim Bias x Block x Session</b>	<b>6.60</b>	<b>970.75</b>	<b>1.84</b>	<b>0.012</b>	<b>0.080</b>	<b>0.002</b>
Experiment x Trial Type x Stim Bias x Block x Session	6.60	970.75	0.95	0.006	0.464	0.001

## References

- Abrahamse, E., Braem, S., Notebaert, W., & Verguts, T. (2016). Grounding cognitive control in associative learning. *Psychological Bulletin*, *142*(7), 693–728. <https://doi.org/10.1037/bul0000047>
- Akçay, Ç., & Hazeltine, E. (2007). Conflict monitoring and feature overlap: Two sources of sequential modulations. *Psychonomic Bulletin & Review*, *14*(4), 742–748. <https://doi.org/10.3758/BF03196831>
- Akçay, Ç., & Hazeltine, E. (2011). Domain-specific conflict adaptation without feature repetitions. *Psychonomic Bulletin & Review*, *18*(3), 505–511. <https://doi.org/10.3758/s13423-011-0084-y>
- Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., van Langen, J., & Kievit, R. A. (2021). Raincloud plots: A multi-platform tool for robust data visualization. *Wellcome Open Research*, *4*, 63. <https://doi.org/10.12688/wellcomeopenres.15191.2>
- Allport, A., & Wylie, G. (1999). Task-switching: Positive and negative priming of task-set. In *Attention, space, and action: Studies in cognitive neuroscience* (pp. 273–296). Oxford University Press.
- Allport, A., & Wylie, G. (2000). *Control of Cognitive Processes: Attention and Performance XVIII*. MIT Press.
- Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. *Attention and Performance 15: Conscious and Nonconscious Information Processing. Attention and Performance Series*, 421–452.
- Antony, J. W., Ferreira, C. S., Norman, K. A., & Wimber, M. (2017). Retrieval as a Fast Route to Memory Consolidation. *Trends in Cognitive Sciences*, *21*(8), 573–576. <https://doi.org/10.1016/j.tics.2017.05.001>
- Arbuthnott, K. (2008). *Asymmetric switch cost and backward inhibition: Carryover activation and inhibition in switching between tasks of unequal difficulty*. [/doiLanding?doi=10.1037%2F1196-1961.62.2.91](https://doi.org/10.1037/2F1196-1961.62.2.91)
- Arechar, A. A., & Rand, D. G. (2020). *Turking in the time of COVID* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/vktqu>
- Arechar, A. A., Kraft-Todd, G. T., & Rand, D. G. (2017). Turking overtime: How participant characteristics and behavior vary over time and day on Amazon

- Mechanical Turk. *Journal of the Economic Science Association*, 3(1), 1–11.  
<https://doi.org/10.1007/s40881-017-0035-0>
- Arrington, C. M., & Logan, G. D. (2004). Episodic and semantic components of the compound-stimulus strategy in the explicit task-cuing procedure. *Memory & Cognition*, 32(6), 965–978. <https://doi.org/10.3758/BF03196874>
- Ashton, J. E., Jefferies, E., & Gaskell, M. G. (2018). A role for consolidation in cross-modal category learning. *Neuropsychologia*, 108, 50–60.  
<https://doi.org/10.1016/j.neuropsychologia.2017.11.010>
- Barceló, F., & Cooper, P. S. (2018). An information theory account of late frontoparietal ERP positivities in cognitive control. *Psychophysiology*, 55(3), e12814.  
<https://doi.org/10.1111/psyp.12814>
- Bejjani, C., & Egner, T. (2021). Evaluating the learning of stimulus-control associations through incidental memory of reinforcement events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 47(10), 1599.
- Bejjani, C., Dolgin, J., Zhang, Z., & Egner, T. (2020). Disentangling the Roles of Cue Visibility and Knowledge in Adjusting Cognitive Control: A Preregistered Direct Replication of the Farooqui and Manly (2015) Study. *Psychological Science*, 31(4), 468–479. <https://doi.org/10.1177/0956797620904045>
- Bejjani, C., Siqi-Liu, A., & Egner, T. (2021). Minimal impact of consolidation on learned switch-readiness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- Bejjani, C., Tan, S., & Egner, T. (2020). Performance Feedback Promotes Proactive but Not Reactive Adaptation of Conflict-Control. *Journal of Experimental Psychology: Human Perception and Performance*, 46(4), 369–387.  
<https://doi.org/10.1037/xhp0000720>
- Bejjani, C., Zhang, Z., & Egner, T. (2018). Control by association: Transfer of implicitly primed attentional states across linked stimuli. *Psychonomic Bulletin & Review*, 25(2), 617–626. <https://doi.org/10.3758/s13423-018-1445-6>
- Biderman, N., Bakkour, A., & Shohamy, D. (2020). What Are Memories For? The Hippocampus Bridges Past Experience with Future Decisions. *Trends in Cognitive Sciences*, 24(7), 542–556. <https://doi.org/10.1016/j.tics.2020.04.004>

- Blais, C., Harris, M. B., Guerrero, J. V., & Bunge, S. A. (2012). Rethinking the role of automaticity in cognitive control. *The Quarterly Journal of Experimental Psychology*, 65(2), 268–276. <https://doi.org/10.1080/17470211003775234>
- Bonnin, C. A., Gaonac'h, D., & Bouquet, C. A. (2011). Adjustments of task-set control processes: Effect of task switch frequency on task-mixing and task-switching costs. *Journal of Cognitive Psychology*, 23(8), 985-997. doi:10.1080/20445911.2011.594435
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108(3), 624-652. doi:10.1037/0033-295X.108.3.624
- Botvinick, M. M., Carter, C. S., Braver, T. S., Barch, D. M., & Cohen, J. D. (2001). *Conflict Monitoring and Cognitive Control*. 29.
- Boynton, M. H., & Richman, L. S. (2014). An online daily diary study of alcohol use using Amazon's Mechanical Turk. *Drug and Alcohol Review*, 33(4), 456–461. <https://doi.org/10.1111/dar.12163>
- Braem, S., & Egner, T. (2018). Getting a Grip on Cognitive Flexibility. *Current Directions in Psychological Science*, 27(6), 470-476. doi:10.1177/0963721418787475
- Braem, S., Bugg, J. M., Schmidt, J. R., Crump, M. J. C., Weissman, D. H., Notebaert, W., & Egner, T. (2019). Measuring adaptive control in conflict tasks. *Trends in Cognitive Sciences*, 23(9), 769–783. <https://doi.org/10.1016/j.tics.2019.07.002>
- Braem, S., Verguts, T., & Notebaert, W. (2011). Conflict adaptation by means of associative learning. *Journal of Experimental Psychology: Human Perception and Performance*, 37(5), 1662–1666. <https://doi.org/10.1037/a0024385>
- Braun, E. K., Wimmer, G. E., & Shohamy, D. (2018). Retroactive and graded prioritization of memory by reward. *Nature Communications*, 9(1), 4886. <https://doi.org/10.1038/s41467-018-07280-0>
- Braver, T. S., & Cohen, J. D. (2000). On the control of control: The role of dopamine in regulating prefrontal function and working. *MIT Press. Making Working Memory Work*, 551–581.
- Brosowsky, N. P., & Crump, M. J. C. (2018). Memory-guided selective attention: Single experiences with conflict have long-lasting effects on cognitive control. *Journal of*

*Experimental Psychology: General*, 147(8), 1134–1153.  
<https://doi.org/10.1037/xge0000431>

- Bugg, J. M. (2014). Conflict-triggered top-down control: Default mode, last resort, or no such thing? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(2), 567–587. doi:10.1037/a0035032
- Bugg, J. M., & Chanani, S. (2011). List-wide control is not entirely elusive: Evidence from picture-word Stroop. *Psychonomic Bulletin & Review; New York*, 18(5), 930–936.
- Bugg, J. M., & Crump, M. J. C. (2012). In Support of a Distinction between Voluntary and Stimulus-Driven Control: A Review of the Literature on Proportion Congruent Effects. *Frontiers in Psychology*, 3. doi:10.3389/fpsyg.2012.00367
- Bugg, J. M., & Hutchison, K. A. (2013). Converging evidence for control of color–word Stroop interference at the item level. *Journal of Experimental Psychology: Human Perception and Performance*, 39(2), 433–449. doi: 10.1037/a0029145
- Bustamante, L., Lieder, F., Musslick, S., Shenhav, A., & Cohen, J. (2021). Learning to Overexert Cognitive Control in a Stroop Task. *Cogn Affect Behav Neurosci*, 21(3), 453–471. <https://doi.org/10.3758/s13415-020-00845-x>
- Cepeda, N. J., Cepeda, M. L., & Kramer, A. F. (2000). Task Switching and Attention Deficit Hyperactivity Disorder. *Journal of Abnormal Child Psychology*, 14.
- Chandler, J., Mueller, P., & Paolacci, G. (2014). Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers | SpringerLink. *Behavioral Research Methods*, 46(1), 112–130.  
<https://link.springer.com/article/10.3758/s13428-013-0365-7>
- Chiu, Y. C., & Egner, T. (2017). Cueing cognitive flexibility: Item-specific learning of switch readiness. *Journal of Experimental Psychology. Human Perception and Performance*, 43(12), 1950–1960. <https://doi.org/10.1037/xhp0000420>
- Chiu, Y. C., & Egner, T. (2019). Cortical and subcortical contributions to context-control learning. *Neuroscience & Biobehavioral Reviews*, 99, 33–41.  
<https://doi.org/10.1016/j.neubiorev.2019.01.019>
- Chiu, Y.-C., & Egner, T. (2019). Cortical and subcortical contributions to context-control learning. *Neuroscience & Biobehavioral Reviews*, 99, 33–41.  
<https://doi.org/10.1016/j.neubiorev.2019.01.019>



- Chiu, Y.-C., Jiang, J., & Egner, T. (2017). The Caudate Nucleus Mediates Learning of Stimulus–Control State Associations. *The Journal of Neuroscience*, *37*(4), 1028–1038. doi:10.1523/jneurosci.0778-16.2016
- Cools, R., & D'Esposito, M. (2011). Inverted-U–Shaped Dopamine Actions on Human Working Memory and Cognitive Control. *Biological Psychiatry*, *69*(12), e113–e125. <https://doi.org/10.1016/j.biopsych.2011.03.028>
- Cools, R., Sheridan, M., Jacobs, E., & D'Esposito, M. (2007). Impulsive Personality Predicts Dopamine-Dependent Changes in Frontostriatal Activity during Component Processes of Working Memory. *The Journal of Neuroscience*, *27*(20), 5506–5514. doi:10.1523/jneurosci.0601-07.2007
- Cooper, P. S., Karayanidis, F., McKewen, M., McLellan-Hall, S., Wong, A. S. W., Skippen, P., & Cavanagh, J. F. (2019). Frontal theta predicts specific cognitive control-induced behavioural changes beyond general reaction time slowing. *NeuroImage*, *189*, 130–140. doi:10.1016/j.neuroimage.2019.01.022
- Cooper, P. S., Wong, A. S. W., Fulham, W. R., Thienel, R., Mansfield, E., Michie, P. T., & Karayanidis, F. (2015). Theta frontoparietal connectivity associated with proactive and reactive cognitive control processes. *NeuroImage*, *108*, 354–363. <https://doi.org/10.1016/j.neuroimage.2014.12.028>
- Cooper, P. S., Wong, A. S. W., McKewen, M., Michie, P. T., & Karayanidis, F. (2017). Frontoparietal theta oscillations during proactive control are associated with goal-updating and reduced behavioral variability. *Biological Psychology*, *129*, 253–264. <https://doi.org/10.1016/j.biopsycho.2017.09.008>
- Crump, M. J. C., & Logan, G. D. (2010). Contextual control over task-set retrieval. *Attention, Perception, & Psychophysics*, *72*(8), 2047–2053. doi: 10.3758/BF03196681
- Crump, M. J. C., & Logan, G. D. (2010). Hierarchical control and skilled typing: Evidence for word-level control over the execution of individual keystrokes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(6), 1369–1380. <https://doi.org/10.1037/a0020696>
- Crump, M. J. C., Gong, Z., & Milliken, B. (2006). The context-specific proportion congruent Stroop effect: Location as a contextual cue. *Psychonomic Bulletin & Review*, *13*(2), 316–321. <https://doi.org/10.3758/BF03193850>

- Crump, M. J., Vaquero, J. M., & Milliken, B. (2008). Context-specific learning and control: The roles of awareness, task relevance, and relative salience. *Consciousness and cognition, 17*(1), 22-36.
- Daly, T. M., & Natarajan, R. (2015). Swapping bricks for clicks: Crowdsourcing longitudinal data on Amazon Turk. *Journal of Business Research, 68*(12), 2603–2609. <https://doi.org/10.1016/j.jbusres.2015.05.001>
- De Baene, W., & Brass, M. (2013). Switch probability context (in)sensitivity within the cognitive control network. *NeuroImage, 77*, 207–214. <https://doi.org/10.1016/j.neuroimage.2013.03.057>
- De Jong, R. (2000). An intention-activation account of residual switch costs. In S Monsell & J. Driver (Eds.), *Control of Cognitive Processes: Attention and Performance XVIII* (pp. 357–376). MIT Press.
- de Vries, M., & Geurts, H. M. (2012). Cognitive Flexibility in ASD; Task Switching with Emotional Faces. *Journal of Autism and Developmental Disorders, 42*(12), 2558–2568. <https://doi.org/10.1007/s10803-012-1512-1>
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*(1), 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- DeSimone, J. C., Weiler, J., Aber, G. S., & Heath, M. (2014). The unidirectional prosaccade switch-cost: Correct and error antisaccades differentially influence the planning times for subsequent prosaccades. *Vision Research, 96*, 17–24. <https://doi.org/10.1016/j.visres.2013.12.005>
- Dibbets, P., Evers, E. A. T., Hurks, P. P. M., Bakker, K., & Jolles, J. (2010). Differential brain activation patterns in adult attention-deficit hyperactivity disorder (ADHD) associated with task switching. *Neuropsychology, 24*(4), 413–423. <https://doi.org/10.1037/a0018997>
- Diedenhofen, B., & Musch, J. (2017). PageFocus: Using paradata to detect and prevent cheating on online achievement tests. *Behavior Research Methods, 49*(4), 1444–1459. <https://doi.org/10.3758/s13428-016-0800-7>
- Dodds, C. M., Clark, L., Dove, A., Regenthal, R., Baumann, F., Bullmore, E., Robbins, T. W., & Müller, U. (2009). The dopamine D2 receptor antagonist sulpiride modulates striatal BOLD signal during the manipulation of information in

working memory. *Psychopharmacology*, 207(1), 35–45.  
<https://doi.org/10.1007/s00213-009-1634-0>

- Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychological Bulletin*, 145(5), 508–535.  
<https://doi.org/10.1037/bul0000192>
- Dreisbach, G., & Fröber, K. (2019). On How to Be Flexible (or Not): Modulation of the Stability-Flexibility Balance. *Current Directions in Psychological Science*, 28(1), 3–9.  
doi:10.1177/0963721418800030
- Dreisbach, G., & Haider, H. (2006). Preparatory adjustment of cognitive control in the task switching paradigm. *Psychonomic Bulletin & Review*, 13(2), 334–338.  
doi:10.3758/BF03193853
- Dreisbach, G., & Haider, H. (2008). That's what task sets are for: Shielding against irrelevant information. *Psychological Research*, 72(4), 355–361.  
<https://doi.org/10.1007/s00426-007-0131-5>
- Dreisbach, G., & Wenke, D. (2011). The shielding function of task sets and its relaxation during task switching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(6), 1540–1546. doi: 10.1037/a0024077
- Dreisbach, G., Haider, H., & Kluwe, R. H. (2002). Preparatory processes in the task-switching paradigm: Evidence from the use of probability cues. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3), 468–483.  
doi:10.1037/0278-7393.28.3.468
- Dudai, Y. (2012). The Restless Engram: Consolidations Never End. *Annual Review of Neuroscience*, 35(1), 227–247. <https://doi.org/10.1146/annurev-neuro-062111-150500>
- Duthoo, W., Abrahamse, E. L., Braem, S., Boehler, C. N., & Notebaert, W. (2014). The heterogeneous world of congruency sequence effects: An update. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.01001>
- Duthoo, W., De Baene, W., Wühr, P., & Notebaert, W. (2012). When Predictions Take Control: The Effect of Task Predictions on Task Switching Performance. *Frontiers in Psychology*, 3. doi:10.3389/fpsyg.2012.00282

- Egner, T. (2008). Multiple conflict-driven control mechanisms in the human brain. *Trends in Cognitive Sciences*, 12(10), 374-380. doi:<https://doi.org/10.1016/j.tics.2008.07.001>
- Egner, T. (2014). Creatures of habit (and control): a multi-level learning perspective on the modulation of congruency effects. *Frontiers in Psychology*, 5. doi:10.3389/fpsyg.2014.01247
- Egner, Tobias. (2007). Congruency sequence effects and cognitive control. *Cognitive, Affective, & Behavioral Neuroscience*, 7(4), 380–390. <https://doi.org/10.3758/CABN.7.4.380>
- Egner, Tobias. (2014). Creatures of habit (and control): A multi-level learning perspective on the modulation of congruency effects. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.01247>
- Eichenbaum, A., Scimeca, J. M., & D'Esposito, M. (2020). Dissociable Neural Systems Support the Learning and Transfer of Hierarchical Control Structure. *The Journal of Neuroscience*, 40(34), 6624–6637. <https://doi.org/10.1523/JNEUROSCI.0847-20.2020>
- Ellefson, M. R., Shapiro, L. R., & Chater, N. (2006). Asymmetrical switch costs in children. *Cognitive Development*, 21(2), 108–130. <https://doi.org/10.1016/j.cogdev.2006.01.002>
- Enkavi, A. Z., Eisenberg, I. W., Bissett, P. G., Mazza, G. L., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Large-scale analysis of test–retest reliabilities of self-regulation measures. *Proceedings of the National Academy of Sciences of the United States of America*, 116(12), 5472–5477. <https://doi.org/10.1073/pnas.1818430116>
- Farooqui, A. A., & Manly, T. (2015). Anticipatory control through associative learning of subliminal relations: invisible may be better than visible. *Psychological science*, 26(3), 325-334.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G\* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, 41(4), 1149-1160.
- Forster, S. E., & Cho, R. Y. (2014). Context Specificity of Post-Error and Post-Conflict Cognitive Control Adjustments. *PLOS ONE*, 9(3), e90281. <https://doi.org/10.1371/journal.pone.0090281>

- Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K., & Awh, E. (2017). Alpha-Band Oscillations Enable Spatially and Temporally Resolved Tracking of Covert Spatial Attention. *Psychological Science*, 28(7), 929–941. <https://doi.org/10.1177/0956797617699167>
- Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory: A computational model. 1(2), 137-160. doi:10.3758/cabn.1.2.137
- Frings, C., Hommel, B., Koch, I., Rothermund, K., Dignath, D., Giesen, C., Kiesel, A., Kunde, W., Mayr, S., Moeller, B., Möller, M., Pfister, R., & Philipp, A. (2020). Binding and Retrieval in Action Control (BRAC). *Trends in Cognitive Sciences*, 24(5), 375–387. <https://doi.org/10.1016/j.tics.2020.02.004>
- Geddert, R., & Egner, T. (2022). No need to choose: Independent regulation of cognitive stability and flexibility challenges the stability-flexibility trade-off. *Journal of Experimental Psychology: General*.
- Gershman, S. J., & Daw, N. D. (2017). Reinforcement Learning and Episodic Memory in Humans and Animals: An Integrative Framework. *Annual Review of Psychology*, 68(1), 101–128. <https://doi.org/10.1146/annurev-psych-122414-033625>
- Geurts, H. M., Begeer, S., & Stockmann, L. (2009). Brief report: Inhibitory control of socially relevant stimuli in children with high functioning autism. *Journal of Autism and Developmental Disorders*, 39(11), 1603.
- Gonthier, C., Braver, T. S., & Bugg, J. M. (2016). Dissociating proactive and reactive control in the Stroop task. *Memory & Cognition*, 44(5), 778-788.
- Goschke, T. (2000). Intentional reconfiguration and J-TI Involuntary Persistence In Task Set Switching. *Control of cognitive processes: Attention and performance XVIII*, 18, 331.
- Goschke, T. (2003). Voluntary action and cognitive control from a cognitive neuroscience perspective. In *Voluntary action: Brains, minds, and sociality*. (pp. 49–85). Oxford University Press.
- Goschke, T. (2013). Volition in Action: Intentions, Control Dilemmas, and the Dynamic Regulation of Cognitive Control. *Action Science: Foundations of an Emerging Discipline*, 409-434.

- Grant, L. D., Cookson, S. L., & Weissman, D. H. (2020). Task sets serve as boundaries for the congruency sequence effect. *Journal of experimental psychology. Human perception and performance*, 46(8), 798-812. <https://doi.org/10.1037/xhp0000750>
- Grent-'t-Jong, T., & Woldorff, M. G. (2007). Timing and Sequence of Brain Activity in Top-Down Control of Visual-Spatial Attention. *PLOS Biology*, 5(1), e12. <https://doi.org/10.1371/journal.pbio.0050012>
- Haaf, J. M., & Rouder, J. N. (2019). Some do and some don't? Accounting for variability of individual difference structures. *Psychonomic Bulletin & Review*, 26(3), 772–789. <https://doi.org/10.3758/s13423-018-1522-x>
- Han, J., Dai, Y., Xie, L., & Li, F. (2018). Brain responses associated with different hierarchical effects on cues and targets during rule shifting. *Biological Psychology*, 134, 52–63. <https://doi.org/10.1016/j.biopsycho.2018.02.010>
- Hazeltine, E., Lightman, E., Schwarb, H., & Schumacher, E. H. (2011). The boundaries of sequential modulations: Evidence for set-level control. *Journal of Experimental Psychology: Human Perception and Performance*, 37(6), 1898-1914. doi:10.1037/a0024662
- Hommel, B. (1998). Event Files: Evidence for Automatic Integration of Stimulus-Response Episodes. *Visual Cognition*, 5(1–2), 183–216. <https://doi.org/10.1080/713756773>
- Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences*, 8(11), 494–500. <https://doi.org/10.1016/j.tics.2004.08.007>
- Hommel, B. (2015). Between persistence and flexibility: The Yin and Yang of action control. In *Advances in motivation science* (Vol. 2, pp. 33-67). Elsevier.
- Jost, K., Mayr, U., & Rösler, F. (2008). Is task switching nothing but cue priming? Evidence from ERPs. *Cognitive, Affective, & Behavioral Neuroscience*, 8(1), 74–84. <https://doi.org/10.3758/CABN.8.1.74>
- Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, 24(5), 849–878. <https://doi.org/10.1017/S0140525X01000103>
- Hutchison, K. A. (2011). The interactive effects of list-wide control, item-based control, and working memory capacity on Stroop performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(4), 851-860. doi:10.1037/a0023437

- Jiang, J., Beck, J., Heller, K., & Egner, T. (2015). An insula-frontostriatal network mediates flexible cognitive control by adaptively predicting changing control demands. *Nature Communications*, 6, 8165. <https://doi.org/10.1038/ncomms9165>
- Jiang, J., Heller, K., & Egner, T. (2014). Bayesian modeling of flexible cognitive control. *Neuroscience & Biobehavioral Reviews*, 46, 30–43. <https://doi.org/10.1016/j.neubiorev.2014.06.001>
- Kandalowski, S. R. M., Seibold, J. C., Schuch, S., & Koch, I. (2020). Examining binding effects on task switch costs and response-repetition effects: Variations of the cue modality and stimulus modality in task switching. *Attention, Perception, & Psychophysics*, 82(4), 1632–1643. <https://doi.org/10.3758/s13414-019-01931-0>
- Karayanidis, F., Jamadar, S., Ruge, H., Phillips, N., Heathcote, A., & Forstmann, B. (2010). Advance Preparation in Task-Switching: Converging Evidence from Behavioral, Brain Activation, and Model-Based Approaches. *Frontiers in Psychology*, 1, 25. <https://doi.org/10.3389/fpsyg.2010.00025>
- Karayanidis, F., Provost, A., Brown, S., Paton, B., & Heathcote, A. (2011). Switch-specific and general preparation map onto different ERP components in a task-switching paradigm. *Psychophysiology*, 48(4), 559–568. <https://doi.org/10.1111/j.1469-8986.2010.01115.x>
- Kazak, A. E. (2018). Editorial: Journal article reporting standards. In (Vol. 73, pp. 1-2). US: American Psychological Association.
- Kenworthy, L., Case, L., Harms, M. B., Martin, A., & Wallace, G. L. (2010). Adaptive Behavior Ratings Correlate With Symptomatology and IQ Among Individuals With High-Functioning Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 40(4), 416–423. <https://doi.org/10.1007/s10803-009-0911-4>
- Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching—A review. *Psychological Bulletin*, 136(5), 849-874. doi:10.1037/a0019842
- King, J. A., Colla, M., Brass, M., Heuser, I., & von Cramon, D. (2007). Inefficient cognitive control in adult ADHD: Evidence from trial-by-trial Stroop test and cued task switching performance. *Behavioral and Brain Functions*, 3(1), 42. <https://doi.org/10.1186/1744-9081-3-42>
- King, J. A., Donkin, C., Korb, F. M., & Egner, T. (2012). Model-based analysis of context-specific cognitive control. *Front Psychol*, 3, 358. doi:10.3389/fpsyg.2012.00358

- King, J. A., Korb, F. M., & Egner, T. (2012). Priming of Control: Implicit Contextual Cuing of Top-down Attentional Set. *Journal of Neuroscience*, 32(24), 8192-8200. <https://doi.org/10.1523/jneurosci.0934-12.2012>
- Kleinsorge, T. (1999). Response repetition benefits and costs. *Acta Psychologica*, 103(3), 295–310. [https://doi.org/10.1016/S0001-6918\(99\)00047-5](https://doi.org/10.1016/S0001-6918(99)00047-5)
- Klug, M, Gramann, K. Identifying key factors for improving ICA-based decomposition of EEG data in mobile and stationary experiments. *Eur J Neurosci*. 2020; 00: 000– 000. <https://doi.org/10.1111/ejn.14992>
- Koch, I., Frings, C., & Schuch, S. (2018). Explaining response-repetition effects in task switching: Evidence from switching cue modality suggests episodic binding and response inhibition. *Psychological Research*, 82(3), 570–579. <https://doi.org/10.1007/s00426-017-0847-9>
- Koch, I., Poljac, E., Müller, H., & Kiesel, A. (2018). Cognitive structure, flexibility, and plasticity in human multitasking – An integrative review of dual-task and task-switching research. *Psychological Bulletin*, 144(6), 557–583. doi:10.1037/bul0000144
- Korb, F. M., Jiang, J., King, J. A., & Egner, T. (2017). Hierarchically Organized Medial Frontal Cortex-Basal Ganglia Loops Selectively Control Task- and Response-Selection. *The Journal of Neuroscience*, 37(33), 7893–7905. <https://doi.org/10.1523/JNEUROSCI.3289-16.2017>
- Lavric, A., Mizon, G. A., & Monsell, S. (2008). Neurophysiological signature of effective anticipatory task-set control: A task-switching investigation. *European Journal of Neuroscience*, 28(5), 1016–1029.
- Leboe, J. P., Wong, J., Crump, M., & Stobbe, K. (2008). Probe-specific proportion task repetition effects on switching costs. *Perception & Psychophysics*, 70(6), 935-945. doi:10.3758/PP.70.6.935
- Lee, W.-T., Hazeltine, E., Jiang, J., & Jiang, J. (2022). Interference and Integration in Hierarchical Task Learning.
- Logan, G. D., & Bundesen, C. (2003). Clever homunculus: Is there an endogenous act of control in the explicit task-cuing procedure? *Journal of Experimental Psychology: Human Perception and Performance*, 29(3), 575–599. <https://doi.org/10.1037/0096-1523.29.3.575>



- Logan, G. D., & Bundesen, C. (2004). Very clever homunculus: Compound stimulus strategies for the explicit task-cuing procedure. *Psychonomic Bulletin & Review*, *11*(5), 832–840. <https://doi.org/10.3758/BF03196709>
- Logan, G. D., & Schneider, D. W. (2006). Interpreting instructional cues in task switching procedures: The role of mediator retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(2), 347–363. <https://doi.org/10.1037/0278-7393.32.3.347>
- Lowe, D. G., & Mitterer, J. O. (1982). Selective and divided attention in a Stroop task. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *36*(4), 684–700. <https://doi.org/10.1037/h0080661>
- Mayr, U. (2006). What matters in the cued task-switching paradigm: Tasks or cues? *Psychonomic Bulletin & Review*, *13*(5), 794–799. doi:10.3758/bf03193999
- Mayr, U., & Keele, S. W. (2000). Changing internal constraints on action: The role of backward inhibition. *Journal of Experimental Psychology: General*, *129*(1), 4–26. doi:10.1037/0096-3445.129.1.4
- Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1124–1140. <https://doi.org/10.1037/0278-7393.26.5.1124>
- Mayr, U., & Kliegl, R. (2003). Differential effects of cue changes and task changes on task-set selection costs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(3), 362–372. doi:10.1037/0278-7393.29.3.362
- Mayr, U., Awh, E., & Laurey, P. (2003). Conflict adaptation effects in the absence of executive control. *Nature Neuroscience*, *6*(5), 450–452. <https://doi.org/10.1038/nn1051>
- McDonald, K., Broderick, W. F., Huettel, S., & Pearson, J. (2019). Bayesian nonparametric models characterize instantaneous strategies in a competitive dynamic game. *Nature Communications*, *10*(1), 1–12. <https://doi.org/10.1101/385195>
- McKewen, M., Cooper, P. S., Wong, A. S. W., Michie, P. T., Sauseng, P., & Karayanidis, F. (2020). Task-switching costs have distinct phase-locked and nonphase-locked EEG power effects. *Psychophysiology*, *57*(5), e13533. doi:10.1111/psyp.13533
- Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(6), 1423–1442.

- Meiran, N. (2000). Modeling cognitive control in task-switching. *Psychological Research*, 63(3), 234–249.
- Meiran, N. (2010). Task Switching: Mechanisms Underlying Rigid vs. Flexible Self-Control. In R. Hassin, K. Ochsner, & Y. Trope (Eds.), *Self Control in Society, Mind, and Brain* (pp. 202–220). Oxford University Press.  
doi:10.1093/acprof:oso/9780195391381.003.0011
- Meiran, N., Chorev, Z., & Sapir, A. (2000). Component Processes in Task Switching. *Cognitive Psychology*, 41(3), 211–253. doi:10.1006/cogp.2000.0736
- Miendlarzewska, E. A., Ciucci, S., Cannistraci, C. V., Bavelier, D., & Schwartz, S. (2018). Reward-enhanced encoding improves relearning of forgotten associations. *Scientific Reports*, 8(1), 8557. <https://doi.org/10.1038/s41598-018-26929-w>
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134–140.  
doi:10.1016/S1364-6613(03)00028-7
- Monsell, S., & Mizon, G. A. (2006). Can the task-cuing paradigm measure an endogenous task-set reconfiguration process? *Journal of Experimental Psychology: Human Perception and Performance*, 32(3), 493–516. doi:10.1037/0096-1523.32.3.493
- Monsell, S., Yeung, N., & Azuma, R. (2000). Reconfiguration of task-set: Is it easier to switch to the weaker task? *Psychological Research*, 15.
- Moss, A. J., Rosenzweig, C., Robinson, J., & Litman, L. (2020). Demographic Stability on Mechanical Turk Despite COVID-19. *Trends in Cognitive Sciences*.  
<https://doi.org/10.1016/j.tics.2020.05.014>
- Murty, V. P., FeldmanHall, O., Hunter, L. E., Phelps, E. A., & Davachi, L. (2016). Episodic memories predict adaptive value-based decision-making. *Journal of Experimental Psychology: General*, 145(5), 548–558.  
<https://doi.org/10.1037/xge0000158>
- Nessler, D. (2012). A new account of the effect of probability on task switching: ERP evidence following the manipulation of switch probability, cue informativeness and predictability. *Biological Psychology*, 18.
- Nicholson, R., Karayanidis, F., Bumak, E., Poboka, D., & Michie, P. T. (2006). ERPs dissociate the effects of switching task sets and task cues. *Brain Research*, 1095(1), 107–123.

- O'Reilly, R. C., & Frank, M. J. (2006). Making Working Memory Work: A Computational Model of Learning in the Prefrontal Cortex and Basal Ganglia. *Neural Computation*, 18(2), 283-328. doi:10.1162/089976606775093909
- Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2010). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, 2011, 156869. <https://doi.org/10.1155/2011/156869>
- Ophir, E., Nass, C., & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences*, 106(37), 15583–15587. <https://doi.org/10.1073/pnas.0903620106>
- Patil, A., Murty, V. P., Dunsmoor, J. E., Phelps, E. A., & Davachi, L. (2017). Reward retroactively enhances memory consolidation for related items. *Learning & Memory*, 24(1), 65–69. <https://doi.org/10.1101/lm.042978.116>
- Permut, S., Fisher, M., & Oppenheimer, D. M. (2019). TaskMaster: A Tool for Determining When Subjects Are on Task. *Advances in Methods and Practices in Psychological Science*, 2(2), 188–196. <https://doi.org/10.1177/2515245919838479>
- Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology*, 72(2), 184–187. [https://doi.org/10.1016/0013-4694\(89\)90180-6](https://doi.org/10.1016/0013-4694(89)90180-6)
- Pfeuffer, C. U., Moutsopoulou, K., Waszak, F., & Kiesel, A. (2019). Execution-based and verbal code-based stimulus–response associations: Proportion manipulations reveal conflict adaptation processes in item-specific priming. *Psychological Research*. <https://doi.org/10.1007/s00426-019-01220-3>
- Qiao, L., Zhang, L., Chen, A., & Egner, T. (2017). Dynamic Trial-by-Trial Recoding of Task-Set Representations in the Frontoparietal Cortex Mediates Behavioral Flexibility. *The Journal of Neuroscience*, 37(45), 11037-11050. doi:10.1523/jneurosci.0935-17.2017
- R Core Team. (2019). *R: A language and environment for statistical computing*. In R Foundation for Statistical Computing. <https://www.R-project.org/>
- Reuss, H., Desender, K., Kiesel, A., & Kunde, W. (2014). Unconscious conflicts in unconscious contexts: The role of awareness and timing in flexible conflict adaptation. *Journal of Experimental Psychology: General*, 143(4), 1701.

- Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124(2), 207–231. doi:10.1037/0096-3445.124.2.207
- Rose NS, LaRocque JJ, Riggall AC, Gosseries O, Starrett MJ, Meyering EE, Postle BR. Reactivation of latent working memories with transcranial magnetic stimulation. *Science*. 2016 Dec 2;354(6316):1136–1139. doi: 10.1126/science.aah7011. PMID: 27934762; PMCID: PMC5221753.
- Ruthruff, E., Remington, R. W., & Johnston, J. C. (2001). Switching between simple cognitive tasks: The interaction of top-down and bottom-up factors. *Journal of Experimental Psychology. Human Perception and Performance*, 27(6), 1404–1419. <https://doi.org/10.1037//0096-1523.27.6.1404>
- Sawaguchi, T., & Goldman-Rakic, P. S. (1991). D1 Dopamine Receptors in Prefrontal Cortex: Involvement in Working Memory. *Science*, 251(4996), 947–950.
- Schmidt, J. R., & De Houwer, J. (2011). Now you see it, now you don't: Controlling for contingencies and stimulus repetitions eliminates the Gratton effect. *Acta Psychologica*, 138(1), 176–186. <https://doi.org/10.1016/j.actpsy.2011.06.002>
- Schmidt, J. R., & Lemerrier, C. (2019). Context-specific proportion congruent effects: Compound-cue contingency learning in disguise. *Quarterly Journal of Experimental Psychology*, 72(5), 1119–1130. <https://doi.org/10.1177/1747021818787155>
- Schmidt, J. R., & Liefooghe, B. (2016). Feature Integration and Task Switching: Diminished Switch Costs after Controlling for Stimulus, Response, and Cue Repetitions. *PLOS ONE*, 11(3), e0151188. <https://doi.org/10.1371/journal.pone.0151188>
- Schmidt, J. R., Crump, M. J., Cheesman, J., & Besner, D. (2007). Contingency learning without awareness: Evidence for implicit control. *Consciousness and cognition*, 16(2), 421–435.
- Schneider, D. W. (2015). Attentional control of response selection in task switching. *Journal of Experimental Psychology: Human Perception and Performance*, 41(5), 1315–1324. doi: 10.1037/xhp0000091.
- Schneider, D. W., & Anderson, J. R. (2010). Asymmetric Switch Costs as Sequential Difficulty Effects. *Quarterly Journal of Experimental Psychology*, 63(10), 1873–1894. <https://doi.org/10.1080/17470211003624010>

- Schneider, D. W., & Logan, G. D. (2005). Modeling Task Switching Without Switching Tasks: A Short-Term Priming Account of Explicitly Cued Performance. *Journal of Experimental Psychology: General*, *134*(3), 343–367. <https://doi.org/10.1037/0096-3445.134.3.343>
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*(1), 1–66. <https://doi.org/10.1037/0033-295X.84.1.1>
- Schumacher, E. H., & Hazeltine, E. (2016). Hierarchical Task Representation: Task Files and Response Selection. *Current Directions in Psychological Science*, *25*(6), 449–454. <https://doi.org/10.1177/0963721416665085>
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*(2), 127–190. <https://doi.org/10.1037/0033-295X.84.2.127>
- Shin, Y., Proctor, R., & Capaldi, E. (2010). A Review of Contemporary Ideomotor Theory. *Psychological Bulletin*, *136*(6), 943–974. <https://doi.org/10.1037/a0020541>
- Shohamy, D., & Daw, N. D. (2015). Integrating memories to guide decisions. *Current Opinion in Behavioral Sciences*, *5*(Supplement C), 85–90. <https://doi.org/10.1016/j.cobeha.2015.08.010>
- Sinai, M., Goffaux, P., & Phillips, N. A. (2007). Cue- versus response-locked processes in backward inhibition: Evidence from ERPs. *Psychophysiology*, *44*(4), 596–609. <https://doi.org/10.1111/j.1469-8986.2007.00527.x>
- Siqi-Liu, A., & Egner, T. (2020). Contextual Adaptation of Cognitive Flexibility is driven by Task- and Item-Level Learning. *Cognitive, Affective & Behavioral Neuroscience*, *20*(4), 757–782. <https://doi.org/10.3758/s13415-020-00801-9>
- Siqi-Liu, A., Egner, T., & Woldorff, M. G. (2022). Neural dynamics of context-sensitive adjustments in cognitive flexibility. *Journal of Cognitive Neuroscience*, *34*(3), 480–494.
- Siqi-Liu, A., & Egner, T. (2022). Task sets define boundaries of learned cognitive flexibility in list-wide switch proportion manipulations.
- Smith, D. M., Dykstra, T., Hazeltine, E., & Schumacher, E. H. (2020). Task representation affects the boundaries of behavioral slowing following an error. *Attention*,

*Perception, & Psychophysics*, 82(5), 2315–2326. <https://doi.org/10.3758/s13414-020-01985-5>

Sohn, M. H., & Anderson, J. R. (2001). Task preparation and task repetition: Two-component model of task switching. *Journal of Experimental Psychology: General*, 130(4), 764.

Spinelli, G., Perry, J. R., & Lupker, S. J. (2019). Adaptation to conflict frequency without contingency and temporal learning: Evidence from the picture-word interference task. *Journal of Experimental Psychology. Human Perception and Performance*, 45(8), 995–1014. <https://doi.org/10.1037/xhp0000656>

Steinhauser, M. (2010). How to correct a task error: Task-switch effects following different types of error correction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(4), 1028–1035. <https://doi.org/10.1037/a0019340>

Stokes, M. G. (2015). ‘Activity-silent’ working memory in prefrontal cortex: a dynamic coding framework. *Trends in cognitive sciences*, 19(7), 394–405.

Surrey, C., Dreisbach, G., & Fischer, R. (2017). Context-Specific Adjustment of Cognitive Control: Transfer of Adaptive Control Sets. *Quarterly Journal of Experimental Psychology*, 70(11), 2386–2401. doi: 10.1080/17470218.2016.1239748

Tambini, A., & Davachi, L. (2019). Awake Reactivation of Prior Experiences Consolidates Memories and Biases Cognition. *Trends in Cognitive Sciences*, 23(10), 876–890. <https://doi.org/10.1016/j.tics.2019.07.008>

Ullsperger, M., Bylsma, L. M., & Botvinick, M. M. (2005). The conflict adaptation effect: It’s not just priming. *Cognitive, Affective, & Behavioral Neuroscience*, 5(4), 467–472. <https://doi.org/10.3758/CABN.5.4.467>

van Langen, J. (2020). *Open-visualizations in R and Python*. <https://github.com/jorvlan/open-visualizations>

van Schouwenburg, M. R., den Ouden, H. E. M., & Cools, R. (2010). The Human Basal Ganglia Modulate Frontal-Posterior Connectivity during Attention Shifting. *Journal of Neuroscience*, 30(29), 9910–9918. <https://doi.org/10.1523/JNEUROSCI.1111-10.2010>

Vandierendonck, A., Liefoghe, B., & Verbruggen, F. (2010). Task switching: Interplay of reconfiguration and interference control. *Psychological Bulletin*, 136(4), 601–626. doi:10.1037/a0019791

- Vijayraghavan, S., Wang, M., Birnbaum, S. G., Williams, G. V., & Arnsten, A. F. T. (2007). Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. *Nature Neuroscience*, *10*(3), 376–384. <https://doi.org/10.1038/nm1846>
- Waskom, M. L., Kumaran, D., Gordon, A. M., Rissman, J., & Wagner, A. D. (2014). Frontoparietal Representations of Task Context Support the Flexible Control of Goal-Directed Cognition. *34*(32), 10743–10755. doi:10.1523/jneurosci.5282-13.2014
- Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus–task bindings in task-shift costs. *Cognitive Psychology*, *46*(4), 361–413. doi: 10.1016/S0010-0285(02)00520-0
- Weidler, B. J., & Bugg, J. M. (2016). Transfer of location-specific control to untrained locations. *Q J Exp Psychol (Hove)*, *69*(11), 2202–2217. <https://doi.org/10.1080/17470218.2015.1111396>
- Wen, T., & Egner, T. (2022). Retrieval context determines whether event boundaries impair or enhance temporal order memory. *Cognition*, *225*, 105145. <https://doi.org/10.1016/j.cognition.2022.105145>
- Wen, T., Geddert, R. M., Madlon-Kay, S., & Egner, T. (2023). Transfer of learned cognitive flexibility to novel stimuli and task sets. *Psychological Science*.
- Whitehead, P. S., Pfeuffer, C. U., & Egner, T. (2020). Memories of control: One-shot episodic learning of item-specific stimulus-control associations. *Cognition*, *199*, 104220. <https://doi.org/10.1016/j.cognition.2020.104220>
- Wimmer, G. E., & Büchel, C. (2016). Reactivation of Reward-Related Patterns from Single Past Episodes Supports Memory-Based Decision Making. *The Journal of Neuroscience*, *36*(10), 2868–2880. <https://doi.org/10.1523/JNEUROSCI.3433-15.2016>
- Wimmer, G. E., & Shohamy, D. (2012). Preference by Association: How Memory Mechanisms in the Hippocampus Bias Decisions. *Science*, *338*(6104), 270–273. <https://doi.org/10.1126/science.1223252>
- Wimmer, G. E., Li, J. K., Gorgolewski, K. J., & Poldrack, R. A. (2018). Reward Learning over Weeks Versus Minutes Increases the Neural Representation of Value in the Human Brain. *Journal of Neuroscience*, *38*(35), 7649–7666. <https://doi.org/10.1523/JNEUROSCI.0075-18.2018>

- Winkler I, Debener S, Müller KR, Tangermann M. On the influence of high-pass filtering on ICA-based artifact reduction in EEG-ERP. *Annu Int Conf IEEE Eng Med Biol Soc.* 2015;2015:4101-5. doi: 10.1109/EMBC.2015.7319296. PMID: 26737196.
- Wisniewski, D., Reverberi, C., Momennejad, I., Kahnt, T., & Haynes, J. D. (2015). The Role of the Parietal Cortex in the Representation of Task-Reward Associations. *Journal of Neuroscience*, 35(36), 12355-12365. doi:10.1523/jneurosci.4882-14.2015
- Woldorff, M. G., Liotti, M., Seabolt, M., Busse, L., Lancaster, J. L., & Fox, P. T. (2002). The temporal dynamics of the effects in occipital cortex of visual-spatial selective attention. *Neural Mechanisms of Executive Control in Cognition*, 15(1), 1–15. [https://doi.org/10.1016/S0926-6410\(02\)00212-4](https://doi.org/10.1016/S0926-6410(02)00212-4)
- Wolff MJ, Jochim J, Akyürek EG, Stokes MG. Dynamic hidden states underlying working-memory-guided behavior. *Nat Neurosci.* 2017 Jun;20(6):864-871. doi: 10.1038/nn.4546. Epub 2017 Apr 17. PMID: 28414333; PMCID: PMC5446784.
- Wong, A. S., Cooper, P. S., Conley, A. C., McKewen, M., Fulham, W. R., Michie, P. T., & Karayanidis, F. (2018). Event-related potential responses to task switching are sensitive to choice of spatial filter. *Frontiers in Neuroscience*, 12, 143.
- Wong, J., & Leboe, J. P. (2009). Distinguishing between inhibitory and episodic processing accounts of switch-cost asymmetries. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 63(1), 8–23. <https://doi.org/10.1037/a0013222>
- Woolgar, A., Hampshire, A., Thompson, R., & Duncan, J. (2011). Adaptive coding of task-relevant information in human frontoparietal cortex. *Journal of Neuroscience*, 31(41), 14592-14599.
- Wu, S., Hitchman, G., Tan, J., Zhao, Y., Tang, D., Wang, L., & Chen, A. (2015). The neural dynamic mechanisms of asymmetric switch costs in a combined Stroop-task-switching paradigm. *Scientific Reports*, 5(1), 10240. <https://doi.org/10.1038/srep10240>
- Yeung, N., & Monsell, S. (2003). Switching Between Tasks of Unequal Familiarity: The Role of Stimulus-Attribute and Response-Set Selection. *Journal of Experimental Psychology. Human Perception and Performance*, 29, 455–469. <https://doi.org/10.1037/0096-1523.29.2.455>



Zhang, R., Stock, A.-K., Fischer, R., & Beste, C. (2016). The system neurophysiological basis of backward inhibition. *Brain Structure and Function*, 221(9), 4575–4587.  
<https://doi.org/10.1007/s00429-016-1186-0>

## Biography

Audrey Siqi-Liu completed her undergraduate degree at Claremont McKenna College, where she graduated *magna cum laude* in 2018 with an honors major in Psychology and a dual major in Literature. She then enrolled as a graduate student in the Cognitive Neuroscience Admitting Program at Duke University, later joining the labs of Dr. Tobias Egner and Dr. Marty Woldorff and affiliating with the Department of Psychology and Neuroscience. There she has investigated the learning mechanisms and neural correlates for adaptive task performance. Her research has been published in *Journal of Cognitive Neuroscience*, *Cognitive, Affective, and Behavioral Neuroscience*, *Journal of Experimental Psychology: Learning, Memory & Cognition*, *Journal of Experimental Psychology: Human Perception and Performance*, and others.