

**PREDICTABLE SWITCHING BETWEEN
RECOGNITION MEMORY AND MAGNITUDE
JUDGEMENT TASKS: EFFECTS AND
IMPLICATIONS**

BY

REBECCA ANNE BURSON

A thesis

**submitted to the Victoria University of Wellington
in fulfilment of the requirements for the degree of
Doctor of Philosophy**

Victoria University of Wellington

2011

Abstract

Task switching and interruption effects—slower and (often) less accurate responses when a task changes compared to that when a task remains the same—have been investigated from both theoretical and applied vantage points (e.g., Altmann & Trafton, 2007; Jersild, 1927; Rogers & Monsell, 1995). The task switching research has typically used simple tasks with high stimulus-response (S-R) overlap, but there is a need to use different methods and tasks to test the boundaries of task switching effects and the theories used to explain them (Logan, 2003). This thesis examined the costs of switching between a recognition memory task, which is a more complex task than those typically used, and a magnitude judgement task (for the number of dots in a spatial array), which is the type of simple task that has been used (e.g., Altmann, 2002; Monsell, Sumner & Waters, 2003). Across seven experiments, participants switched between the recognition and magnitude tasks in predictable 1, 2, or 4-trial runs. The first two experiments examined task switching effects on recognition memory performance, with Experiment 2 investigating whether specific recognition processes (i.e., recollection and/or familiarity) were affected by switching tasks. Experiment 3 investigated the recovery from a task switch for both tasks and included a visual, task switch “reminder” cue in an attempt to improve switching performance. Finally, Experiments 4A, 4B, 5A, and 5B examined evidence for two well-known task switching phenomena, the practice effect and the preparation effect. The results led to four critical conclusions: (1) switching between two tasks with minimal S-R overlap produced significant RT and accuracy switch effects; (2) the cost to recognition memory accuracy did not reflect an impairment to controlled recollection processes; (3) the magnitude and persistence of task switching effects changed as a function of practice within an experiment; and (4) there was little evidence that participants began to switch tasks in advance of stimulus presentation. The results reported in this thesis provide a clear example of task switching driven by the type of stimulus (word or dots), where a change in stimulus type (i.e., from word to dots array or vice versa) initiated the time consuming process of retrieving/activating the appropriate task set. Future research will need to clarify whether the stimulus-driven nature of switching between the recognition and magnitude tasks remains when using different task switching paradigms and when S-R overlap is reintroduced.

Preface

This thesis begins by introducing the reader to recognition memory and briefly discussing theoretical standpoints on the processes contributing to recognition decisions. After discussing experimental manipulations that affect recognition memory performance in Chapter One, the effects of periodic interruptions to memory are discussed in Chapter Two. Chapter Three suggests that these interruptions can be interpreted as task switches. After introducing the reader to various methods of exploring task switching and a selection of common task switching effects, the chapter closes by briefly discussing theoretical accounts of task switching. Chapter Four reviews previous task switching literature that has used memory tasks and introduces the two tasks that will be used in this thesis, a recognition memory task and an unrelated magnitude task. The chapter closes by presenting the aims of this thesis and an overview of the general method to be used to examine the cost of switching between these tasks. Chapters Five and Six focus solely on recognition memory performance (RT and corrected accuracy). To establish the most effective way of investigating task switching effects with this relatively novel task (the recognition task), Chapter Five assesses the cost of switching between the two tasks in 1-, 2-, and 4-trial runs. Chapter Six focuses on accuracy switch effects and attempts to determine whether the cost to recognition memory accuracy can be isolated to a specific memory process (i.e., controlled recollection). With a more general focus on task switching performance and manipulations that affect it, Chapters Seven and Eight present data from both tasks. In Chapter Seven, performance across 4-trial runs is analysed for both tasks with the goal of establishing whether recovery from a task switch is immediate or protracted across multiple trials. Additionally, the benefit of providing a visual cue to remind participants of their position within a run is assessed. Chapter Eight presents four experiments that examine evidence for two common task switching effects—preparation effects and practice effects. Specifically, these experiments aim to establish whether participants can use the time available prior to the start of a trial to prepare for a known, predictable task switch. Finally, Chapter Nine summarises the key findings of the experiments before discussing how these findings can be interpreted within recognition memory and task switching theory. The chapter closes by discussing future directions for research examining the cost of switching between two tasks that do not share stimulus sets, such as the recognition and magnitude tasks.

Acknowledgements

First and foremost, I want to thank my supervisor, Dr Todd Jones. It has been an awesome journey and your knowledge, support, and encouragement is truly appreciated. Thank you to the members of CogLab (Jan Lauwereyns, Gina Grimshaw, Matt Crawford, and Matt Weaver in particular) at Victoria University of Wellington, who provided valuable feedback as I developed my research. I would also like to acknowledge Associate Professor Paul Atchley, who provided valuable comments on parts of this research. Thanks to Dr Matt Gerrie for the incredible encouragement and advice about navigating the world of the PhD student. While based at Washington University in St Louis, Professor Kathleen McDermott and her cognitive lab (especially Katie Arnold and Bridgid Finn) provided excellent support and feedback.

This thesis was supported by a Victoria University PhD scholarship. I also owe thanks to the Faculty of Science for their financial support that allowed me to attend the 50th Psychonomic Society conference to present data from this thesis. Finally, I was the grateful recipient of a Fulbright-Ministry of Research, Science and Technology Graduate Award, which funded my time as a visiting student researcher at Washington University in St Louis.

Mum and Dad, thank you. Through the ups and downs you were always there with an encouraging word, even when you didn't understand what I was talking about. Christine and Tony Hoggart—you opened your home to me and treated me like a daughter. Thank you for your love and support. Finally, I have to thank Jason Edwards. Jase, I can never thank you enough for your calming influence, your patience, and your never-ending encouragement. Thank you for reminding me of what Tolkien (by way of Gandalf) said: “a wizard [or thesis] is never late... he arrives precisely when he means to.” With love, Rebecca.

Table of Contents

Abstract.....	iii
Preface	v
Acknowledgements	vii
Table of Contents	ix
List of Figures.....	xiii
List of Tables	xv
Chapter One.....	1
Memory	1
Recognition Memory	2
Experimental Manipulations that Affect Recognition Performance	7
Chapter Two	11
Interruptions and Memory	11
Chapter Three	17
Task Switching	17
Origin of Task Switching Research and Theory	18
Methods of Investigating Task Switching	19
Key Task Switching Effects and Manipulations	23
Task Switching Theories and Models	30
Chapter Four	39
Task Switching and Episodic Memory.....	39
Justification for Current Studies	42
General Methodology	42
The Aims of This Thesis	43
Chapter Five	45
Experiment 1: Establishing Task Switching Effects	45
Method.....	47
Participants	47
Materials	47

Design and Procedure	47
Results	50
Accuracy	50
Reaction Times	51
Exploring Practice Effects	52
Discussion.....	54
Summary.....	55
Chapter Six	57
Experiment 2: Task Switching and Level of Processing Effects.....	57
Method.....	59
Participants	59
Materials	60
Design.....	60
Procedure.....	60
Results	63
Corrected Accuracy	63
Reaction Times	65
Discussion.....	67
Summary.....	70
Chapter Seven.....	71
Experiment 3: Recovery from a Task Switch and Transition Cuing.....	71
Method.....	73
Participants	73
Materials	73
Design.....	73
Procedure.....	74
Results	76
Accuracy.....	76
Reaction Times	81
Discussion.....	87

Effects of Providing a Sequence Cue	87
Performance Changes Within a Run	88
Chapter Eight	91
Experiments 4A – 5B: Preparation and Practice	91
General Method	95
Participants	95
Materials	96
Design	96
Procedure	96
Experiment 4A	99
Results and Discussion	100
Accuracy	100
Reaction Times	103
Experiment 4B	108
Results and Discussion	111
Accuracy	111
Reaction Times	114
Conclusions for Experiments 4A and 4B	120
Experiment 5A and 5B	121
Experiment 5A	122
Results and Discussion	122
Accuracy	122
Reaction Times	126
Experiment 5B	132
Results and Discussion	132
Accuracy	132
Reaction Times	136
Conclusions for Experiments 5A and 5B	141
Discussion	145
Practice Effects	146
Preparation Effects	147

Summary.....	152
Chapter Nine.....	153
General Discussion	153
Key Findings	153
Interpreting the Findings: Task Set Reconfiguration	159
Interpreting the Findings: Cognitive Control Model	161
Alternative Sources of Switch Cost.....	163
Challenging Definitions of Task Switching	164
Future Directions	166
Conclusions	167
Appendix A	169
Stimuli	169
Appendix B.....	183
Hits and False Alarms for Experiments 1 – 5B	183
Appendix C.....	191
Signal Detection Analyses for Experiment 2	191
Appendix D	193
Supplementary Tables for Experiments 4A – 5B	193
References	197

List of Figures

<i>Figure 2.1</i>	The design of the seven-trial runs used in Jones et al. (2011).	15
<i>Figure 2.2</i>	Performance on critical recognition memory trials collapsed across Experiment 1A and 1B of Jones et al. (2011).....	16
<i>Figure 3.1</i>	An illustration of the timing parameters in the explicit task cuing paradigm.	22
<i>Figure 3.2</i>	Reprint of Rogers and Monsell (1995)	25
<i>Figure 3.3</i>	The executive control process and the task process of Rubinstein, Meyer, and Evans (2000) two-stage reconfiguration model of task switching	32
<i>Figure 5.1</i>	Trial and task progression for the 1-, 2-, and 4-trial run conditions and the control run condition in Experiment 1.	49
<i>Figure 5.2</i>	Mean RTs for correct critical recognition decisions in Experiment 1 as a function of run condition.	52
<i>Figure 5.3</i>	Mean RTs of correct critical recognition memory decisions in Experiment 1 as a function of test third and run length condition....	53
<i>Figure 6.1</i>	Mean corrected accuracy of recognition decisions in Experiment 2 as a function of task transition and level of processing at study	64
<i>Figure 6.2</i>	Mean RTs for correct recognition memory decisions for deeply encoded words in Experiment 2 as a function of task transition and test block	67
<i>Figure 7.1</i>	Mean corrected accuracy of recognition memory decisions in Experiment 3 as a function of run position	78
<i>Figure 7.2</i>	Mean accuracy of magnitude judgements in Experiment 3 as a function of run position.....	80
<i>Figure 7.3</i>	Mean RTs for correct recognition memory decisions in Experiment 3 as a function of run position and test block	83
<i>Figure 7.4</i>	Mean RTs for correct magnitude judgements in Experiment 3 as a function of run position.....	86
<i>Figure 8.1</i>	Mean RT switch costs for Experiment 4A as a function of task and item type	106

<i>Figure 8.2</i>	A prototypical cumulative RT distribution illustrating a preparation effect	109
<i>Figure 8.3</i>	Cumulative RT distributions of correct recognition memory trials for Experiment 4A as a function of RSI and task transition.....	110
<i>Figure 8.4</i>	Mean RT switch costs for Experiment 4B as a function of RSI and test block.....	117
<i>Figure 8.5</i>	Mean RT switch costs for Experiment 4B as a function of RSI, task and item type.....	118
<i>Figure 8.6</i>	Cumulative RT distributions of correct recognition memory trials for Experiment 4B as a function of RSI and task transition.....	119
<i>Figure 8.7</i>	Mean RT switch costs for Experiment 5A as a function of RSI and test block.....	130
<i>Figure 8.8</i>	Cumulative RT distributions of correct recognition memory trials for Experiment 5A as a function of RSI and task transition.....	131
<i>Figure 8.9</i>	Mean accuracy of magnitude judgements in Experiment 5B as a function of RSI and transition.....	135
<i>Figure 8.10</i>	Mean RT switch costs for Experiment 5B as a function of test block and item type.....	139
<i>Figure 8.11</i>	Mean RT switch costs for Experiment 5B as a function of RSI and item type	140
<i>Figure 8.12</i>	Cumulative RT distributions of correct ‘new’ recognition memory trials for Experiment 5B as a function of RSI and task transition ..	142
<i>Figure 8.13</i>	Cumulative RT distributions of correct ‘old’ recognition memory trials for Experiment 5B as a function of RSI and task transition ..	143

List of Tables

Table 5.1	<i>Means and Standard Deviations for Correct Critical Recognition RTs in Experiment 1 as a Function of Item Type and Run Condition</i>	51
Table 6.1	<i>Means and Standard Deviations for Corrected Recognition Accuracy in Experiment 2 as a Function of Test Block, Level of Processing at Study, and Task Transition</i>	63
Table 6.2	<i>Means and Standard Deviations for Correct Recognition Decisions in Experiment 2 as a Function of Test Block, Item Type, and Task Transition.....</i>	66
Table 7.1	<i>Means and Standard Deviations for Corrected Recognition Accuracy in Experiment 3 as a Function of Position in Run and Sequence Cue Condition</i>	77
Table 7.2	<i>Means and Standard Deviations for Magnitude Judgement Accuracy in Experiment 3 as a Function of Position in Run and Sequence Cue Condition</i>	79
Table 7.3	<i>Means and Standard Deviations for Magnitude Judgement Accuracy in Experiment 3 as a Function of Position in Run and Test Block...</i>	81
Table 7.4	<i>Means and Standard Deviations for Correct Recognition RTs in Experiment 3 as a Function of Position in Run, Test Block, and Sequence Cue Condition.....</i>	82
Table 7.5	<i>Means and Standard Deviations for Correct Magnitude Judgement RTs in Experiment 3 as a Function of Position in Run, Test Block, and Sequence Cue Condition.....</i>	85
Table 8.1	<i>Characteristics of Experiments in Chapter 8</i>	97
Table 8.2	<i>Means and Standard Deviations for Corrected Recognition Accuracy in Experiment 4A as a Function of RSI, Test Block, and Transition.....</i>	100
Table 8.3	<i>Means and Standard Deviations for Magnitude Judgement Accuracy in Experiment 4A as a Function of RSI, Test Block, Transition, and Decision</i>	102

Table 8.4	<i>Means and Standard Deviations for the RT Switch Costs in Experiment 4A as a Function of RSI, Test Block, Task and Decision</i>	105
Table 8.5	<i>Means and Standard Deviations for Corrected Recognition Accuracy in Experiment 4B as a Function of RSI, Test Block, and Transition.....</i>	111
Table 8.6	<i>Means and Standard Deviations for Magnitude Judgement Accuracy in Experiment 4B as a Function of RSI, Test Block, Transition, and Decision</i>	113
Table 8.7	<i>Means and Standard Deviations for the RT Switch Costs in Experiment 4B as a Function of RSI, Test Block, Task and Decision</i>	116
Table 8.8	<i>Means and Standard Deviations for Corrected Recognition Accuracy in Experiment 5A as a Function of RSI, Test Block, and Transition.....</i>	122
Table 8.9	<i>Means and Standard Deviations for Magnitude Judgement Accuracy in Experiment 5A as a Function of RSI, Test Block, Transition and Decision</i>	124
Table 8.10	<i>Means and Standard Deviations for the RT Switch Costs in Experiment 5A as a Function of RSI, Test Block, Task and Decision</i>	129
Table 8.11	<i>Means and Standard Deviations for Corrected Recognition Accuracy in Experiment 5B as a Function of RSI, Test Block, and Transition</i>	132
Table 8.12	<i>Means and Standard Deviations for Magnitude Judgement Accuracy in Experiment 5B as a Function of RSI, Test Block, Transition, and Decision</i>	134
Table 8.13	<i>Means and Standard Deviations for the RT Switch Costs in Experiment 5B as a Function of RSI, Test Block, Task and Decision</i>	138
Table 9.1	<i>Summary of Accuracy and RT Switch Costs for the Recognition and Magnitude Tasks Across Experiments 1 – 5B.....</i>	154
Table A.1	<i>Critical Words from Experiment 1.....</i>	169
Table A.2	<i>Filler and Buffer Words from Experiment 1</i>	170

Table A.3	<i>Critical Words from Experiment 2 and 4A – 5B.....</i>	172
Table A.4	<i>Buffer and Filler Words from Experiment 2 and 4A – 5B.....</i>	177
Table A.5	<i>Critical Words from Experiment 3.....</i>	178
Table A.6	<i>Buffer and Filler Words from Experiment 3.....</i>	181
Table B.1	<i>Means and Standard Deviations for Critical Recognition Hit and False Alarm Rates in Experiment 1 as a Function of Run Condition</i>	183
Table B.2	<i>Means and Standard Deviations for Recognition Hit and False Alarm Rates in Experiment 2 as a Function of Test Block and Transition.....</i>	184
Table B.3	<i>Means and Standard Deviations for Recognition Hit and False Alarm Rates in Experiment 3 as a Function of Run Position and Test Block</i>	185
Table B.4	<i>Means and Standard Deviations for Recognition Hit and False Alarm Rates in Experiment 4A as a Function of RSI, Transition, and Test Block.....</i>	186
Table B.5	<i>Means and Standard Deviations for Recognition Hit and False Alarm Rates in Experiment 4B as a Function of RSI, Transition, and Test Block.....</i>	187
Table B.6	<i>Means and Standard Deviations for Recognition Hit and False Alarm Rates in Experiment 5A as a Function of RSI, Transition, and Test Block.....</i>	188
Table B.7	<i>Means and Standard Deviations for Recognition Hit and False Alarm Rates in Experiment 5B as a Function of RSI, Transition, and Test Block.....</i>	189
Table C.1	<i>Means and Standard Deviations for A' in Experiment 2 as a Function of Test Block, Level of Processing at Study, and Task Transition.....</i>	191
Table D.1	<i>Means and Standard Deviations for Correct RTs in Experiment 4A as a Function of Test Block, Transition, RSI, Task and Decision ..</i>	193
Table D.1	<i>Means and Standard Deviations for Correct RTs in Experiment 4B as a Function of Test Block, Transition, RSI, Task and Decision ..</i>	194
Table D.1	<i>Means and Standard Deviations for Correct RTs in Experiment 5A as a Function of Test Block, Transition, RSI, Task and Decision ..</i>	195

Table D.1 *Means and Standard Deviations for Correct RTs in Experiment 5B
as a Function of Test Block, Transition, RSI, Task and Decision ..196*

Chapter One

Memory

The ability to recall or recognise stimuli, events, or actions that we have experienced before is an essential part of the human condition, and all of us rely on this ability—our memory—across a variety of situations every day. Given its importance to everyday life, considerable thought has been given and much research has been conducted to explore factors that influence memory performance and to understand the mechanisms of memory function. In one of the earliest accounts of memory in modern psychology, James (1890, p. 648) defined memory as “the knowledge of a former state of mind after it has already once dropped from consciousness.” James was careful to differentiate between events that have happened so recently as to have not yet left consciousness (“primary memory”), and those that have left consciousness and need to be reinstated (“memory proper”), effectively purporting a two-stage model of memory (James, p.644, 648). Likewise, in Atkinson and Shiffrin’s (1968) two-stage model of memory, information enters the mind through sensory registers, before flowing into the short-term store—a limited capacity, temporary buffer where information could be manipulated, rehearsed and/or outputted. Information to be held for later retrieval is transferred from the short-term store to the long-term store, where it can be held indefinitely (Atkinson & Shiffrin).

Multi-system models have evolved beyond the Atkinson and Shiffrin (1968) model as experimental evidence has challenged its assumptions. However, it is generally accepted that when information enters the mind/brain, if it is to be successfully recalled or recognised at a later time it must be processed in such a way that it enters long-term memory/storage. This fact remains whether the resulting memories are explicit or implicit in nature, and whether they relate to personal events and history or are of a general, semantic nature. In order to advance our knowledge of memory processes and develop more accurate theories about the mechanisms involved, we need to be able to test what has been successfully transferred to and stored in long-term memory in a controlled experimental fashion.

This thesis is particularly interested in episodic memory. Rather than reflecting memory for an element of knowledge, an episodic memory consists of “multi-feature

representations in which numerous different kinds of information... [are] bound together with the individual's awareness of personal experiences in subjective time" (Schacter & Tulving, 1994, p.28). Tests of episodic memory assess the ability of participants to explicitly recall or recognise things (be they objects, events, or otherwise) that they have previously experienced. Experimental studies of episodic memory have largely relied on a study-test procedure, first presenting item(s) to-be-remembered and later testing memory for these items using recall or recognition memory tests. Conducting a test of memory recall is fairly straightforward—participants are instructed to recall items that they have previously encountered, sometimes aided by cues (cued recall), in the order they were studied (serial recall) or in any order (free recall). Performance is reported as the number of items recalled as a function of study list size (proportion recalled) and can also take into account items produced at test that were not part of the original study material (intrusions).

Recognition Memory

An alternative method of examining episodic memory is to use a recognition memory test. Items are presented to participants, who are instructed to decide if each item was previously studied (i.e., old) or not (i.e., new). Another method involves presenting multiple items at once and instructing participants to select item(s) that were presented at study (forced choice recognition: Hockley, 1992; Shepard, 1967).

Recognition memory tests may solicit participants' subjective experiences during the recognition attempt (Gardiner, 1988; Tulving, 1985) or participants' confidence in their recognition decision (Pollack, 1959; Ratcliff & Starns, 2009). Recognition decisions provide at least two raw data points to experimenters: the response itself and the time taken to make it (reaction/response time; RT). Memory researchers typically emphasise responses themselves, particularly the proportion of old items correctly called 'old' (hit rate) and the proportion of new items incorrectly called 'old' (false alarm rate).

Corrected accuracy—calculated by subtracting the false alarm rate from the hit rate—is often reported instead of raw hit and false alarm rates, as it provides a measure of accuracy that attempts to control for bias towards answering old regardless of item status. Corrected accuracy scores fall within the range of -1 to 1 with 0 representing chance performance and positive values representing above chance performance.

Models of recognition memory.

Recognition memory performance can also be described using the signal detection parameters of discriminability (d' or A') and decision criterion/bias (β or C), that are calculated using hit and false alarm data. That is, making a recognition memory decision can be interpreted as a simple task of discriminating a signal (old memory trace plus noise) from noise (latent activation of new items). Old and new items are represented by two normal distributions of memory strength, with an item's specific position in the distribution governed by characteristics of the encoding event and by forgetting. A recognition decision is made by comparing an item's activation to a decision criterion, such that an old decision is elicited when activation exceeds this point (Banks, 1970; Green & Swets, 1974; Snodgrass & Corwin, 1988; Swets, 1973). In standard signal detection models of recognition memory, old and new distributions have equal variance. Studying a set of words produces a constant increase in familiarity for all items (Yonelinas, Dobbins, Szymanski, Dhaliwal, & King., 1996). However, there is evidence that the variance of the old item distribution is larger than the new item distribution, which has led to the development of unequal variance signal detection (UVSD) models of recognition memory (Slotnick & Dodson, 2005; Slotnick, Klein, Dodson, & Shimamura, 2000; Wixted & Stretch, 2004).

Using signal detection theory to model recognition memory performance leads to a critical assumption—that decisions are based on a single variable of signal strength, often called *familiarity*. Manipulations that affect recognition performance do so by changing the discriminability of targets and lures, and/or changing the decision criterion so participants' responses are more or less conservative (Swets, 1973). That is, signal detection models of recognition memory are *single process* models (although see Wixted 2007a, 2007b for a two process signal detection model). Other single process models of recognition memory assert that recognition decisions are based on assessments of how similar the target is to items stored in memory. Global matching models such as Search of Associative Memory (SAM: Gillund & Shiffrin, 1984), Retrieving Efficiently from Memory (REM: Malmberg, 2008; Shiffrin & Steyvers, 1997), and Theory of Distributed Associative Memory (TODAM 2: Murdock, 1997, 2006), as well as multiple trace models such as Hintzman's (1988) multiple trace simulation model (MINERVA 2) posit that successful recognition relies on matching features of a target probe with features of the relevant trace in memory. During

encoding, a memory trace (sometimes called an ‘image’ or a ‘vector’) of the item’s features is created. When a test probe is presented, its features are logged and compared with traces in memory. When a sufficient number of features match and there are no mismatching features, a target probe can be identified as ‘old’ (Gillund & Shiffrin, 1984; Klein, Shiffrin, & Criss, 2007; Shiffrin & Steyvers, 1997).

More recent matching models such as REM (Retrieving Effectively from Memory) have added extra parameters to account for a wider range of memory phenomena. However, the addition of event traces, time parameters, and other variables to global matching models has not changed the key assumption that memory decisions are made by calculating a single value representing similarity between a target and memory trace. At their core, global matching models and multiple trace models, particularly as they apply to recognition memory, are single process models (Klein et al., 2007; Rotello & Heit, 1999).

By comparison, dual process models of recognition assert that a second process independent of familiarity (or similarity matching) contributes to recognition memory performance (Atkinson & Juola, 1973, 1974; Gardiner, 1988; Wixted, 2007a; Yonelinas, 1994). Where familiarity is considered to be fast and automatic, this second process (often called *recollection*) is considered to be consciously driven and slow (Heathcote, Raymond, & Dunn, 2006; Rugg & Curran, 2007; Yonelinas, 2002; Yonelinas & Jacoby, 1995). To illustrate this distinction, Tulving (1985) drew a distinction between explicit remembering of an event and a more general feeling of knowing an event occurred. Specifically, “even when a person does not *remember* an event, she may *know* something about it” (Tulving, p.6). In a typical study-test recognition paradigm, delaying the final test decreased the proportion of old recognition decisions accompanied by conscious remembering of the study event and increased the proportion of old decisions accompanied by feelings of simply knowing the word had been studied (Tulving). Similarly, some manipulations at encoding/study have been reported to affect the proportion of ‘remember’ responses, but not ‘know’ responses (Gardiner). Dissociations in subjective experiences support the claim that ‘remembering’ and ‘knowing’ can be considered to reflect conceptually and functionally distinct memory processes. That is, ‘remember’ responses have been used as a measure of recollection whereas ‘know’ responses have been used as a measure of familiarity (see Gardiner). However, it is important to acknowledge that alternative explanations of the process(es) underlying ‘remember’ and ‘know’ judgements have

been proposed. A single process model of recognition can account for remember/know dissociations by using two decision criteria reflecting two levels of confidence or memory strength (Donaldson, 1996; Dunn, 2004).

In the dual-process recognition theory proposed by Atkinson and Juola (1973, 1974), memory decisions are made using a combination of judgements of familiarity and controlled memory searches. When a participant is presented with a memory probe, a measure of familiarity is generated and compared to upper and lower decision criterion. If familiarity was above the high criterion or below the low criterion, 'old' or 'new' decisions are made immediately. However, if the value of familiarity fell between the two decision points, participants initiated a controlled search of memory for the target. As such, Atkinson and Juola (1973, 1974) proposed that decisions based solely on familiarity were fast but error-prone, whereas decisions based on an exhaustive search of memory were slower but almost always correct.

Since Atkinson and Juola (1973, 1974) there have been several versions of dual-process models (for a review, see Yonelinas, 2002). Many of these models still include elements of signal detection theory, such as the dual-process signal detection (DPSD) model developed and refined by Yonelinas (1994, 1997, 1999, 2001b). In this model, familiarity is considered to be a continuous variable explained by an equal-variance signal detection model. By contrast, recollection is formalised as a threshold variable, where the study episode (or details about it) are either recalled or not. Familiarity and recollection are considered to act independently of one another. Therefore, the probability of successfully recognising a studied item is the sum of the probability of recollecting the study episode and the probability that familiarity exceeds the decision criterion (Yonelinas, 1994).

Evidence for a DPSD account of recognition memory has been reported using a number of experimental procedures. For example, Yonelinas (1994) showed that although signal detection models predict symmetrical receiver operating characteristics (ROCs), recognition memory ROCs are often asymmetrical. Specifically, they are skewed up and to the left of the plot, indicating that high confidence responses—those that Yonelinas would argue represent decisions based on recollection processes—deviate from the signal detection model (for a strength-based explanation of asymmetric ROCs, see Wais, Wixted, Hopkins & Squire, 2006). Level-of-processing manipulations (Craik & Lockhart, 1972) have produced ROCs consistent with the DPSD model of recognition memory (Yonelinas, et al., 1996), as have divided attention manipulations

(Yonelinas, 2001a). Likewise, process dissociation procedures (Jacoby, 1991, 1998) have also provided evidence for dual-process accounts of recognition memory (e.g., Jacoby, 1999; Jacoby, Debnar & Hay, 2001; Yonelinas, 2001a; Yonelinas & Jacoby, 1995). Critically, Yonelinas (2001a; Yonelinas, Aly, Wang & Koen, 2010) has reported that remember/know procedures, process dissociation procedures, and ROC functions have produced very similar estimates of recollection and familiarity, indicating that they were measuring the same two underlying processes (a strong sign of convergent validity).

Similar to the DPSD model, the source activation confusion (SAC) model (Diana, Reder, Arndt & Park, 2006; Park, Reder, & Dickison, 2005; Reder et al., 2000;) posits that recollection is all-or-nothing (threshold) and familiarity is continuous. The SAC model states that studying a word creates a memory of the study event (an episode node) which is associated with information about the experimental environment (experimental context node), conceptual information about the word itself (concept node) and contextual information specific to that trial (specific context node). A recognition memory test activates the experimental context node, and the presentation of a test word activates its concept node (and sometimes the specific context node). This activation spreads throughout the network, including to the episode node. Responses driven by recollection of the study event are the product of episodic node activation and are all-or-nothing. By contrast, responses driven by familiarity are made when the activation of the concept node exceeds a set criterion (Diana et al., 2006). As the resting activation of a concept node is governed by frequency and recency of exposure, some items (i.e., high frequency words) are more likely to produce old responses even when they were not studied. Responses based on familiarity are more prone to error, whereas recollection is considered highly accurate (although see Park et al., 2005 for an account of false recollection).

In an alternative conceptualisation of dual process recognition, the multi-process unequal variance signal detection model (UVSD; Wixted & Stretch; Wixted, 2007a, 2007b) claims that, like familiarity, recollection is a graded phenomenon and can be partial in nature (i.e., Dodson, Holland & Shimamura, 1998). The USVD model posits that two processes (at least) underlie recognition performance and that they are all expressed as continuous variables in a signal detection model. An activation value is calculated by summing recollection and familiarity strength, and this activation value is compared against a criterion to produce a recognition decision. Critically, the controlled

recollection process still contributes to recognition performance even though decisions are modelled using a signal detection distribution.

It is, however, important to acknowledge that there are situations where a second process such as recollection is not needed to explain memory performance. One example of this is face recognition memory, where the discrimination of old and new faces as well as the false recognition of conjunction faces (a recombination of two studied faces to create a new face) can be adequately explained using a single process (familiarity-based) account of recognition memory (Jones & Bartlett, 2009; Jones, Bartlett & Wade, 2006). By contrast, similar experiments using words have shown that recollection does play a role in recognition performance (Jones & Jacoby, 2001). That is, the use or role of a consciously controlled recollection process may also depend on what one is trying to recognise: it is bound by experimental conditions.

Experimental Manipulations that Affect Recognition Performance

Participants' performance on recognition memory tests can be affected by a number of different manipulations at encoding (study) and retrieval (test). Encoding manipulations affect the strength, integrity, or elaborateness of the memory trace, leading to differences in memory performance at retrieval. For example, increasing the number of items to be studied decreases accuracy (Pollack, 1959), as does dividing attention with a secondary task (Craik, Govoni, Naveh-Benjamin, & Anderson, 1996). By contrast, increasing the time spent studying an item (McClelland & Chappell, 1998; Ruiz, Soler, & Dasi, 2004) or the number of times an item is presented (Jones, 2005; McClelland & Chappell, 1998) increases memory strength and thus accuracy. Characteristics of the items to be studied can also affect memory. For example, high frequency words are more likely to be falsely recognised as 'old' than low frequency words (Heathcote, Ditton, & Mitchell, 2006; Joordens & Hockley, 2000).

A particularly effective encoding manipulation that will be used in this thesis is the level-of-processing manipulation (Craik & Lockhart, 1972). The level-of-processing framework proposed that the strength of a memory was determined by how an item is encoded. Optimal strengthening of a memory trace was attained by processing an item beyond its physical characteristics to its semantic meaning (Craik & Lockhart). This hypothesis was tested by Craik and Tulving (1975), who manipulated level of processing at study. During an incidental study phase, participants made deep semantic (e.g., is the word a mammal?) or shallow perceptual (e.g., is the word presented in

upper-case letters?) encoding judgements. When participants' memory for these words was tested, a robust level-of-processing effect (accuracy deeply encoded words > accuracy shallowly encoded word) was observed across recognition and recall tests, for incidental and intentional learning, in the laboratory and in a classroom setting, with and without monetary incentive.¹ Although research soon established that memory performance was actually optimised by matching retrieval conditions and processes at study and test (Morris, Bransford, & Franks, 1977; Roediger, Gallo, & Geraci, 2002; Weldon & Roediger, 1987), the level-of-processing manipulation has remained an often-used, standard manipulation in memory research.

A variety of manipulations at retrieval also affect memory performance. Recognition performance can be affected by encouraging conservative or lenient response biases (Curran, DeBuse & Leynes, 2007) or by constraining when participants can make their responses, thus biasing processes available to make the memory decision (Hintzman & Curran, 1994; Odegard, Koen, & Gama, 2008; Rotello, Macmillan, & Van Tassel, 2000). Delaying the final test allows for memory traces to decay, thus decreasing accuracy (Hockly & Consoli, 1999; Tulving, 1985). Another method of decreasing memory accuracy is to divide participants' attention; instructing participants to perform the recognition task and a second task simultaneously (Hicks & Marsh, 2000; Jacoby, 1991; Jones & Jacoby, 2001; Knott & Dewhurst, 2007a, 2007b).

Divided attention manipulations (full vs divided attention) at retrieval are relevant to the present research because of their similarity to the attention shifts required in task switching, the topic of this thesis. Divided attention at retrieval has the potential to demand considerable attentional resources as participants are required to monitor multiple tasks and resolve instances of interference. It requires that participants maintain two sets of instructions in working memory—one for each task—so that the appropriate response can be made when a stimulus is presented. For example, Jacoby's (1991) test phase required participants in a divided attention condition to identify visually-presented words as old or new, while also monitoring aurally-presented digit strings for a target sequence. Through process dissociation procedures, Jacoby (1991) concluded that divided attention at retrieval decreased recognition performance overall, but was particularly disrupting to recollection processes.

¹ An additional phonemic level of processing (e.g., 'does the word rhyme with weight?') was also used by Craik and Tulving (1975), producing better performance than the perceptual encoding conditions but worse performance than the semantic encoding condition.

A number of studies have confirmed Jacoby's (1991) observation of decreased recognition accuracy under divided attention and are in general agreement that these effects are either limited to or much more strongly observed for recollection than for familiarity (Hicks & Marsh, 2000; Jones & Jacoby, 2001; Knott & Dewhurst, 2007a). Although Hicks and Marsh reported significant divided attention effects for secondary tasks ranging from adding successive digit pairs to random number and letter generation, there is also evidence that the complexity of the divided attention task or the degree of interference (or similarity) a divided attention task has with the primary recognition task is also important (Fernandes & Moscovitch, 2000; Skinner & Fernandes, 2008). Indeed, observing a divided attention effect at retrieval is by no means guaranteed (see Craik et al., 1996; Naveh-Benjamin, Craik, Gavrilescu & Anderson, 2000).

Chapter Two

Interruptions and Memory

While divided attention experiments (as well as other dual-task paradigms) involve the simultaneous monitoring of two sets of stimuli, research on interruptions has investigated the impact of a temporary shift away from a primary task to complete a secondary task. The onset of the interruption task cues the temporary suspension of the primary task, and returning to the primary task after an interruption can take time and cognitive effort, resulting in poorer performance (Altmann & Trafton, 2002, 2007; Gillie & Broadbent, 1989; Trafton, Altmann, Brock, & Mintz, 2003). There are a small set of experiments that have directly examined the effect of interruptions on episodic memory (see Edwards & Gronlund, 1998; Gillie & Broadbent; Jones, Burson, & Aronsen, 2011). However, Gillie and Broadbent and Edwards and Gronlund used an interruption paradigm quite unlike standard episodic memory paradigms and obtained data not easily reconciled with accuracy and RT measures typical in the episodic memory literature.

Gillie and Broadbent (1989) trained participants to navigate a computer-generated town with the goal of collecting a list of five or seven objects from a variety of locations in a fixed order or at the participants' discretion. Twelve trials were completed by participants. Six trials were interrupted half way through the list so that participants could perform a secondary task (simple math, free recall memory test, or math with coded digits) and six trials were completed without interruption. Gillie and Broadbent compared the average time to collect an item and the number of participant requests for help (i.e., to clarify what items were to be collected) for the first and second halves of each trial as a function of interruption condition (no interruption, interruption). Gillie and Broadbent asserted that prior to the interruption (i.e., in the first half of the trial) both groups should perform the same. By contrast, for the second half of trials the mean time should be longer and the mean requests for help should be greater for interruption trials than those for no-interruption trials. That is, Gillie and Broadbent hypothesised that interruption effects would be illustrated by an interaction of trial type (interruption, no-interruption) and trial half (first half, second half).

When the interruption was a simple math task, no interruption effects were observed regardless of interruption length (30 seconds or 165 seconds). Likewise, when the interruption task required participants to study and then recall a short list ($N = 32$) of words, although the data showed a pattern consistent with an interruption effect, it failed to reach statistical significance. Finally, Gillie and Broadbent (1989) examined the effect of a more complex and cognitively demanding math interruption task. During the interruption, participants solved math problems where the digits were replaced with letters (e.g., $EF - DG = ?$) by using a supplied cypher. Gillie and Broadbent argued that these actions required substantial mental processing and placed working memory under significant load—much more than the earlier math and recall tasks. Consistent with an interruption effect, the average collection time for the second half of interruption trials was significantly slower than that for the first half (before the interruption) or either half of the no-interruption trials. Following an interruption, participants took longer to recall from memory the items they needed to collect and then locate and select them from within the town. No interruption effect was observed in request-for-help data, but this null effect may have been due to floor effects. At the very least, Gillie and Broadbent showed that when the interruption task was sufficiently demanding of cognitive resources (such that there is little time to rehearse information from the primary task) access to episodic memory was slowed.

Edwards and Gronlund (1998) used a similar method to Gillie and Broadbent (1989), training participants to navigate their way through a ‘town’ on a computer screen to collect a list of 10 items from various locations in a specified order. Six trials were completed by participants. The collection order was manipulated between participants, such that one group collected the same items in the same order for every trial (fixed order) and the other group collected the same items in a different order for every trial (arbitrary order). The sixth trial was different from the earlier trials. After collecting the fifth item, participants performed an interruption task for three minutes. The interruption task was either similar to the primary task (planning a route through the town to collect 15 new items or rearrange anagrams) or dissimilar to the primary task (performing addition on a matrix of numbers). Prior to resuming the primary task, participants answered memory probes about their location (or in Experiment 3, item collected) prior to the interruption as well as the next item they were meant to collect.²

² Experiment 1 also administered these probes as recognition probes, but this manipulation had no impact on performance.

It is important to note that these probe questions were never asked in the absence of an interruption and, as such, Edwards and Gronlund did not have a baseline measure of performance. Instead, they compared performance on probe questions to chance, which was assessed as 1/19 for the location probe and 1/10 for the item probes, based on the number of possible locations or items. After completing every trial, participants were instructed to recall (in order) the items they had collected.

Across all three experiments reported by Edwards and Gronlund (1998), participants recalled a mean of 8/10 items in the correct order. In Experiment 1, participants in the fixed order condition answered the post-interruption probe questions more accurately ($M = 0.59$ for location probe, $M = 0.52$ for item probe) than participants in the arbitrary order condition ($M = 0.44$ for location probe, $M = 0.28$ for item probe). Accuracy was consistently better when the interruption task was different to the primary task, compared to that when the interruption task was similar to the primary task—there was an interruption similarity effect. Together with Gillie and Broadbent (1989), Edwards and Gronlund's interruption similarity effects are consistent with the divided attention literature which emphasises the importance of secondary task complexity and similarity to the primary task for observing divided attention effects (Fernandes & Moscovitch, 2000; Skinner & Fernandes, 2008). Even in the absence of a baseline (no-interruption) condition, Edwards and Gronlund's findings suggested that participants' episodic memory can be affected by an interruption. Participants found it particularly difficult to remember what item they were meant to collect next following an interruption that required them to plan an item search through a village or to rearrange anagrams.

Interruption effects on episodic memory may have been produced as an unintended by-product in a recognition memory experiment by Jones and Atchley (2006). Jones and Atchley investigated the role of familiarity and recollection in the production and prevention of memory errors in a continuous recognition memory paradigm. A continuous recognition paradigm does not separate study and test phases. Instead, words are presented twice in a single list. Recognition memory decisions are elicited on every trial, producing a 'new' response for the first presentation of a word and an 'old' response for the second presentation of a word. The word stimuli in the experiment were compound words (e.g., blackmail, jailbird) that could be recombined to produce new 'conjunction' lures (e.g., blackbird). Along with identifying old and new words, participants were instructed to identify conjunction lures and in such cases

to verbally report one of the parent words that were used to create the conjunction lure. This recall-to-reject strategy reflects the successful use of recollection to avoid saying old to a conjunction lures.

In Experiments 2A and 2B, Jones and Atchley (2006) manipulated the lag (number of intervening trials) between the first and second presentation of a word and measured how the accuracy of 'old' judgements was affected by the length of this lag. Each experiment used three lags: a five-trial lag common to both experiments and two lags that were different for each experiment. Jones (personal communication, 2005) noted that accuracy for old words at lag five was significantly lower in Experiment 2A than that in 2B (0.66 and 0.82 respectively), and suggested that this difference could be due to a slight difference in the construction of the test lists. For Experiment 2A, a conjunction lure was occasionally presented during the five-trial lag, but not for Experiment 2B. Therefore, in Experiment 2A participants were sometimes interrupted from making recognition decisions to explicitly recall an earlier parent word, and this interruption may have had a negative impact on recognition performance.

Jones et al. (2011) tested this hypothesis more rigorously in two experiments. The size and persistence of interruption effects on recognition memory performance were examined using a continuous recognition memory task that was periodically interrupted (~ 6% of all trials) with a secondary episodic or semantic memory task. Stimuli for the recognition memory task were the same compound words used in Jones and Atchley (2006) excluding conjunctions lures (that is, blackmail and jailbird were used, but not blackbird). Stimuli for the interruption task were Xes presented in one of five colours. On each recognition trial, participants were instructed to judge whether the word had been presented earlier in the experiment (i.e., it was 'old') or not (i.e., it was 'new'). The semantic interruption task required participants to identify the colour of the X stimuli and the recall interruption task required participants to recall the word presented on the preceding trial. Dispersed throughout each experiment were 'runs' of seven trials, illustrated in Figure 2.1. A run was either interrupted (containing a single interruption on Trial 5 or 6) or uninterrupted (containing only recognition memory trials). Of primary interest was performance on Trial 7 of a run, which as a consequence of experimental design either followed an interruption by one or two trials (I+1 and I+2 conditions respectively) or followed only recognition trials (control conditions). This critical trial was either old (word presented on Trial 1) or new (word not previously presented).

Trial	I + 1	I + 2	Control
1	wheelchair	wheelchair	wheelchair
2	drawstring	drawstring	drawstring
3	dragonfly	dragonfly	dragonfly
4	eyesore	eyesore	eyesore
5	fireplace	XXXXXXXX	fireplace
6	XXXXXXXX	snakeskin	fireplace
7*	wheelchair	wheelchair	wheelchair

Figure 2.1. The design of the seven-trial runs used in Jones et al. (2011). Trial 7 (marked by *) is the critical trial. The Xes were presented in five different colours and were used for the interruption task. In the runs represented here, the critical trial is 'old', as the word was presented on Trial 1.

Experiment 1A and 1B examined recognition memory performance when it immediately followed an interruption (I+1 condition) or not (control condition). Jones et al. (2011) reported significantly lower accuracy in the I + 1 condition compared to that in the control condition (7 – 14% lower). Participants' ability to discriminate between old and new words was impaired following an interruption, regardless of whether the interruption task was semantic or episodic. An interruption effect was also reported for RT performance. As illustrated in Figure 2.2, RTs for correct recognition trials were significantly slower following an interruption than that for the control condition where no interruption was experienced. This pattern of worse performance following an interruption held when incorrect responses were included in the analysis. The interruption effect was significantly larger for the recall interruption task than that for the colour naming interruption task.

In their second experiment, Jones et al. compared recognition performance across three conditions—control, I+1, and I+2—to investigate whether interruption effects persisted beyond the trial immediately following an interruption (the I+1 trial). In Experiment 2, critical trials were always old words. Although the interruption effect on recognition memory hits (correctly identifying old words) was significant, this effect

was isolated to the I+1 trial. By contrast, interruption effects on recognition RTs were significant for I+1 (~ 315 ms) and I+2 (~ 82 ms) conditions. Recognition memory RTs were slower for the I+1 condition than that for the I+2 condition and both interruption conditions were slower than the control condition. That is, there was a gradual recovery from the interruption. Unlike the Experiments 1A and 1B, the type of interruption task did not affect the size of the RT interruption effect. This may have been a consequence of Experiment 2 using only ‘old’ words for critical trials, as new words showed larger task effects in Experiment 1A and 1B. Combined, the experiments reported by Jones et al. provided good evidence that interrupting a continuous recognition test to perform a brief, unrelated task had a negative effect on recognition performance. Immediately following an interruption, recognition memory decisions were slower and less accurate. Furthermore, recovery from an interruption was not immediate, as RTs remained slower than control conditions two trials after the interruption.

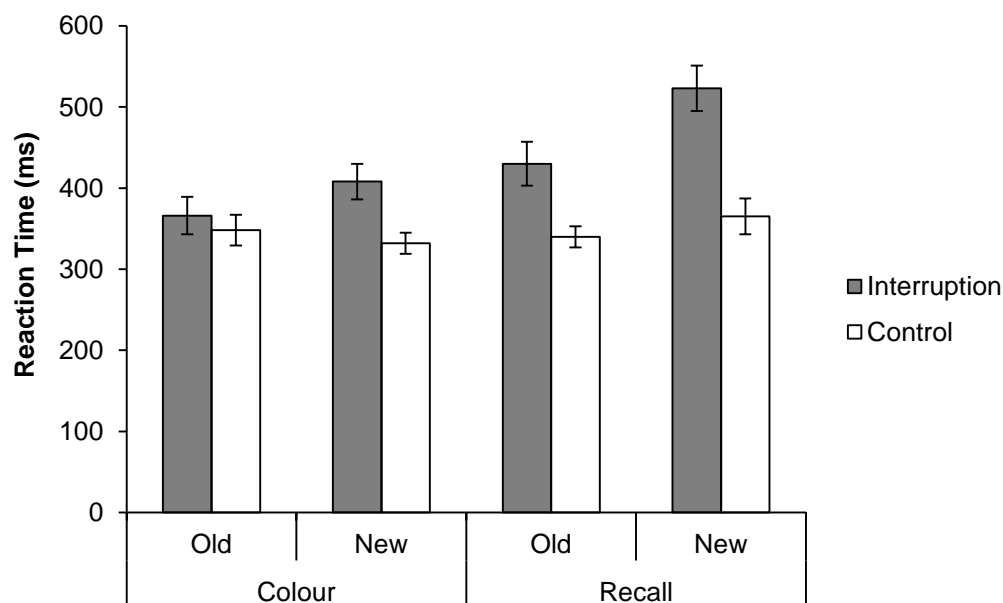


Figure 2.2. Performance on critical recognition memory trials collapsed across Experiment 1A and 1B of Jones et al. (2011). RTs are presented as a function of interruption task, word status, and run type (interruption or control). Recognition decisions were made after a 1250-ms response delay.

Chapter Three

Task Switching

The research presented in this thesis was borne out of the findings of Jones et al. (2011), where recognition memory performance was slower and less accurate after an interruption. Jones et al. (2011) proposed that the interruptions to the continuous recognition task could be interpreted as infrequent task switches. The principle reason for a task switching interpretation of Jones et al. is methodological. In a typical interruption experiment, an on-going process is interrupted before it can be completed and it must be resumed at the end of the interruption. Participants have to suspend a goal they are in the process of completing and upon returning to this primary task they must remember what they were doing and what they were planning to do next (Altmann & Trafton, 2002, 2007; Edwards & Gronlund, 1998; Gillie & Broadbent, 1989; Trafton, Altmann, Brock & Mintz, 2003). By contrast, each recognition trial in Jones et al. was a discrete encoding and retrieval event and was completed *before* the onset of the interruption task. After the interruption, participants did not have to recall what process they were involved in or what they had planned to do next. Instead, they had to adjust the goals and rules they were using from those appropriate for the interruption task (i.e., identify colour, recall word) to those appropriate for the recognition task (i.e., is the word ‘old’?). Thus, successful performance in Jones et al.’s continuous recognition task required participants to occasionally switch between two different sets of task rules.

Task switching research is based on the same basic methodology as described above—participants complete sequences of two or more tasks that have different goals and rules (often called *task sets*). Each trial requires that the appropriate task set is used to produce the correct response to a stimulus. A trial (and thus, task) is completed before the next trial begins, which may or may not require switching task sets. Switching between task sets is assumed to take time and cognitive resources, and to negatively affect performance. That is, compared to task repetition trials (where trial N and N-1 are the same task), participants are slower and often less accurate on task switch trials (where trial N and N-1 are different tasks). This phenomenon is called a *switch effect* (Monsell, 2003; Schneider, 2007; Vandierendonck, Liefvooghe & Verbruggen, 2010).

Origin of Task Switching Research and Theory

The origin of task switching research can be traced to Jersild (1927), who examined task switching performance by comparing the time taken to complete a block of a single task (pure block, AAAAAA) to the time taken to complete a block of two tasks that alternated (mixed blocks, ABABAB). In one experiment, Jersild compared the completion times for pure and mixed blocks of an addition task (adding 6 to the target digit) and a subtraction task (subtracting 3 from the target digit). Completion times for mixed blocks were reliably slower than pure blocks, a phenomenon Jersild called a *shift cost* but today is called a *mixing cost*. Mixing costs were observed when participants alternated between tasks that used the same stimulus sets but not when used different stimulus sets (e.g., when participants switched between reporting opposites of words and performing math operations on numbers). In fact, Jersild reported *faster* completion times for mixed lists compared to pure lists when two tasks used different stimulus sets. Jersild concluded that mixing costs were driven by the ambiguity of stimuli—that they could elicit responses for more than one task—and that resolving this ambiguity is what led to longer completion times.

Spector and Biederman (1976) further examined the importance of stimulus ambiguity. For tasks using different stimulus sets, the advantage for mixed lists compared to pure lists was eliminated when stimuli were presented one-at-a-time compared to all-at-once on a piece of paper (although there was still no significant mixing cost). Spector and Biederman agreed with Jersild's (1927) explanation of this result: presenting task-unambiguous stimuli on a single piece of paper allowed participants to preview upcoming stimuli and plan responses, minimising or eliminating switch effects. Furthermore, Spector and Biederman also reported that presenting task cues (i.e., '+' for an addition task) with stimuli reduced completion times for mixed lists when tasks shared stimulus sets. Thus, manipulations that reduced stimulus ambiguity and eliminated the need for participants to remember the task sequence produced smaller mixing costs, leading Spector and Biederman to conclude (consistent with Jersild) that switch effects were largely driven by the need to resolve conflict when selecting the appropriate task to perform.

The methods used by Jersild (1927) and Spector and Biederman (1976), as well as their interpretation of their results, strongly shaped the development of the task switching literature. Subsequent task switching research has used simple tasks that share

a stimulus set, such as performing mathematic operations on digit stimuli (i.e., Baddeley, Chincotta & Adlam, 2001; Emerson & Miyake, 2003) or making decisions about size, identity or colour of stimuli (Altmann, 2006; Mayr & Kliegl, 2003; Miyake, Emerson, Padilla & Ahn, 2004; Posse, Waszak & Hommel, 2006; Rubin & Meiran, 2005). In using the same stimulus set for both (or all) tasks in an experiment, ambiguity (or stimulus-response (S-R) overlap) is maintained—a factor considered essential to some task switching theories (see Allport, Styles & Hsieh, 1994; Allport & Wylie, 2000). Post Jersild and Spector and Biederman, there is only a small set of task switching research that has used different stimuli for each task or tasks with limited S-R overlap (e.g., Allport et al., 1994; Altmann, 2002; Meiran, 2000a; Rogers & Monsell, 1995), however they have either failed to produce switch effects (Allport et al., 1994) or reported switch effects of smaller magnitude (Allport et al.; Meiran 2000a; Rogers & Monsell). However, the fact that costs are still observed in these situations suggests that task switching effects are more than just an index of conflict/interference resolution.

Methods of Investigating Task Switching

The earliest method used to examine task switching was the mixed and pure list method (Jersild, 1927) described earlier in this introduction. Comparisons of mixed and pure lists have also been used in more modern task switching research (e.g., Allport et al., 1994; Baddeley et al., 2001; Emerson & Miyake, 2003). However, this method is not without limitations, most notably that participants are under the additional cognitive load of maintaining two task sets in mixed lists but not in pure lists (Rogers & Monsell, 1995). As a consequence, differences in performance between mixed and pure blocks represent a mixing cost component in addition to a task switching component. With the resurgence in task switching research in the mid-1990's, new methods of examining task switching performance were developed in large part to eliminate such limitations and to produce better measures of cognitive control (Meiran, 1996; Rogers & Monsell). These new methods were the alternating runs paradigm (Rogers and Monsell), the explicit task cuing paradigm (Meiran, 1996), and more recently, the voluntary switching paradigm (Arrington & Logan, 2004, 2005). Generally speaking, these paradigms reflect increasing attempts to control the processes contributing to switching performance and when they can be engaged. All of these experimental paradigms have positive and negative aspects, and the choice of method will rely both on the purpose of an experiment and the theoretical outlook of the experimenter.

Alternating runs paradigm.

Since its inception (Rogers & Monsell, 1995), the alternating runs paradigm has been used by many different researchers to investigate switch effects (e.g., Altmann, 2002, 2006, 2007; Brown, Lehmann & Poboka, 2006; de Jong, 2000, 2001; Hunt & Klein, 2002; Koch, 2003, 2005; Lindsen & de Jong, 2010; Minear & Shah, 2008; Monsell, Sumner & Waters, 2003; Posse et al., 2006). In this paradigm, participants complete a set number of trials (a *run*) of one task before switching tasks and completing a run of the alternative task (i.e., AA BB AA BB). Although a 2-trial run length is most often used, experimenters have used longer run-lengths or multiple run lengths to further understand the cognitive processes involved in switching tasks (e.g., Altmann, 2002; Altmann & Gray, 2002; Kieffaber & Hetrick, 2005; Monsell et al., 2003). Regardless of run length, task switching effects are calculated by comparing RT and accuracy performance for the first trial in a run (task switch) to the subsequent trial(s) in a run (task repetition).

As task progression is entirely predictable, participants can use their memory of the task sequence to anticipate what task they will perform next. Similar to the mixed versus pure blocks paradigm, participants can perform the tasks without explicit trial-by-trial instructions. However, experimenters have often used task cues in addition to this inherent predictability, providing participants with an external reminder of the upcoming tasks' identity and allowing them to reorient themselves if they lose track of the task order. For example, Rogers and Monsell (1995) used spatial location—a clockwise progression through four quadrants of a square—as task cues. Alternatively, Koch (2003) used geometric frames (square or diamond) as task cues—after a response to trial N was made, the task-relevant frame for the N+1 trial was presented prior to stimulus onset. Regardless of whether external cues are used or not, the key characteristic of the alternating runs paradigm is the constant, predictable pattern of switch and repetition trials that allows for RT and accuracy switch effects to be easily calculated at the level of individual trials.

Rogers and Monsell (1995) claimed that one of the primary advantages of using an alternating runs paradigm was that the cognitive load of maintaining task sets was kept constant. Unlike the mixed versus pure list paradigm, performance on task switch and task repetition trials was assessed within one block. Without doubt, the paradigm represented a significant contribution to task switching methodology and likely helped

revive interest in task switching as a method of examining cognitive control. There are researchers who maintain that the alternating runs paradigm is not without faults (see Altmann, 2002 for one account). Some have suggested that task switching effects that are calculated using this method are artificially inflated by ‘restart costs’—a slowing of performance at the beginning of a run regardless of whether a run uses the same or different task from the previous run (Altmann & Gray, 2002; Allport & Wylie, 2000; Gopher, Armony & Greenshpan, 2000). However, as reports of this ‘restart cost’ have generally been limited to experiments where instructions or cues are presented prior to the start of each run (a variation called an intermittent instruction paradigm), it is debatable whether this is a large concern to the alternating runs paradigm as a whole.

Explicit task cuing paradigm.

First introduced by Meiran (1996), the explicit task cuing paradigm differs from the alternating runs paradigm in that the progression of tasks is *not* predictable. Instead, cues are presented on each trial that instruct participants as to what task to perform. The sequence of task cues (and thus tasks) across a block/experiment is pseudo-random, with ‘task switch’ and ‘task repetition’ trials identified based on the identity of the N-1 task. Switch effects are calculated by comparing performance for trials identified as task switches to performance for trials identified as task repetitions. The cues can vary from more transparent to less transparent: from providing responses (e.g., ‘more/less’), a description of the task (e.g., ‘magnitude’), a letter associated with the task (e.g., ‘m’), or a symbol or colour that needs to be interpreted (e.g., the colour blue cues the magnitude task). Researchers have even utilised cues that inform participants about the transition (i.e., switch cue, repeat cue) rather than task identity (Schneider & Logan, 2007). As such, the cognitive effort and/or processes involved in interpreting cues and then switching tasks can vary, which in turn affects the nature and magnitude of the switch effect (Emerson & Miyake, 2003; Miyake et al., 2004).

In the explicit task cuing paradigm, cues and stimuli can be presented simultaneously, or the cue can precede the stimulus (that is, cues can vary in temporal location). Three timing parameters—illustrated in Figure 3.1—are important when the cue is presented prior to the stimulus. The duration between the trial N response and trial N + 1 cue is called the response-to-cue interval (RCI). This interval represents a

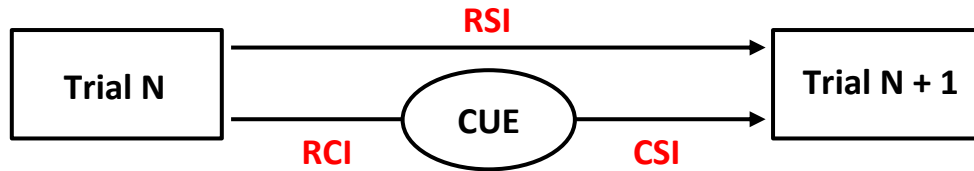


Figure 3.1. An illustration of the timing parameters in the explicit task cuing paradigm. RSI = response to stimulus interval. RCI = response to cue interval. CSI = cue to stimulus interval.

relatively passive period of time where participants do not know what task they will be performing next. The time between the presentation of a cue and the stimulus of trial N+ 1 represents the cue-to-stimulus interval (CSI). This interval is of critical importance as participants have knowledge of the upcoming task and therefore the opportunity to begin to switch task sets (Meiran, 1996). Finally, both the RCI and the CSI combined represent the response-to-stimulus interval (RSI), representing the total time between trials, and (some would argue) the time available for the trial N task set to decay passively.

The explicit task cuing paradigm arguably provides benefits over alternating runs and mixed versus pure list paradigms with its ability to disentangle active and passive processes contributing to switch effects (Mayr & Keele, 2000; Meiran, 1996). Indeed, the flexibility and control over timing in this paradigm has inevitably contributed to its popularity in task switching research today. However, this paradigm is not without limitations. Experimental evidence has suggested that the presentation of a cue does not mean that it will be utilised straight away (Verbruggen, Liefvooghe, Vandierendonck & Demanet, 2007). Furthermore, Mayr and Kliegl (2003) and Logan and Bundesen (2003) noted that cue switches and task switches are confounded in the explicit task cuing paradigm. That is, whenever the current (‘online’) task changed the cue changed as well, and therefore the effects of these two changes could not be disentangled. When this limitation is controlled for, the observed RT switch effects are somewhat smaller (see Logan & Bundesen, 2003; Mayr, 2006; Mayr & Kliegl, 2003; Monsell & Mizon, 2006; Schneider & Logan, 2011).

Voluntary task switching paradigm.

A more recent contribution to task switching methodology is the voluntary task switching paradigm (Arrington & Logan, 2004). This paradigm allows participants to

choose what task they want to perform on any given trial, with the proviso that they perform each task equally often. As such, it is argued that all processes associated with switching tasks are initiated by the participant instead of by cues. Responses for each task are mapped onto different keys and task repetitions and switches are identified post hoc. Critically, Arrington and Logan (2004) have shown that, although it is the participant who makes a choice to switch tasks, significant switch effects are still observed. The voluntary task switching has primarily been used by Arrington and colleagues (e.g., Arrington & Logan, 2004, 2005; Arrington, Weaver & Pauker, 2010), but other researchers have used the paradigm (e.g., Forstmann, Brass, Koch & von Cramon, 2006; Mayr & Bell, 2006). Again, this task switching paradigm has its own limitations: although Arrington and Logan (2004) argued that switch effects reported using this paradigm reflected endogenous, participant driven task switching, more recent evidence has shown that task switching behaviour in this paradigm can be influenced by external factors (see Demanet, Verbruggen, Liefoghe & Vandierendonck, 2010).

Key Task Switching Effects and Manipulations

New task switching paradigms were developed largely in an attempt to control the processes involved in switching task sets and to discriminate between theoretical explanations of task switching effects. However, equally important was the systematic manipulation of other experimental variables (e.g., timing of trials). Over the history of task switching research, researchers have reported a large number of experimental manipulations that have produced unique changes in performance (see Vandierendonck et al., 2010 for a comprehensive review). This section summarises the several key task switching effects: preparation effects, practice effects, and run length effects. These effects are highlighted here because of their relevance to the experiments that will be presented in this thesis. Preparation effects and practice effects are the product of longer RSIs/CSIs and experience with switching tasks, respectively. They both produce decreases in the size of RT switch effects. By contrast, run length effects are primarily focused on changes in performance across a run of repetition trials for a single task.

Preparation effects.

Rogers and Monsell (1995) proposed that with sufficient time (and knowledge of the upcoming task), participants could begin to switch task sets in advance of stimulus onset. Under these circumstances the appropriate task set would be active

when the target stimulus was presented, and thus responses could be made more quickly. That is, increasing the RSI (or CSI in the explicit cuing paradigm) should decrease the size of the RT switch cost (switch RT – repetition RT) and accuracy switch cost (repetition accuracy – switch accuracy). Rogers and Monsell tested this hypothesis in a series of experiments that utilised a two-trial alternating runs paradigm and five different RSIs ranging from very short (200 ms) to substantially longer (1200 ms). Participants completed multiple blocks of trials over two days, switching between making parity decisions (is the digit odd or even?) and consonant/vowel distinctions (is the letter a consonant or a vowel?) to compound stimuli (e.g., ‘A9’). As can be seen in Figure 3.2, when RSI varied randomly within a block (Experiment 2), the magnitude of the switch cost was unaffected by RSI. However, when RSI was varied across blocks (Experiment 3 & 4), the RT (and accuracy) switch cost decreased with longer RSIs up to 600 ms, but showed no further decrease beyond that RSI length (there was a *residual switch cost*). Although the RT switch costs were smaller when stimuli contained attributes relevant to only one task (low S-R overlap: ‘#9’) compared to that when stimuli contained attributes relevant to both tasks (high S-R overlap: ‘A9’), both conditions showed the same trend of decreasing RT switch costs as RSI lengthened. Rogers and Monsell concluded that when provided with additional time prior to stimulus onset, participants would use that time to change (or, reconfigure) task sets in advance, and called the resulting decrease in RT and accuracy switch costs a ‘preparation effect’.

Meiran (1996) was quick to point out that there was an alternative explanation for Rogers and Monsell’s (1995) preparation effect: increasing the RSI meant that the currently irrelevant (‘offline’) task set had longer to decay and, therefore, preparation effects could be the product of less interference. To address this alternative explanation, Meiran (1996) and Meiran, Chorev and Sapir (2000) used the explicit task cuing paradigm, systematically varying RCI and CSI to assess the contributions of passive decay and anticipatory task switching to preparation effects. On each trial, participants were presented with a happy face (a circle with eyes and smile) in one of four quadrants in a square frame, with task cues (arrows) positioned to the left/right or top/bottom of the frame. The two tasks were to identify the location of the face horizontally (left or right) or vertically (up or down) and respond using spatially congruent keys on a numeric keypad. When the RSI was kept constant, Meiran (1996) reported that the RT switch cost was significantly smaller at a long (1716 ms) CSI compared to that at a

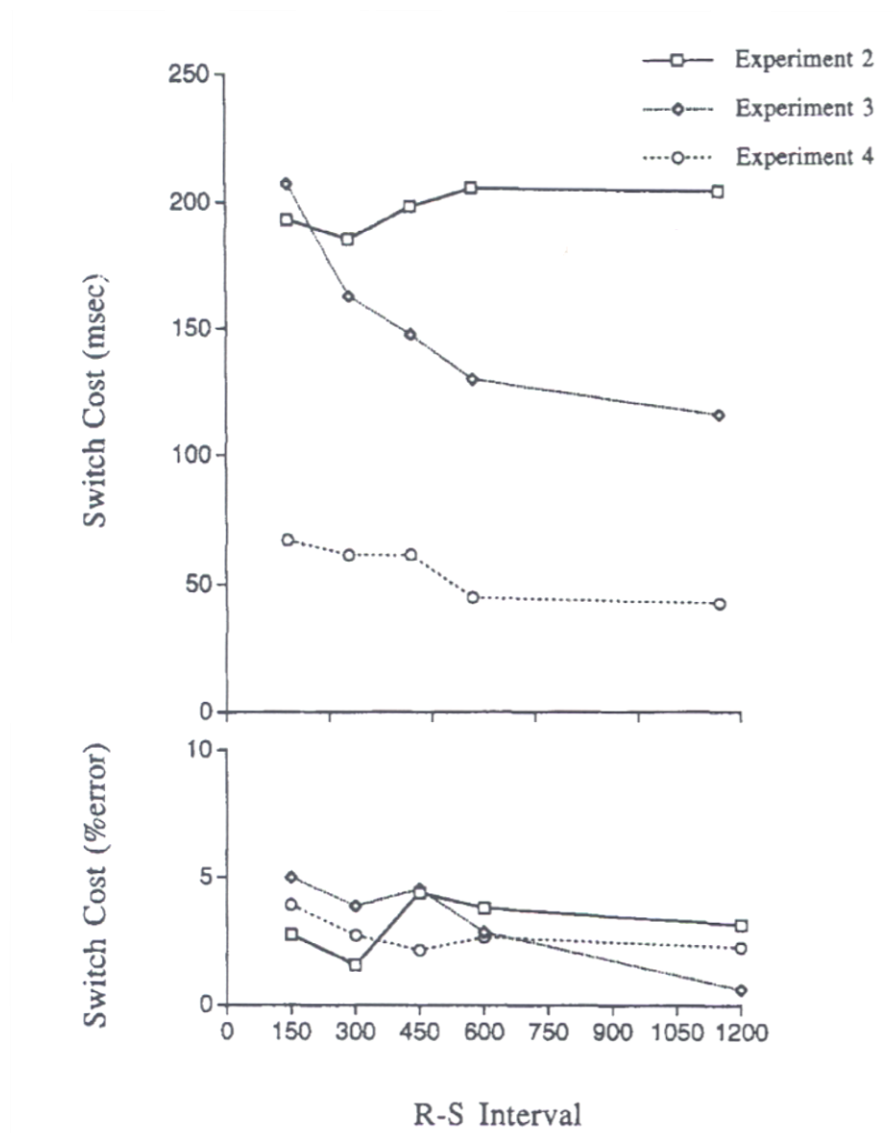


Figure 3.2. Reprint of Rogers and Monsell (1995) The cost in mean reaction time and error rate of switching task as a function of response to stimulus (R-S) interval, in Experiment 2 (varied R-S interval), Experiment 3 (constant R-S interval) and Experiment 4 (constant R-S interval). Experiments 3 and 4 used high stimulus-response overlap stimuli, Experiment 4 used low stimulus-response overlap stimuli. R. D. Rogers & S. Monsell. Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 134 (2), 207-231, 1995, APA Publisher. Reprinted with permission.

short (216 ms) CSI, indicating that pre-stimulus anticipatory task switching contributed to decreases in the RT switch cost. Furthermore, lengthening the RCI (with a constant CSI) also decreased the RT switch cost, indicating that passive decay (or *dissipation*) also contributed to the size of RT switch costs (Meiran et al., 2000). However, RCI and CSI manipulations provided optimal benefits at different time intervals. Meiran et al. (2000) reported that regardless of the accompanying RCI, increases in CSI up to 500 ms produced a sizable decrease in switch cost. By contrast, the RT switch cost showed a gradual but consistent decrease in size as the RCI was lengthened up to 2000 ms. That is, changes in the size of the RT switch cost due to decay and due to advanced preparation were independent. Critically, neither long RCIs or long CSIs (or both) eliminated the RT switch effects. Consistent with Rogers and Monsell's finding, there was a residual component to RT switch effects that was insensitive to timing manipulations. Meiran et al. (2000) concluded that at least three components could be identified in measures of switch costs: a decay component, a preparation component, and a residual component. With agreement across the alternating runs paradigm (Rogers & Monsell) and the explicit cuing paradigm (Meiran, 1996; Meiran et al., 2000), the validity of the preparation effect as a meaningful phenomenon in task switching was confirmed. Increasing the time prior to stimulus onset where participants know what task they are to perform (i.e., CSI in the explicit task cuing paradigm, RSI in the alternating runs paradigm) produces a significant decrease in switch cost beyond what can be explained by the decay of offline task sets.

The preparation effect has been observed in numerous studies using different tasks, stimuli, and experimental paradigms (see Altmann, 2004b for an overview) and has been used to explore a number of research questions. The explicit task cuing paradigm is favoured by many researchers (e.g., Altmann, 2004b; Koch 2003; Logan & Bundesen, 2003; Meiran, 1996; Verbruggen et al., 2007), but there are still researchers who have continued to utilise the alternating runs paradigm (e.g., de Jong, 2000, 2001; Lindsen & de Jong, 2010; Monsell et al., 2003; Nieuwenhuis & Monsell, 2002). More recently there has been interest in using the voluntary task switching paradigm (e.g., Liefoghe, Demanet & Vandierendonck, 2009) to understand the processes underlying preparation effects. However, it is important to note that producing a preparation effect is by no means guaranteed. Just as Rogers and Monsell's (1995) failed to observe preparation effects when RSI varied randomly within a block, researchers using other experimental paradigms have also reported cases where longer RSIs/CSIs do not reduce

switch costs (e.g., Allport et al., 1994; Koch, 2001, 2003; Altmann, 2004b). The preparation effect is sensitive to changes in method, and certain prerequisites (e.g., manipulating RSI/CSI within subjects) may need to be met in order for it to be observed (see Altmann, 2004b).

Preparation and the probability of task switch.

In response to conflicting findings reported when using an explicit task cuing paradigm with two cues per task (see Logan & Bundesen, 2003; Mayr & Kliegl, 2003), Monsell and Mizon (2006) proposed that participants may use information about the probability of a task switch occurring on a given trial to adjust their task switching and preparation strategies. That is, if participants perceived there to be a high probability that the next trial would require a switch in tasks they could change task sets *prior* to cue presentation; task switching could be anticipatory. Consistent with this proposal, Monsell and Mizon reported that as the probability of a task switch increased, the magnitude of the RT switch cost decreased to the point where the length of the CSI no longer had any effect on the size of the RT switch cost (i.e., the preparation effect was eliminated). Thus, Monsell and Mizon (2006) provided strong evidence that (at least when multiple cues were used for each task) participants are sensitive to experimental characteristics such as task switch probability.³ In particular, Monsell and Mizon's explanation of probability effects suggests that the process of switching tasks can be initiated before explicit, external information about the upcoming task is presented. If the preparation effect is explained as the consequence of pre-stimulus task set switching or activation, findings such as these force one to acknowledge that beyond providing time for reconfiguration to occur, participants' ability or willingness to reconfigure in advance of stimulus onset can be affected by any number of experimental manipulations (see Verbruggen et al., 2007, for another example). That is, whether (and how) advanced preparation occurs depends on a number of experimental factors.

Practice.

Although practice does produce an overall speeding of responses, the term 'practice effect' here refers specifically to the decrease in the size of a switch effect over blocks of trials (Jersild, 1927; Koch, 2005; Kramer, Hahn, & Gopher, 1999; Kray & Lindenberger, 2000; Meiran, 1996; Minear & Shah, 2008) or over days (Rogers &

³ Schneider and Logan (2006) replicated Monsell and Mizon's (2006) findings but posited a very different theoretical explanation that did not rely on reconfiguration or participants' sensitivity to probability of task switches.

Monsell, 1995; Monsell et al., 2003). The role of practice has been acknowledged from the very earliest task switching research. Jersild (1927) compared the effects of practice on completion times for mixed and pure blocks. In one experiment, participants performed naming and number substitution tasks in response to visual forms in mixed and pure blocks. Completion times for ten successive mixed and pure blocks were plotted and visually compared. Although both block types showed decreases in completion times across the ten blocks, this improvement was much larger for mixed blocks than for pure blocks (improvement of 13.8 and 6 seconds respectively). Practice decreased the size of the mixing cost. Jersild proposed that in addition to the individual task sets, practice produces a “more comprehensive mental set” that encompasses both tasks and the processes involved in shifting between them, eventually leading to more efficient performance on mixed blocks (p. 58).

Along with their manipulations of S-R overlap (‘#9’ versus ‘A9’) and RSI, Rogers and Monsell (1995) included analyses of performance across two days of testing (~768 trials per day). Significant decreases in RT switch cost were reported from day one to day two in all experiments, ranging from 31 to 96 ms in size. By contrast, practice effects on accuracy were only significant in two of the six experiments. Rogers and Monsell generally downplayed practice effects—noting that the decrease in RT switch cost was relatively small, especially when expressed as a proportion of baseline performance. Similar practice effects were reported by Meiran (1996), thus establishing that practice benefits task switching performance across a variety of paradigms including those most common in the literature. Meiran (1996) proposed that practice effects reflect a strengthening of associations between task cues and task sets, speeding changes in task sets.

Practice effects have been used as a tool to gauge the impact of other manipulations. Researchers have questioned whether the benefits of practice can be transferred to situations that use different tasks (Minear & Shah, 2008) or sequence predictability (Koch, 2005). However, practice effects have largely been considered an extraneous variable that needs to be controlled rather than an effect of interest per se. Researchers have attempted to control for practice by excluding the first trials (or blocks of trials) of an experiment from their analyses of task switching performance. For example, Altmann (2004b) excluded the first 175 trials from his analyses and Monsell et al. (2003) excluded the first 336 trials from theirs. These numbers are by no means unusual. One can quickly conclude that practice effects are often considered a nuisance

variable to be controlled instead of a source of information about task switching processes.

Repetition trial performance.

RT switch costs are generally calculated by subtracting repetition trial performance from switch trial performance. The repetition trial used to calculate this measure is usually the trial immediately following the switch (e.g., position two in a run). However, this method of calculating switch costs is based on the assumption that the performance cost of switching tasks is isolated to the first trial in a run (i.e., performance on all repetition trials is equivalent). There is evidence to support this assumption from studies that have used the alternating runs paradigm (Altmann, 2002; Altmann & Gray, 2002; Keele & Rafal, 2000; Rogers & Monsell, 1995). However, evidence from the explicit task cuing paradigm (Salthouse, Fristoe, McGuthry & Hambrick, 1998) and theories of task switching that implicate interference (e.g., Allport et al., 1994) would suggest that recovery from a task switch is more gradual. Monsell et al. (2003) reconciled these competing results by identifying the predictability of task switches as an essential quality that affects whether switch costs are isolated to the first (switch) trial or not. Participants made magnitude and parity decisions to digits, switching between the two tasks in a 4-trial alternating runs paradigm (predictable task switches) or an explicit task cuing paradigm (random task switches). When task switches were predictable, switch effects were isolated to the first trial in the run and the mean RTs and accuracies of position two, three, and four were roughly equal. By comparison, when task switches were unpredictable the mean RTs across the four trials of a single task decreased significantly with each subsequent trial to position three. Monsell et al. concluded that participants used their knowledge of task predictability to moderate their preparedness: when a task switch is likely to be followed by another task switch, participants will voluntarily ‘release’ the task set rather than keep it fully engaged.

Within-run slowing.

Within-run slowing refers to a gradual increase in RTs across repetition trials in a run of a task (Altmann, 2002; Altmann & Gray, 2002). Altmann and Gray (2002) used an intermittent instruction paradigm: a task cue was presented at the beginning of a run and participants continued to perform this task until the next cue appeared. Participants made magnitude and parity judgements to digit stimuli across a range of run lengths,

with mean run length manipulated between subjects ($M_{\text{short}} = 10$, $M_{\text{long}} = 20$). Repetition trials (position two onwards) became progressively slower further into a run of a task for both the short and the long run length conditions. Similar effects were reported for accuracy, as error rates became progressively higher further into a run of a task. The rate of slowing depended on the average run length. Responses slowed more rapidly when the average run length was short than that when the average run length was long.

Altmann (2002) and Altmann and Gray (2002, 2008) explained within-run slowing as a consequence of task set decay. When participants switched to a new online task, the appropriate task set was activated so that its rules could be used to generate a response. Across a run, the level of activation for the online task set decayed. If the online task set needed to be sampled to produce the appropriate response, this sampling was likely to take longer and be more error prone for later positions in a run where the task set was more decayed. Within run slowing has been observed when run length was entirely unpredictable (Altmann & Gray, 2002) as well as when it remained constant (Altmann, 2002; Poljac, Koch & Bekkering, 2009). Furthermore, Altmann (2002) reported observing within-run slowing when tasks used different stimulus sets and when spatial cues were provided as an aid to monitor task sequence. According to Altmann (2002), the observation of within-run slowing is not the result of anticipatory processes and is not dependent on the maintenance of task sequences in memory.

Task Switching Theories and Models

Early interpretations (e.g., Jersild, 1927; Spector & Biederman, 1966) of task switching emphasised the importance of S-R overlap (or ambiguity), and posited that RT switch effects reflected time needed to resolve this ambiguity. This explanation (and task switching research as a whole) remained unchanged until the early 1990s when the publication of two influential papers by Allport et al. (1994) and Rogers and Monsell (1995) reinvigorated interest in task switching and prompted the development of new task switching theories. Models of task switching can be broadly classified by whether they rely on consciously controlled processes (e.g., Meiran, 2000b; Rogers & Monsell; Rubinstein, Meyer & Evans, 2001) or whether they emphasise unconscious, automatic processes such as interference or priming (e.g., Allport et al.; Allport & Wylie, 2000; Altmann & Gray, 2008; Logan & Bundesen, 2003).

Task set reconfiguration.

Reconfiguration models of task switching originated with Rogers and Monsell (1995) and were largely developed to explain preparation effects and residual switch costs. In essence, Rogers and Monsell proposed that performance on a switch trial required the completion of two distinct stages where the online task set was loaded into working memory and used to produce responses. The first stage is endogenously controlled, time consuming, and can be completed in advance of stimulus presentation. The second stage is exogenously controlled and stimulus-locked. Full reconfiguration of task sets cannot be achieved until the stimulus has been presented (thus, producing residual switch costs). Rogers and Monsell's model has spurred the development of more detailed conceptual and computational reconfiguration models (Mayr & Kliegl, 2000; Meiran, 2000a, 2000b; Rubinstein et al., 2001; Ruthruff, Remington & Johnston, 2001; Sohn & Anderson, 2001). However, a critical characteristic of all reconfiguration models is that they hypothesise that switch trials require *additional* processes be completed (relative to repetition trials), thus producing slower and less accurate performance.

Meiran (2000a, 2000b) states that task switch effects are observed when stimuli, motor responses, or other elements are ambiguous within the context of an experiment. Examples of this ambiguity would be if a single stimulus (e.g., a happy face in the top-left corner) is associated with more than one task, or a single motor response (e.g., pressing the number 7 on the keyboard) is appropriate for more than one task. Performing in such an environment relies on the use of task sets that clarify which of the multiple interpretations of a stimulus or response is appropriate. In Meiran's (2000a, 2000b) reconfiguration model, a task set consists of a series of algorithms—most critically, stimulus and response task sets—that are configured independently of each other and often at different points in time between trial N and trial N+1. The course of a trial begins with the reconfiguration of the stimulus set as soon as information about the upcoming task is available (e.g., when a task cue is presented) and is used to direct attention towards the relevant dimension(s) of the upcoming stimulus. Response task sets are reconfigured after stimulus task sets and constrain response options to those appropriate to the task. When the stimulus is presented, task-appropriate responses are compared to the stimulus representation to produce a measure of potency (similarity), and the most potent response option is selected for execution. After a response is made,

stimulus and response task sets remain activated until a cue (externally or internally generated) prompts a new case of reconfiguration (Meiran, 2000a, 2000b).

Rubinstein et al. (2001) conceptualised performance in task switching experiments using two general processes (executive control processes and task-driven processes) that advance in parallel, with some elements occurring prior to stimulus onset but others occurring after stimulus onset. The progression of these two processes is illustrated in Figure 3.3. Executive control processes contain two stages: goal shifting and rule activation. The goal shifting stage is responsible for monitoring tasks by keeping the current ('online') goal in working memory and removing a goal when it is no longer relevant. By contrast, the rule activation stage is responsible for retrieving the appropriate task rule. For example, if a trial required participants to perform a magnitude task, the current goal would be 'make a magnitude decision', whereas the current rule would be 'if digit is numerically larger than 5 then its magnitude is more, otherwise its magnitude is less'. Rubinstein et al. stated that both goal setting and rule activation stages can be completed prior to stimulus onset, although this is less likely for rule activation. Furthermore, there are situations (short RSIs, stimuli that include explicit task cues) where these stages can become exogenously initiated.

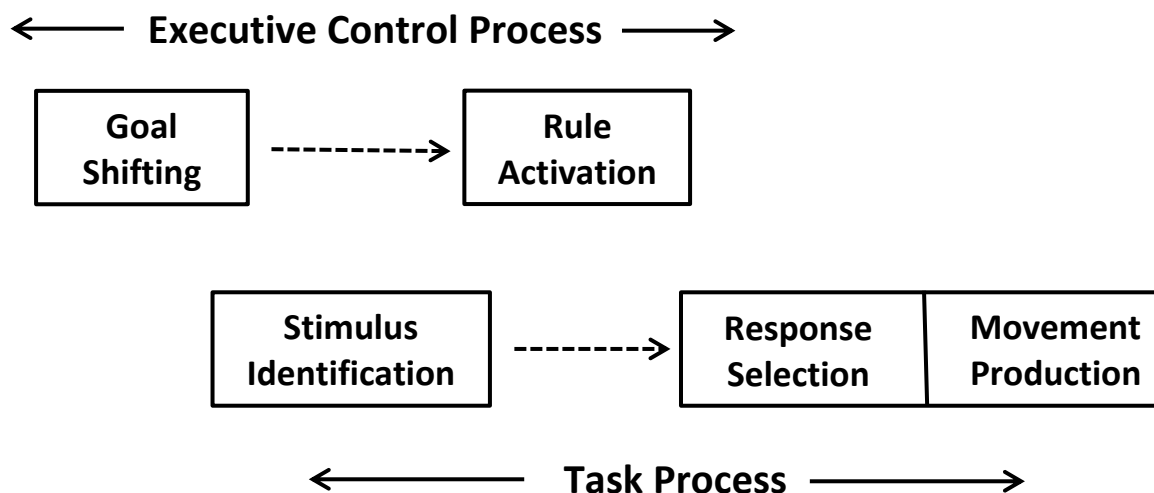


Figure 3.3. The executive control process and the task process of Rubinstein et al.'s (2000) two-stage reconfiguration model of task switching. J. S. Rubinstein, D. E. Meyer, and J. E. Evans. Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance*, 27(4), 763-797, 2001, APA publishing, Adapted with permission.

Task-driven processes consist of three stages: stimulus identification, response selection, and movement production. The stimulus identification stage involves encoding stimulus features into working memory. Stimulus identification is followed by the response selection stage, where the currently active rule (see rule activation) is used to generate the appropriate conceptual response. Finally, the conceptual response is translated into a motor response during the movement production stage. Continuing with the earlier example, upon presentation of the stimulus (e.g., ‘7’), its features—including its value—are encoded in working memory where the appropriate conceptual response of ‘more’ is produced using the current task rules. Finally, ‘more’ is translated into a motor response of pressing ‘z’ with the left index finger.

According to Rubinstein et al.’s (2001) model, the difference between repetition and switch trials is that repetition trials involve only the task processes, whereas switch trials also involve executive control processes. In its computational form, Rubinstein et al.’s model uses additive factor logic (Sternberg, 1969) in that each stage of a process (both executive control and task) must be completed in succession and the RT on a given trial reflects the sum of completion times for contributing stages. Therefore, experimental manipulations that affect different stages and/or processes should have additive effects, whereas those that affect the same stages should produce interactions. For example, Rubinstein et al. manipulated rule complexity (easy or difficult math tasks) and task cuing (providing or withholding math operation cue) within the same experiment. As rule complexity was expected to affect rule activation and response selection stages, and presenting an external cue was expected to affect goal shifting and response selection stages, they expected to (and indeed, did) observe an interaction for completion times of mixed lists. Similar use of additive factor logic has confirmed that manipulations of task expectancy and task recency affected different stages (goal shifting and response selection respectively) of reconfiguration (Ruthruff et al., 2001).

Research has provided strong evidence that switching tasks involves the reconfiguration of task sets, some of which can be initiated endogenously and some of which are locked to the presentation of a stimulus (Meiran, 2000a, 2000b; Rogers & Monsell, 1995; Rubinstein et al., 2001). However, these reconfiguration models do not rule out the impact of other more automatic processes on task switching effects. Indeed, researchers such as Sohn and Anderson (2001) have produced models of task switching performance that involve contributions from executive control processes as well as automatic processes such as repetition priming. As it stands today, most researchers

who support the role of endogenous controlled reconfiguration in task switching are also willing to accept the following: (a) task sets need to be ‘loaded’ into working memory from long term memory in order to be utilised, and it is this process which is called reconfiguration (Mayr & Kliegl, 2000; Meiran, 1996); (b) with predictable task progressions or the use of instructional cues, reconfiguration can be initiated in advance of stimulus presentation; (c) some aspect of reconfiguration cannot occur until the stimulus is presented (see de Jong, 2001, for an exception); and (d) there are components of switch effects that reflect interference and inhibition processes that exist independent from any reconfiguration (see Mayr & Keele, 2000; Meiran, 1996; Arbuthnott & Frank, 2000).

Proactive interference and negative priming.

Task set inertia theory.

Task set inertia theory proposes that when a participant switches tasks and there is an element of S-R overlap, proactive interference from the offline task set negatively affects performance. Time is required to overcome this proactive interference from competing offline task sets, resulting in task switching costs and mixing costs (Allport et al., 1994). In task set inertia theory, task switching effects are interpreted wholly within an interference framework, without reference to or reliance on consciously controlled processes unique to switch trials (i.e., reconfiguration). Task set inertia theory was developed to explain results reported in Allport et al. that used the mixed versus pure block paradigm. In Allport et al., the stimuli were arrays of digits within a rectangle (e.g., ‘4’ presented six times) and Stroop colour-words (e.g., the word ‘blue’ presented in red). For word stimuli, participants read the colour name (dominant task) or identified the colour of the word (non-dominant task). For number stimuli, participants identified the value of the digit (dominant) or the number of digits in the array (non-dominant). Consistent with Jersild’s (1927) findings, the mean time to complete a mixed list was significantly longer than that to complete a pure list (that is, there was a mixing cost), and the magnitude of this slowing was greater when the mixed lists used the same stimuli for both tasks. More importantly, mixed lists that used different stimulus sets for each task (e.g., word reading and digit value tasks) only produced mixing costs when previous lists had used the same stimuli for different tasks (e.g., participants had previously identified the colour of words).

Allport et al. (1994) compared the size of switch costs when switching from dominant to non-dominant task (e.g., word reading to colour naming) and vice versa to examine whether the relative automaticity of tasks affected the cost of switching tasks. The size of the RT switch cost was significantly larger when switching to the dominant task compared to that when switching to the non-dominant task (i.e., there was an asymmetric switch cost). Allport et al. suggested that as the dominant task (e.g., word reading) was easy and natural to perform, it required little additional activation to become the 'online' task set. By contrast, the non-dominant task (e.g., colour naming) was more difficult and had a lower level of latent activation and required more activation to become the 'online' task set. The additional activation to the non-dominant task set resulted in higher levels of interference when it was no longer relevant, thus producing larger switch costs for the dominant word reading task.

Negative priming.

Where Allport et al. (1994) established that persisting activation from a prior task set can interfere with performance and produce switch effects, the work of Allport and Wylie (2000) and Wylie and Allport (2000) found that this interference was largely due to the inappropriate retrieval of task sets and responses previously associated with a *specific stimulus*. That is, interference is the product of a specific stimulus (e.g., 'blue' presented in red) prompting the retrieval of a specific response (e.g., 'blue'), rather than a stimulus set (e.g., colour words) prompting the retrieval of a task set (e.g., 'read the word'). Allport and Wylie used Stroop stimuli including incongruent colour words, neutral colour stimuli and neutral words (as in Allport et al., 1994). In the experiments of interest (Experiments 3 – 5), participants completed three phases: (1) a baseline block of the word reading task; (2) a block of the colour naming task (interference condition) or a block of rest (control condition); and (3) a final block containing two runs of word reading with instructions between the runs. Note that participants never switched tasks within a block and only those in the interference condition performed more than one task in the whole experiment. RT performance for Stroop words in the first run of phase three was slower overall for the interference condition than the control condition. Having previously made colour naming decisions participants produced slower responses. In another experiment, Allport and Wylie (2000) manipulated S-R associations at a stimulus level. During the second phase of the experiment, only half of the colour-word stimuli were presented for participants to make colour naming responses. Thus, only half of the Stroop colour words developed competing colour

naming S-R associations—they were negatively primed. Phase three performance was compared for colour-words that were negatively primed in phase two and colour-words that were not negatively primed. For the first trial in a run, the mean RT for negatively primed colour-words was significantly slower than that for the unprimed colour-words. Allport and Wylie (2000) concluded that a significant portion of the interference from alternative task sets is stimulus specific, particularly on switch trials.

Waszak, Hommel and Allport (2003, 2004) posited that the effects of negative priming are largely isolated to switch trials because the activation of the online task is relatively weak and is therefore particularly vulnerable to irrelevant S-R associations. When stimuli are associated with multiple tasks they develop multiple S-R associations over the course of an experiment. Consequentially, the presentation of a stimulus can result in the spontaneous retrieval of S-R associations specific to that item that are currently counterproductive. This type of spontaneous (yet incorrect) retrieval of S-R associations produces the longer RTs observed on switch trials (Allport & Wylie, 2000; Koch & Allport, 2006; Waszak et al., 2003, 2004; Wylie & Allport, 2000).

Cognitive control model—activation and decay.

The Cognitive Control Model (CCM) of task switching (Altmann & Gray, 2008) frames task switching performance within the ACT-R memory model (Anderson, 1996). Put simply, the CCM model posits that the successful completion of a trial (be that a task switch or a task repetition trial) requires that the appropriate task set is sampled from memory and applied to the current stimulus. The CCM model does not rely on (nor support the existence of) consciously controlled reconfiguration processes. When a task cue is presented, it creates/updates and activates a representation in episodic memory called a *task code* (or task set). The activation of the ‘online’ task code and the most active ‘offline’ task code are represented using signal detection theory (Green & Swets, 1974) by two normally distributed activation densities (Altmann & Gray, 2008). The point at which the online and offline task codes intersect is called the *retrieval threshold* and activation above this threshold means that the online task is *accessible*. When the appropriate task set is accessible, it can be used to guide task performance and select the appropriate response to the presented stimulus. When one switches tasks, a greater amount of activation (and thus time) is required to reach the retrieval threshold and overcome interference from the offline task set, producing task switching effects.

In its computational form, the CCM model uses multiple parameters, such as the retrieval threshold, baseline activation levels of task codes, and activation of offline task codes to simulate data for a number of critical task switching phenomena. Within-run slowing has already been discussed in this chapter, with its effects representing changes in decay rates for task codes (Altmann & Gray, 2002, 2008). Another phenomenon that the CCM model has been used to explain is the preparation effect—manipulations of RSIs/CSIs change when the activation of a task code is initiated. In brief, identifying a cue and then activating the online task code to a level above threshold takes multiple encoding cycles (time). With long preparation intervals, the likelihood of cue identification and task code activation being achieved prior to stimulus onset is higher. Thus, fewer encoding cycles are likely to be needed after stimulus presentation, producing faster RTs (Altmann & Gray, 2008).

Compound cue retrieval model.

The compound cue retrieval model explains task switch effects when an external task cue is presented (see Logan & Bundesen, 2003; Mayr & Kliegl, 2003). Logan and Bundesen (2003) noted that by combining information from the cue and stimulus, the correct answer can be retrieved directly from long term memory (LTM) without needing to reconfigure task sets. For example, when completing a magnitude task, a cue of ‘magnitude’ and a target of ‘7’ produces a compound cue (‘magnitude-7’) that can be used to retrieve the correct answer ‘more’ from memory. According to the compound cue retrieval model, the presentation of a task cue initiates the process of comparing that cue to representations held in memory. When trial N and trial N-1 cues are the same, representations of the cue are held in LTM and short-term memory (STM). However, when trial N and trial N-1 cues are different, representations of the cue are held only in LTM. As the rate of processing (or encoding) a cue depends on the memory systems it is represented in, cue repetitions (held in STM and LTM) produce faster processing rates (and shorter encoding times) than that for cue switches (held only in LTM). A shorter cue encoding time means that the correct response can be retrieved from LTM more quickly, thus decreasing RTs and the size of switch effects. Thus the compound cuing model, like the CCM model, argues that the same processes occur for repetition and switch trials. The difference between switch and repetition trial performance represents a *benefit* of repeating a task rather than a cost of switching tasks. Subsequent work by this group of researchers (Logan & Bundesen, 2004; Logan & Schneider, 2010;

Logan, Schneider & Bundesen, 2007; Schneider & Logan, 2005, 2006, 2007) has continued to develop and provide evidence to support the compound cue retrieval model, claiming that (at least for the explicit task cuing paradigm), additional reconfiguration processes are unnecessary to explain task switching performance.

Summary of task switching theories.

The task switching models outlined above can be broadly classified by whether they include a role for endogenous cognitive control process(es) (Mayr and Kliegl, 2000, 2003; Meiran, 1996; Meiran, 2000a, 2000b; Rogers & Monsell, 1995; Rubinstein et al., 2001) or not (Allport & Wylie, 2000; Altmann & Gray, 2002, 2008; Logan & Bundesen, 2003; Wylie & Allport, 2000). Generally speaking, reconfiguration models of task switching tend to be more accepting of additional parameters of interference or inhibition, although they fall short of formally including such processes in their models (i.e., Mayr and Kliegl, 2000; Rubinstein et al., 2001). By contrast, Altmann and Gray (2002, 2008) have argued that task switching effects can be explained without reference to consciously controlled reconfiguration processes. Similar statements have been made by proponents of the compound cuing model of task switching (Logan & Bundesen, 2003, 2004; Logan, Schneider & Bundesen, 2007) and task set inertia / negative priming theorists (Allport & Wylie; Wylie & Allport).

Chapter Four

Task Switching and Episodic Memory

In the opening of Chapter 3, the interruptions effects reported by Jones et al. (2011) were interpreted from a task switching perspective. It was suggested that the periodic interruption of the recognition task for a single trial was functionally very similar to an infrequent task switch. From this perspective, the slower RTs and lower accuracy for recognition decisions that followed an interruption trial (recall or colour identification) reflected the cost of switching from the interruption task set back to the recognition memory task set. Of particular interest to this thesis is that Jones et al. showed that participants' ability to correctly identify old and new words was significantly impaired when they had recently switched tasks. This thesis seeks to further explore this observation and to use task switching and recognition memory methods to investigate task switching using a task arguably more complex than those typically used in the task switching literature (that is, the recognition task).

Three studies have reported switch effects on explicit memory (Kavcic, Krar & Doty, 1999; Mayr & Kliegl, 2000; Werkle-Bergner, Mecklinger, Kray, Meyer & Duzel, 2005). Mayr and Kliegl (2000) reported significant task switching effects when participants switched between making source recognition decisions and semantic memory decisions. In the initial phase, participants repeatedly studied a short list of words ($N = 16$) and were then tested for their memory of each word's location and colour. The aim was for participants to learn each word's colour and location to a set criterion—a word was dropped from subsequent study-test cycles when its colour and location had been correctly recalled four consecutive times. In the test phase, these same words were used as stimuli for the completion of four tasks: recalling the colour at study, recalling the location at study, reporting the size of the object, and reporting whether the object is living or non-living. The first two tasks required episodic memory whereas the second two tasks required semantic memory. Four blocks (two tasks per block) using 4-trial alternating runs (AAAA BBBB) were presented to participants in the test phase. Each test block used a different combination of semantic and episodic tasks. Mayr and Kliegl (2000) reported significant RT switch costs for all four combinations of tasks, although this cost was larger when switching to an episodic task

(range ~1200 ms to 1350 ms) than that when switching to a semantic task (range ~950 – 1050 ms). Similarly, participants committed more errors on switch trials relative to repetition trials, particularly when the online task was episodic. These effects were also observed using an explicit task cuing procedure. Mayr and Kliegl (2000) concluded that tasks with large long term memory (LTM) retrieval demands (e.g., episodic memory tasks) took longer to install in working memory. Without adequately long preparation intervals or cues that eliminated the need to retrieve the task set from LTM, large switch costs were observed. However, for the interests of this thesis the critical finding of Mayr & Kliegl (2000) was that switching tasks had a negative effect on source recognition performance (both RT and accuracy), even though stimuli had been thoroughly studied prior to the test phase.

Werkle-Bergner et al. (2005) used a recognition task and a source recall task in their examination of the electrical signals in the brain (event related potentials, ERP) associated with switching between tasks. The experiment consisted of eight study-test blocks. During study, 30 words were presented one at a time and participants were instructed to memorise the word and its font, as well as judge whether the word contained the letter 'a'. There were two tasks at test: identifying words as "old" or "new" (general task, G), and judging whether the target word was presented in the same or different font relative to its study presentation (specific task, S). Half of the study-test blocks used only one task at test (pure blocks) and half used 2-trial alternating runs (mixed blocks: GG SS...). For both types of test block, geometric frames (rectangles or ovals) were presented 300 ms before target onset as task cues. In addition to ERP data, Werkle-Bergner et al. reported behavioural data consistent with mixing effects and switch effects. The mean RT was slower and the mean accuracy was lower for mixed blocks than those for pure blocks. Within the mixed blocks, comparing performance on repetition and switch trials also produced significant RT (~134 ms) and accuracy (~0.13) switch effects. The magnitude of these switch effects did not differ significantly for either task. Thus, Werkle-Bergner et al. (2005) provided evidence that recognition memory performance could be affected by switching between two types of episodic memory tasks.

Mayr and Kliegl (2000) and Werkle-Bergner et al. (2005) both reported RT and accuracy switch effects on episodic memory performance. As is common in the task switching literature, the tasks they were switching between used the same stimulus sets. By comparison, Kavcic et al. (1999) investigated the effect of switching between

different sets of visual stimuli (but maintaining the same task). The experiment used a continuous recognition paradigm. Participants were instructed to respond whether each stimulus had been presented earlier in the experiment (i.e., it was old) or not (it was new). The stimuli were words, nonwords, faces, coloured butterflies, coloured abstract images, and monochromatic abstract images (20 per stimulus set). Whether stimulus set switching was predictable or unpredictable was manipulated between participants. In the predictable condition stimulus type switched every trial, whereas in the unpredictable condition stimulus type switched pseudo-randomly. As there was only one task—to make recognition memory decisions—no task cues were used. More often than not, Kavcic et al. did not observe significant set switching effects for recognition accuracy. However, with unpredictable stimulus switches (but not predictable) between particular stimulus sets (nonwords following words, butterflies following faces), reliable switch effects were observed (accuracy switch cost of 0.06 and 0.11 respectively). The effect of set switches on recognition RTs was considerably more reliable. Kavcic et al. reported that stimulus set switches produced a significant increase in RT, regardless of switch predictability (RT switch cost 24 – 92 ms).

Kavcic et al. (1999) argued that stimulus set switch effects were comparable to measures of residual switch costs and that these switch effects represented the time taken to end one processing mode (e.g., processing faces) and initiate the next processing mode (e.g., processing words). Similar conclusions can be drawn when Kavcic et al.'s experiment is interpreted using Meiran's (2000a, 2000b) model of task switching. Each stimulus in Kavcic et al. uniquely belonged to one stimulus set—they were univalent. By comparison, “old” and “new” motor responses (i.e., the keys pressed to respond) were bivalent. The same motor responses were made in response to memory searches for word, face, and butterfly stimuli. In essence, switching from face stimuli to butterfly stimuli affected a change in response set from ‘respond old face or new face’ to ‘respond old butterfly or new butterfly’. With univalent stimuli and bivalent responses, Meiran's (2000a) model of task switching hypothesises that performance costs will reflect the residual component of a switch cost. Thus, Kavcic et al. provided evidence that stimulus switches can negatively affect recognition memory performance and may be considered a type of task switching.

Justification for Current Studies

The work of Mayr and Kliegl (2000), Werkle-Bergner et al. (2005) and Kavcic et al. (1999) confirmed that task switching and stimulus set switching can negatively affect performance on episodic memory tasks. However, it is difficult to attribute these switch effects to a failure of memory proper rather than to interference from other task sets associated with stimuli or negative priming from earlier stimulus-response associations. Mayr and Kliegl's (2000) relatively overlearned, very small set size could have produced an underestimation of the real cost to memory. Their use of forced recall for source information and the absence of new words at test meant that Mayr and Kliegl's (2000) results do not give an indication of participants' ability to discriminate between old and new words. Finally, although source recognition and item recognition are arguable related, they reflect qualitatively different types of memory (Glisky, Polster & Routhieaux, 1995) and may be differently affected by task switches. Finally, although Werkle-Bergner et al. established that item recognition performance was lowered on switch trials they did not examine what this observation meant for memory processes. These studies have used memory tasks as tools to examine task switching rather than using task switching as a tool to examine memory. This thesis attempts to cast the investigative light in both directions: to examine how episodic memory is affected when participants switch tasks as well as how the process of switching tasks is affected when the task itself involves retrieving information from episodic memory.

General Methodology

All of the experiments in this thesis used a study-test recognition memory design. In a study phase, participants were presented with words to study, sometimes with no further instructions and sometimes with the addition of a semantic or perceptual judgement task. A study phase was followed by a test phase, where participants switch between a recognition memory task and a magnitude task in 1-, 2-, and 4-trial alternating runs. Given the transparency of the two tasks and that task repetitions and switches were always predictable, external task cues are generally not supplied (although Experiment 3 uses a visual sequence cue in some conditions).

For the magnitude task, an array of dots was presented to participants who were instructed to judge whether there were more or less than specific number of dots. The choice of secondary task was motivated by several factors. Firstly, I wanted a secondary task that would use different stimuli to the primary recognition memory task and that

would elicit responses that could not be applied to the word stimuli. This choice to use a magnitude task limited the S-R overlap between the two tasks so that the only element shared across the tasks was a motor response (responses to both tasks were made on the same keys on the keyboard). An additional consequence was that it was impossible for participants to inadvertently perform the wrong task—a magnitude judgement could not be made for word stimuli, and an old/new decision could not be made for dot arrays. Any differences in recognition accuracy between repetition and switch trials would represent a failure on the part of long term memory as a result of switching tasks. Although novel, dot arrays have been used in task switching experiments before; Ruthruff et al. (2001) used dot stimuli arranged as they would be represented on a die in place of numbers for a parity task. In this thesis, the magnitude task used randomly positioned dots within a predefined area of the monitor such that none of the dots overlapped. The random positioning of the dots discourages overestimation of the number of dots in a display (Ginsburg, 1978), as does ensuring that dots are not overlapping or connected (He, Zhang, Zhou & Chen, 2009). More importantly, the random positioning of dots ensures that, like the word stimuli in the recognition memory task, each magnitude judgement is made in response to a unique array of dots. The use of unique stimuli for every trial during the study phase seriously limits the effects of positive or negative S-R priming (Allport & Wylie, 2000; Wylie & Allport, 2000).

The Aims of This Thesis

This thesis examines how task switching and recognition memory methods and theories can be combined to increase our understanding of these disciplines separately. Additionally, it presents an opportunity to investigate task switching effects with relatively novel tasks in a novel experimental design. Logan (2003) stated that the advancement of task switching theory and understanding would require researchers to expand their horizons beyond the simple tasks and standard task switching methods typically used in the literature. This thesis aims to contribute toward that goal by examining the effects of switching between making recognition decisions to word stimuli (recognition task) and magnitude judgements to dot stimuli (magnitude task). The experiments in this thesis were developed to address three primary research aims:

1. To establish that task switching effects can be observed when switching between two tasks that utilise different stimuli and arguably different cognitive processes.

2. To examine how task switching methods can be used in conjunction with recognition memory theory to establish the locus of costs to memory performance.
3. To examine the characteristics of two common task switching effects—preparation effects and practice effects—within the more naturalistic experimental conditions produced when switching between recognition and magnitude tasks.

Experiments 1 and 2 establish that task switching effects can be observed when participants switch between the recognition task and the magnitude task. They develop a reliable experimental method of examining task switching effects on recognition performance and seek to confirm that the observed switch effects are consistent with those observed in earlier task switching research. Importantly, Experiment 2 also examines whether switching tasks differentially affects familiarity or recollection processes. Experiment 3 examines how performance changes over a 4-trial run for the purpose of examining recovery from a task switch (Monsell et al., 2003) and within-run slowing (Altmann, 2002; Altmann & Gray, 2008). The value of providing a visual sequence cue to remind participants of an upcoming task switch is also investigated. The focus of Experiments 4A, 4B, 5A, and 5B is on two commonly observed phenomena in task switching research: preparation effects and practice effects. Each experiment comprised two study-test blocks and each test block was associated with a single response-to-stimulus interval (RSI). With the use of different stimulus sets for each task and the absence of external cues to scaffold performance, these experiments aim to study preparation and practice effects in an experimental design that more closely resembles everyday life experience. Critically, these experiments represent an attempt to establish a boundary for when preparation effects are observed.

Performance on the secondary magnitude task is not reported for Experiments 1 and 2, as they focus on task switching effects as they relate to recognition memory performance and theory. Although recognition memory performance remains the primary interest, Experiments 3 – 5B report task switching performance for both the recognition and magnitude tasks as these experiments focus on more general task switching effects. The primary measures of performance are RT and accuracy as a function of task transition (switch or repetition trial). To preview the general findings, switching between tasks consistently resulted in significant RT switch costs. Across all experiments, accuracy of switch trials was lower than that for repetition trials and this difference reached significance more often than not.

Chapter Five

Experiment 1: Establishing Task Switching Effects

Experiment 1 was designed to address the first aim of this thesis, namely to establish whether task switch effects could be observed when switching between a recognition memory task and an unrelated secondary task with limited stimulus-response (S-R) overlap. During the study phase, participants studied a list of words. In a test phase, participants switched between making recognition memory (R) decisions to word stimuli and magnitude (M) decisions to dot stimuli. The two tasks were presented in an alternating runs design, and run length (1-, 2-, or 4-trial runs) was manipulated between participants. The primary purpose of manipulating run length was to establish the most efficient and effective way of assessing task switching effects with these particular tasks.

Jones et al. (2011) reported significant RT and accuracy interruption effects when the interruption was unpredictable and tasks used different S-R sets. However, the interruptions were relatively rare and were only present for one trial. When using 1-trial runs (mixed blocks), some researchers have struggled to observe task switching effects with tasks that do not share stimulus sets (e.g., Jersild, 1927). When they have been observed, they have been eliminated with a very small amount of practice (Allport et al., 1994). By contrast, Rogers and Monsell (1995) observed significant RT and accuracy switch effects that persisted over two days of testing when using a 2-trial alternating runs paradigm and tasks with limited S-R overlap (e.g., ‘#9’ for the parity task and ‘A&’ for the letter task). Given these mixed results, it is important to assess the size and resilience of switch effects across a variety of run lengths to establish the most effective way of testing task switching effects with these types of tasks.

For a 1-trial run condition (i.e., a mixed block) every trial is a switch trial, making it impossible to compare switch and repetition performance within participants when only a single test block is used. In order to keep the analysis of switch effects consistent across run conditions it was therefore necessary to include an additional condition where critical trials were repetition trials. Traditionally, performance on mixed blocks has been contrasted with performance on task-pure blocks (e.g., Allport et al., 1994; Baddeley et al. 2001; Jersild, 1927). However, this method confounds task

switching and cognitive load, as participants must maintain task sets for two tasks in mixed blocks and one task in pure blocks (Rogers & Monsell, 1995). Also, when task cues are not provided, mixed blocks require participants to monitor task sequence—what they have just performed and what they are to perform next—but pure blocks do not require this additional cognitive process (Baddeley et al., 2001). To avoid these confounds, repetition trial performance was measured by a separate run condition (hereafter referred to as the control run condition) that used a 2-trial alternating runs design offset by one trial. This approach meant that there were two groups that used a 2-trial alternating runs: the 2-trial run condition where critical trials were switch trials (MM RR MM...) and the control run condition where the critical trial was a repetition trial (M RR MM R...). It is important to note that comparisons of switch and repetition trial performance could be made within a run for the 2-trial and 4-trial conditions—these run conditions contained switch and repetition trials. However, in this experiment all analyses of task switching effects (RT and accuracy) were made by comparing critical trial performance in the experimental conditions (1-, 2-, and 4-trial run conditions) to critical trial performance in the control run condition. This held the position of critical switch and repetition trials constant across all run conditions and maintained a common point of comparison for all experimental conditions. Additionally, this meant all analyses of task switching effects were made between participants, as opposed to some within (2- and 4-trial runs) and some between (1-trial run).

The primary hypothesis for Experiment 1 was that switch effects would be observed for recognition memory RT and perhaps for accuracy. That is, recognition memory performance on the critical trials would be slower and less accurate for the 1-, 2-, and 4-trial run conditions (where they were switch trials) compared to that for the control run condition (where they were repetition trials). An additional interest that was explored in this experiment was whether the length of a run had an impact on performance on task switch trials. The Cognitive Control Model (CCM, Altmann & Gray, 2008) posits that the activation of a task set decays over time. As such, one could hypothesise that the longer a task set has been ‘offline’, the more activation (and thus, time) the task set should require to reach threshold. Despite this seemingly logical assumption, research that has manipulated run length has either failed to report if performance on the switch trial is affected by the number of trials that have passed since its last use (Altmann, 2002; Altmann & Gray, 2002) or has reported no significant

differences (Monsell et al., 2003). Therefore, the effect of run length on task switching performance was explored with a tentative hypothesis that, for the three experimental conditions (1-, 2-, and 4-trial run conditions), critical recognition memory RTs may increase as run length increases.

Method

Participants

Ninety-six undergraduates (24 per run length condition) from Victoria University of Wellington participated for research credit towards an introductory psychology paper. All participants in the final data set: (1) had normal or corrected-to-normal vision, (2) spoke English from early childhood, and (3) had mean RTs for correct critical recognition trials within 3 *SD* of the group mean. Participants who failed to meet the above criteria were replaced to achieve proper counterbalancing.

Materials

Words ($N = 128$) from the English Lexicon Project (Balota et al. 2007) were used in the experiment. Thirty of these words were used to create two critical lists (see Appendix A) that were balanced for mean word frequency (Kucera & Francis, 1967, $M = 75.83$), word length ($M = 6.50$), and syllables ($M = 1.97$). The remaining words had similar properties and were used as filler words in the study and/or test phases. Dot stimuli were black circles (diameter = 35 mm). Words were presented centrally on a white screen in black bold 18-pt Courier New font. The experiment was conducted using E-Prime software (Schneider, Eschman & Zuccolotto, 2002).

Design and Procedure

The experiment comprised a study phase and a test phase. During the study phase, participants viewed and attempted to learn a list of words. During the test phase participants switched between the recognition memory task and the magnitude task in predictable 1-, 2-, or 4-trial runs. Critical recognition trials were manipulated in a 4 (run condition: 1, 2, 3, control) x 2 (item type: old, new) mixed design. Participants sat approximately 0.5 meters from the monitor at individual computer stations. One to four participants were run per session. General instructions for the experiment and instructions for the study phase were presented visually and verbally, after which participants worked at their own pace.

Study Phase.

The study phase consisted of 68 trials: 4 primacy buffer words, 15 words from a critical list, 45 non-critical words, and 4 recency buffer words. Words from the critical list were evenly dispersed every four words throughout the study phase, excluding the first and last four study trials, which were used for primacy and recency buffer words. Participants were instructed to study each word as they would be tested on their memory later in the experiment. A study trial began with the presentation of a brief (100 ms) focal stimulus (+++) presented in the centre of the monitor. Immediately after the focal stimulus the target word was presented in the same position for 3000 ms. An intertrial interval (ITI; 500 ms) with a blank screen occurred between a study word and the focal stimulus on the following trial.

Test Phase.

The test phase comprised a practice test and the actual test. Participants alternated between a recognition memory task (R) and a magnitude task (M) every one, two, or four trials, as illustrated in Figure 5.1. The actual test for the 1-, 2-, and 4-trial run conditions were constructed so that critical trials were switch trials. The actual test for the control run condition was constructed so that the critical trial was a repetition trial. Participants were informed that they would be alternating between a magnitude task and a recognition memory task in a predictable sequence. The length of each sequence—which varied depending on run condition—was described verbally and illustrated visually to participants.

The 16-trial practice test used buffer words from the study phase or new words otherwise not presented in the experiment. The 240-trial actual test used words from the studied critical list (critical ‘old’ recognition trials), the nonstudied critical list (critical ‘new’ recognition trials), studied non-critical words (other ‘old’ recognition trials), and non-studied words (other ‘new’ recognition trials). The allocation of critical list to item type (old, new) was balanced across participants. During the actual test, the presentation of the first critical word occurred after four trials, and every eight trials subsequently. The actual test was designed so that ‘old’ and ‘new’ recognition responses were randomly distributed across the critical recognition trials. For the magnitude task, the actual test used an equal number of ‘more’ and ‘less’ magnitude trials (four and six dots

Trial	1-trial	2-trial	4-trial	Control
1	RGN	RGN	MAG	RGN
2	MAG	RGN	MAG	MAG
3	RGN	MAG	MAG	MAG
4	MAG	MAG	MAG	RGN
5	RGN	RGN	RGN	RGN
6	MAG	RGN	RGN	MAG
7	RGN	MAG	RGN	MAG
8	MAG	MAG	RGN	RGN

Figure 5.1. Trial and task progression for the 1-, 2-, and 4-trial run conditions and the control run condition in Experiment 1. RGN = recognition memory trial, MAG = magnitude judgement trial. The critical trial (trial 5) is presented in bold. Critical trials are switch trials for 1-, 2-, and 4-trial run conditions, and repetition trials for the control condition.

respectively). A single list of magnitude trials was constructed with ‘more’ and ‘less’ magnitude responses distributed pseudo-randomly. Magnitude trials in the actual test selected trials from this list sequentially in the pattern relevant to the run condition (i.e., 1-, 2-, or 4-trials in a row).

A recognition memory trial began with the brief (100 ms) presentation of a focal stimulus (+++) in the centre of the monitor. Immediately after the focal stimulus the target word was presented in the same location. Old and new judgements could be made at any time after the word was presented using the *f* and *j* keys on the keyboard. The allocation of keys to responses was balanced across participants. A trial terminated when a response was made, and the next trial began after a 500-ms ITI of a blank screen.

A magnitude trial began with the brief (100 ms) presentation of a focal stimulus (+++) in the centre of the monitor. Immediately after the focal stimulus the dot targets (four or six) were presented. The location of dots on any given trial was random, with the proviso that no dots overlapped. Participants were instructed to judge whether there

were more or less than five dots and to respond using the *f* (more) or *j* (less) keys on the keyboard. The allocation of key to magnitude judgement response was not balanced as there was no interest in magnitude performance or in comparing performance for ‘more’ and ‘less’ decisions. A trial terminated when a response was made, and the next trial began after a 500-ms ITI of a blank screen.

After receiving instructions about the two tasks, participants completed the practice test. Participants were then reminded about the response allocations for the recognition and magnitude tasks prior to starting the actual test. Participants were debriefed and thanked for their participation after completing the experiment.

Results

All statistical analyses were conducted on critical recognition memory trials. Data from the practice test were excluded from analyses. Statistical analyses of RT performance used participants’ median RTs for correct critical recognition memory trials. An alpha level of 5% was used for all statistical tests.

Accuracy

Corrected accuracy was calculated by subtracting the false alarm rate from the hit rate for critical recognition trials. The mean hit and false alarm rates for this and all subsequent experiments can be found in Appendix B. Accuracy for the control (task repetition) condition was moderate ($M = 0.50$, $SD = 0.19$) and well above chance (0), indicating that participants were capable of discriminating between old and new words. Accuracy for the 1-, 2-, and 4-trial run (task switch) conditions was very similar and showed good discrimination between old and new words. Of these three conditions, mean corrected accuracy of the 1-trial run condition ($M = 0.54$, $SD = 0.23$) was highest, followed by the 2-trial run condition ($M = 0.47$, $SD = 0.17$), and the 4-trial run condition ($M = 0.43$, $SD = 0.14$). Accuracy in the 2- and 4-trial run conditions was numerically lower than that for the control condition, consistent with a switch effect on recognition accuracy. However, accuracy in the 1-trial run condition was higher than the control condition and inconsistent with an accuracy switch effect.

A Oneway ANOVA was conducted on corrected accuracy to examine task switching effects. The main effect of run condition was not significant ($F(3, 92) = 1.52$, $p = 0.215$), indicating that there was no accuracy switch effect for the recognition task and no difference between the switch trials for the different run lengths. The lack of a significant task switching effect may have in part been due to the large variance in

performance, particularly in the 1-trial run condition which is arguably a special case in task switching (i.e., it is a mixed block as opposed to an alternating run). However, even when the 1-trial run condition was excluded from the analysis, the main effect of run condition failed to reach significance ($F(3, 92) = 1.07, p = 0.348$). Likewise, combining the 2- and 4-trial run conditions (switch trials) and comparing to the control run condition (repetition trials) using a planned difference contrast did not produce a statistically significant task switching effect on recognition accuracy ($p = 0.256$).

Reaction Times

RT analyses utilised correct recognition responses, resulting in the exclusion of 26% of the critical recognition memory trials. These exclusion rates are higher than reported in the task switching literature, but reflect the relative difficulty of the recognition task in comparison to typical tasks used to investigate task switching effects. Therefore, the exclusion rate is not overly concerning. The means of median RTs for the recognition task are reported in Table 5.1 as a function of item type. To be clear, the RTs reported for the 1- 2- and 4-trial run conditions represented switch trial performance and the RTs reported for the control condition represented repetition trial performance. The mean RTs for the control run condition were considerably faster than those for the 1-, 2-, and 4-run conditions, providing evidence of an RT switch effect. The RTs tended to slow as the run length increased, although compared to the

Table 5.1

Means and Standard Deviations for Correct Critical Recognition RTs in Experiment 1 as a Function of Item Type and Run Condition.

Run Condition	Item Type			
	Old		New	
	M	SD	M	SD
1-trial run	1199	374	1261	408
2-trial run	1410	544	1321	423
4-trial run	1454	583	1454	509
Control	884	179	995	227

Note. Mean RTs for 1- 2- and 4-trial run conditions are switch trials. Mean RTs for the control run condition are repetition trials.

difference in RT between repetition and switch trials, these increments were relatively subtle.

To confirm the presence of a RT switch effect a 2 (item type: old, new) x 4 (run condition) mixed ANOVA was conducted on correct critical recognition RTs. There was a significant main effect of run condition [$F(3, 92) = 7.60$, $MSE = 319649.73$, $p < 0.001$, $\eta_p^2 = 0.20$], illustrated in Figure 5.2. Post-hoc comparisons using Games-Howell corrections for unequal variance were used to examine group differences. The mean RT for the control condition ($M = 939.34$, $SD = 189.61$) was significantly faster than that for the 1-trial condition ($M = 1230.28$, $SD = 352.55$), 2-trial condition ($M = 1365.56$, $SD = 460.79$) and 4-trial condition ($M = 1454.18$, $SD = 516.46$). The RTs for the 1-, 2-, and 4-trial run conditions did not differ significantly. That is, although there was a significant switch effect on correct recognition RTs there was no evidence that run length affected the size of the RT switch effect. Neither the main effect of item type ($F = 0.46$, $p = 0.50$), nor the interaction of item type and condition ($F = 1.93$, $p = 0.13$) were significant.

Exploring Practice Effects

As practice switching tasks has been shown to decrease the size of the RT

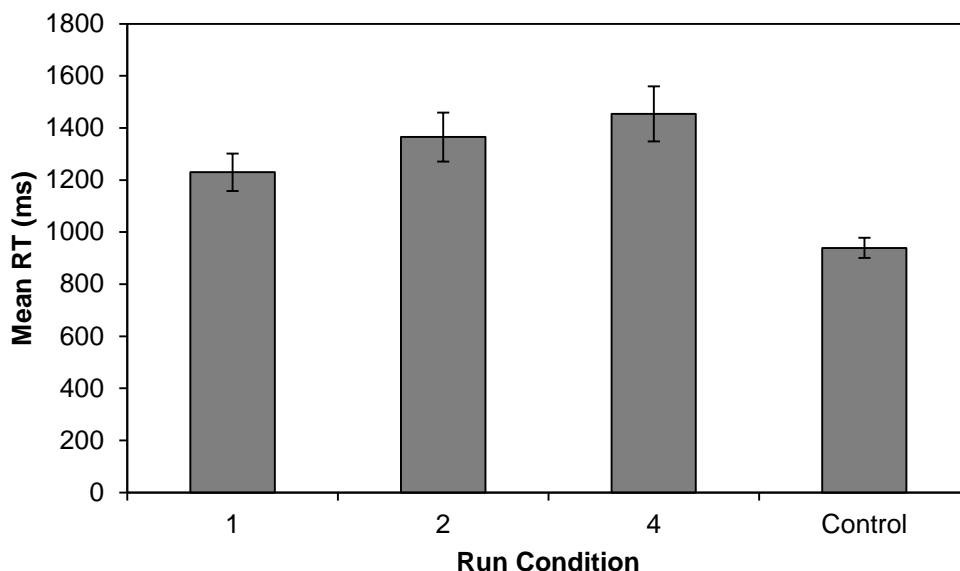


Figure 5.2. Mean RTs for correct critical recognition decisions in Experiment 1 as a function of run condition. The mean RTs for 1- 2- and 4-trial run conditions are switch trials. The mean RT for the control condition is a repetition trial. Error bars represent S.E.

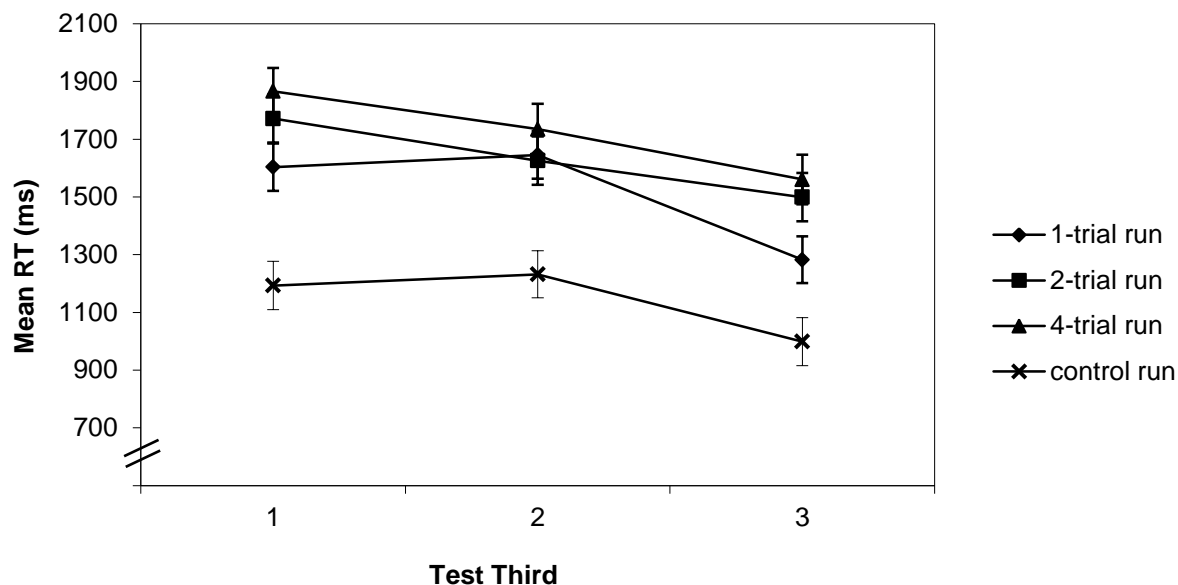


Figure 5.3. Mean RTs of correct critical recognition memory decisions in Experiment 1 as a function of test third and run length condition. The mean RTs for 1- 2- and 4-trial run conditions are switch trials. The mean RTs for the control condition are repetition trials. Error bars represent S.E.

switch effect (e.g., Koch, 2005; Meiran, 1996; Minear & Shah, 2008; Rogers & Monsell, 1995) it is worth examining if (and how) performance on the critical recognition trials changed over the course of the experiment. To examine evidence for this change in performance across the course of the experiment, raw RTs for correct critical recognition memory trials were coded post-hoc for test third (1st = critical trials 1-10; 2nd = critical trials 11-20; 3rd = critical trials 21-30). Given the exploratory nature of this enquiry, statistical analyses were not performed on these data. Instead, the mean recognition RT for each run condition was plotted as a function of test third so that a visual assessment of practice effects could be made.⁴ As can be seen in Figure 5.3, the mean RTs were faster for the control run condition (critical trials = repetition trials) than any of the other run conditions (critical trials = switch trials) at each test third, confirming the presence of a RT switch effect. Similar to the RT analyses performed on participants' median RTs, the 4-trial run condition was slowest at each third. However, given the large variance in performance and the relatively small difference between the experimental run conditions compared to the large difference between the experimental

⁴ Practice effects on task switching performance will be examined more purposefully in Chapter 8.

and control run conditions (i.e., the switch effect), it is difficult to interpret this as support for a run length effect.

With regard to the effects of practice within the experiment, the mean RTs for critical recognition memory decisions trended downwards across the thirds in a linear fashion for all of the run conditions. More importantly, the decrease in mean RT for critical recognition memory decisions across test thirds appeared larger for the 1-, 2-, and 4-trial run conditions (where critical trials were switch trials) than that for the control run condition (where critical trials were repetition trials). The 1-, 2-, and 4-trial run conditions did not seem to differ largely from one another in the size of the decrease across test thirds. It may be that there were too few trials and that performance was too variable to observe any sizable differences across run lengths. Thus, Experiment 1 provided at least some visual confirmation of practice effects. Not only did experience with the recognition task speed correct recognition responses, but this improvement was somewhat larger for switch trials than that for repetition trials.

Discussion

Experiment 1 examined the effect of a task switch on recognition memory performance when participants alternated between a recognition memory task and a magnitude task in 1-, 2-, or 4-trial runs. Mean RTs for correct recognition responses on critical trials were significantly slower for switch trials (1-, 2-, and 4-trial run conditions) compared to repetition trials (control condition). That is, consistent with previous research (Rogers & Monsell, 1995; Jones et al., 2011), there was a significant RT switch effect when participants switched between tasks with limited S-R overlap.

However, recognition memory accuracy did not show a significant switch effect. Unlike in Jones et al. (2011), participants' ability to discriminate between old and new words at test was not significantly affected by a task switch. This failure to produce a statistically significant difference in recognition memory accuracy may be the result of the method used to examine it. The critical comparison of switch and repetition trials was made between participants, and as such could have lacked statistical power. Indeed, accuracy in the 2- and 4-trial run conditions was numerically lower for critical trials than that for the control condition (3% and 7% lower respectively). It is also worth noting that the 1-trial run condition—where accuracy for critical (switch) trials was numerically higher than that for critical (repetition) trials in the control condition—is essentially a mixed block (Jersild, 1927). There is evidence that mixed blocks using

different stimulus sets for each task do not always show mixing or switch effects (Jersild; Allport et al., 1994).

When the effect of run condition was analysed using participants' median RTs, the trend towards increasing switch RTs with longer run lengths was not significant. Despite the Cognitive Control Model (Altmann & Gray, 2008) of task switching proposing that task sets decay at a steady rate, the length of time (i.e., number of trials) since a task set was last used did not significantly affect recognition memory performance on switch trials. This finding suggests that either the proposed decay of task set activation did not occur, or it was extremely slow such that the longest run length (4 trials) was not long enough to observe significant run length effects. It is, however, worth noting that this failure to observe a significant run length effect is consistent with Monsell et al. (2003), who also reported no significant effect of run length on switch RTs.

Finally, a visual comparison of critical recognition memory RTs across experiment thirds showed that overall RT performance improved as participants gained experience within the experiment. Consistent with a practice effect (e.g., Koch, 2005; Minear & Shah, 2008), the size of this improvement from the first to final thirds appeared larger for switch trials than that for repetition trials. The amount or frequency of task switches had no sizable impact on task switching performance or its improvement across experiment thirds.

Summary

In Experiment 1, significant RT switch effects were observed for recognition memory performance but an accuracy switch effect was not observed. Finally, a visual comparison of critical recognition decision RTs across experiment thirds showed evidence of practice effects.

Chapter Six

Experiment 2: Task Switching and Level of Processing Effects

Given that Jones et al. (2011) reported a sizable interruption effect on recognition accuracy, the non-significant accuracy switch effect in Experiment 1—particularly in the 2- and 4-trial run conditions—may have been due to a lack of statistical power from comparing performance between participants. Experiment 2 addressed this concern by using a within-participants design and comparing task switch and task repetition trials to assess the effect of task switching on recognition memory performance. However, the interest in task switching effects on recognition performance went beyond establishing that there is a cost to participants' ability to discriminate between old and new words. A number of manipulations at encoding and retrieval are known to differentially affect recollection or familiarity processes (see Yonelinas, 2002 for a review). Of particular note, divided attention at retrieval has been shown to affect recollection more than familiarity (e.g., Hicks & Marsh, 2000; Jacoby, 1991; Jones & Jacoby, 2001). As such, it is important to establish if specific memory processes are differentially affected by task switching.

As reported in the introduction, Craik and Lockhart (1972) asserted that memory for a specific item was largely determined by how it was studied or 'processed'. That is, encoding conditions that encouraged elaborative (deep) processes, such as judgements of semantic meaning (e.g., 'is the word a type of fish?') produced superior memory performance at test compared to encoding conditions that encouraged less elaborative (shallow) processes, such as judgements of perceptual form (e.g., 'is the word presented in upper-case letters'). Craik and Tulving (1985) obtained level of processing (LOP) effects across a range of memory tests (e.g., recall, recognition) and experimental situations (e.g., laboratory, classrooms). Memory was superior for deeply studied words compared to shallowly studied words.

Even today, the level-of-processing (LOP) effect remains an important and influential finding in cognitive psychology (see Craik, 2002; Roediger, Gallo & Geraci, 2002; Watkins, 2002). The LOP effect is important to the current experiment because research has consistently reported that, relative to shallow encoding, deep encoding strategies not only improve memory accuracy, but also increase the use of controlled

recollection to make recognition decisions (e.g., Gardiner, 1988; Rajaram, 1993; Yonelinas et al., 1996; Boldini, Russo & Avons, 2004). The increase in production of recollection-based recognition responses elicited by deep levels of processing has also been reported to be relatively large in comparison to other encoding manipulations such as study duration (Yonelinas, 2002).

Although LOP effects are considered very robust, they are not always observed (e.g., Morris et al., 1977) and can be eliminated (Boldini et al., 2004). For example, LOP effects can be eliminated by controlling participants' ability to access/use recollection by using a response signal delay (RSD) with a 400-ms response window. In a RSD condition, participants are instructed to make their recognition responses after a specified interval of time has passed after stimulus onset. Boldini et al. examined the effects of a LOP manipulation and the effects of matching study and test modality (i.e., encoding specificity) across 7 RSDs ranging from 100 ms to 3000 ms. The study-test modality and level of processing variables were manipulated within participants, and the RSD at test variable was manipulated between participants. In the study phase, words ($N = 80$) were presented one at a time visually or aurally. Participants rated the pleasantness of a word on a three-point scale (deep encoding condition) or repeated the word aloud (shallow encoding condition). In the test phase, participants made recognition memory decisions to studied and unstudied words ($N = 160$). A response signal (asterisks above and below a word) was presented to participants after the RSD had passed, signalling that participants had a 400 ms window to make a response. Participants were instructed to respond using the keyboard if they had seen or heard the presented word during the study phase and to withhold any response if they had not (a go/no-go paradigm).

Boldini et al. (2004) reported that matching study-test modalities produced superior recognition accuracy than that for different study-test modalities at shorter RSDs (< 300 ms) but not at longer RSDs (> 300 ms). An initial difference in familiarity produced superior performance when modalities matched at shorter RSDs but this advantage was obscured by the contribution from recollection at longer RSDs. More important was the finding that, although overall recognition accuracy was always above chance, deep encoding produced superior recognition accuracy than that for shallow encoding at longer RSDs (> 300 ms) but not shorter RSDs (< 300 ms). Boldini et al. interpreted these findings as evidence that LOP effects were driven by the contribution of recollection, which took time to access. For shorter RSDs there was insufficient time

for recollection to be used when making recognition decisions, and thus no LOP effect was produced. By contrast, for longer RSDs, the slow, consciously controlled recollection process was available for use and produced a large LOP effect.

Evidence from Boldini et al. (2004) showed that the controlled recollection process took time to access and utilise so that it could contribute to performance in the deep LOP condition. If switching tasks is a time consuming and cognitively demanding process—which is arguably the position of most task switching theories—then it stands to reason that the recollection process may be affected by a task switch. Therefore, the purpose of Experiment 2 was to examine task switching effects on recognition memory accuracy, with a particular interest in whether participants' ability to use recollection was affected by a task switch. In the study phases of Experiment 2, participants alternated between identifying words as representing living or non-living things (deep encoding) and reporting whether words contained the letter 'e' (shallow encoding). During test phases, participants switched between making recognition memory decisions to words and magnitude judgements to dots in 2-trial alternating runs.

Three key research interests produced three hypotheses for Experiment 2. First, it was hypothesised that accuracy would be lower and RTs would be slower for switch trials compared to those for repetition trials, reflecting accuracy and RT switch effects respectively. Second, it was hypothesised that, at least for repetition trials, a LOP effect would be observed. Deep encoding at study should produce higher recognition accuracy than shallow encoding at study. Taking into account the findings of Boldini et al. (2004) and the idea that divided attention is argued to have a greater impact on recollection than familiarity (Yonelinas, 2002), there is reason to expect task switching to affect accuracy for words studied in the deep encoding condition more than that for words studied in the shallow encoding condition. Thus, the final hypothesis for Experiment 2 was that the LOP effect would be reduced or eliminated on task switch trials, reflecting impaired recollection processes. That is, task transition (at test) and level of processing (at encoding) would interact.

Method

Participants

The participants were 32 undergraduate psychology students from Victoria University of Wellington who: (1) had normal or corrected-to-normal vision, (2) had spoken English since early childhood, and (3) met minimum performance requirements.

The performance requirements were: (a) above chance repetition trial accuracy for the magnitude task and the deep encoding recognition condition, (b) mean RTs for correct trials within 3 *SD* of the group mean, and (c) fewer than 20% timeouts for either task. Participants who failed to meet the above criteria were replaced to achieve proper counterbalancing. Remuneration was credit towards a research requirement of an introductory psychology class.

Materials

Concrete nouns ($N = 150$) from the Toronto Word Pool (Friendly, Franklin, Hoffman & Rubin, 1982) were used to create 16 lists of 8 words (see Appendix A). Lists were closely matched for syllables ($M = 2.00$), length ($M = 6.03$), imagery ($M = 5.66$), concreteness ($M = 6.27$), and frequency (Kucera & Francis, 1967, $M = 11.54$). Lists were constructed so that the correct responses to deep and shallow encoding judgements were roughly equal (e.g., approximately half of words in a list had no letter 'e'). The 16 lists of critical words were separated into 2 groups of 8 lists (lists A-H, and lists I-P). Each group was used for one study-test block (i.e., lists A-H for study-test 1, lists I-P for study-test 2). To be clear, while word lists cycled through conditions *within* a study-test block, they did not cycle between study-test blocks. The allocation of list group to the first or second study-test block was balanced across participants. Buffer and filler words ($N = 22$) had similar linguistic characteristics to list words. Words were presented centrally in black bold 24-pt Courier New font on a white background. Dot stimuli were black circles (diameter = 8 mm). The experiment was conducted using E-Prime software (Schneider et al., 2002).

Design

The critical manipulation in Experiment 2 was the level of processing (LOP) at encoding/study. At test, the recognition memory task was manipulated in a 3 (item type: old-deep encoding, old-shallow encoding, new) x 2 (transition: repetition, switch) repeated measures design. The magnitude task was manipulated in a 2 (item type: more, less) by 2 (transition: repetition, switch) repeated measures design.

Procedure

Participants sat approximately 0.5 meters from the monitor at individual computer stations. One to four participants were run per session. The experiment was conducted in a single session that comprised two study-test blocks. The choice to use

two blocks was driven by a desire to increase accuracy and reduce participant fatigue. In each study phase, participants studied words under deep and shallow encoding conditions. In each test phase, participants switched between the recognition memory task and the magnitude task in 2-trial alternating runs.

Study phases.

Each study phase consisted of 42 trials: 5 primacy buffer words, 32 words selected in an alternating pattern from 4 lists (for each level of processing: 1 list for repetition transitions at test and 1 list for switch transitions at test), and 5 recency buffer words. Participants were instructed to make a level of processing decision for each word presented and to study each word for a later memory test. That is, learning in this experiment was intentional. Although Lockhart and Craik (1990) preferred the use of incidental learning conditions for LOP manipulations, the effect has been reported in experiments that use intentional learning conditions (see Roediger et al., 2002; Postman & Kruesi, 1977; Hogan, Kelly & Craik, 2006). The deep encoding trials required participants to decide if the word represented a living or non-living thing and to respond using the *f* and *j* keys respectively. Shallow encoding trials required participants to decide if the word contained the letter 'e' and to respond using the *f* and *j* keys respectively.

A study trial began with the presentation of a brief (100 ms) focal stimulus (+++) in the centre of the monitor. Immediately after the focal stimulus a word was presented for 3000 ms in the same location. Participants alternated between the deep and shallow encoding trials, aided by the presentation of response statements (i.e., 'living' and 'non-living', or 'e', and 'no e') below the word. The spatial location of the response options corresponded to the location of the *f* or *j* response. After the 3000-ms word presentation was complete, an ITI with a blank screen occurred for 500 ms. Although alternating between deep and shallow encoding conditions provided an element of task switching during encoding/study, previous research has reported that switching between deep and shallow LOP at study has no effect on recognition performance at test (Hogan et al., 2006).

Test phases.

Each of the two test phases comprised a practice test and the actual test. Participants completed a recognition memory task (R) and a magnitude task (M) in a 2-trial alternating runs sequence (RR MM RR MM...). The practice test consisted of 10

trials, starting with the magnitude task. Recognition trials in the practice test used buffer words from the study phase or new words otherwise not presented in the experiment. The actual test consisted of 130 trials, again starting with the magnitude task. For the recognition memory task, old target words were from the four lists presented in the immediately preceding study phase. Four lists not previously presented were used for new words. One deep encoding list, one shallow encoding list, and two ‘new’ lists were used for each position (position 1, position 2) in a run of the recognition task. The allocation of lists to item status and position within each test phase was counterbalanced across participants. For each position in a run of the magnitude task, half of the trials contained eight dots (‘more’ trials), and half of the trials contained six dots (‘less’ trials). All iterations of ‘more’ and ‘less’ trials within a 2-trial run were presented pseudo-randomly throughout the actual test such that correct magnitude judgements were not predictable within or across runs and were not useful to predict the status of an upcoming recognition trial.

A recognition trial began with the brief (100 ms) presentation of a focal stimulus (+++) in the centre of the monitor. Immediately after the focal stimulus the target word was presented in the same location. A response delay of 400 ms passed before 8 asterisks (*****) appeared above and below the target word, indicating that old/new judgements could be entered using the *f* and *j* keys on the keyboard. The allocation of keys to responses was balanced across participants. The response delay prevented participants from responding prematurely and increased the availability of recollection (Boldini et al. 2004). After 1000 ms, exclamation marks (!!!!!!!) replaced the asterisks for 600 ms to warn participants that an immediate response was required to avoid a timeout. This response deadline was used in order to constrain RTs and to reduce variance in participants’ performance. The trial terminated when a response was made or the warning period ended. The next trial began after a 500-ms ITI of a blank screen. A magnitude trial began with the brief (100 ms) presentation of a focal stimulus (+++) in the centre of the monitor. Immediately after the focal stimulus the dot targets (6 or 8 dots) were presented randomly within a central 100 mm (length) x 70 mm (height) section of the display with the proviso that dots did not overlap. Participants were instructed to judge whether there were more or less than seven dots and to respond using the *f* (more) and *j* (less) keys on the keyboard. As in Experiment 1, the allocation of key to magnitude judgement response was not balanced. When a response was made (or 2000 ms had elapsed), the trial was terminated. The next trial began after a 500-ms

Table 6.1.

Means and Standard Deviations for Corrected Recognition Accuracy in Experiment 2 as a Function of Test Block, Level of Processing at Study, and Task Transition.

	Test Block			
	Test 1		Test 2	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Deep				
Repetition	0.59	0.24	0.68	0.22
Switch	0.52	0.31	0.60	0.25
Shallow				
Repetition	0.45	0.28	0.58	0.21
Switch	0.37	0.30	0.50	0.20

ITI of a blank screen.

Full instructions were presented prior to each test phase. After receiving instructions about the two tasks and the structure of the test phase, participants completed a practice test. Participants were then reminded about the response allocations for the magnitude and recognition tasks prior to starting the actual test. Participants were debriefed and thanked for their participation after completing the experiment.

Results

The analyses reported in Experiment 2 focused on performance on the recognition memory task. Data from the practice tests and the first two trials of the actual tests were excluded from analyses. Statistical analyses of RT performance used participants' median RTs for correct trials in each condition. An alpha level of 5% was used for all statistical tests.

Corrected Accuracy

Corrected accuracy (hit rate – false alarm rate) was calculated at a participant level for each level of processing (LOP) at encoding, task transition, and test block. These data are reported in Table 6.1. Across both test blocks and LOP conditions the

corrected accuracy of switch trials was lower than that of repetition trials, demonstrating an accuracy switch effect. The size of this switch effect showed no noticeable difference for deeply or shallowly encoded words. Corrected accuracy showed typical LOP effects, with higher accuracy for the deep encoding condition than that for the shallow encoding condition. The size of this LOP effect was very similar for repetition and switch trials within each test. Finally, the data also showed a general improvement in accuracy from Test 1 to Test 2, particularly for the shallow encoding condition.

Switch effects and LOP effects were examined using a 2 (test block) x 2 (LOP) x 2 (transition) repeated measures ANOVA. Demonstrating an accuracy switch effect, the mean corrected accuracy for repetition trials ($M = 0.57$, $SD = 0.18$) was significantly higher than that for switch trials ($M = 0.50$, $SD = 0.21$), $F(1, 31) = 7.42$, $MSE = 0.05$, $p = 0.011$, $\eta_p^2 = 0.19$. Participants' ability to correctly identify words as old or new was significantly impaired following a task switch. The mean corrected accuracy for semantically encoded words (deep LOP: $M = 0.60$, $SD = 0.20$) was significantly higher than that for perceptually encoded words (shallow LOP: $M = 0.48$, $SD = 0.17$), $F(1, 31) = 29.80$, $MSE = 0.03$, $p < 0.001$, $\eta_p^2 = 0.49$. That is, there was a significant LOP effect. However, as can be seen in Figure 6.1, the size of the LOP effect was similar for repetition and switch trials: the interaction of LOP and transition was not significant

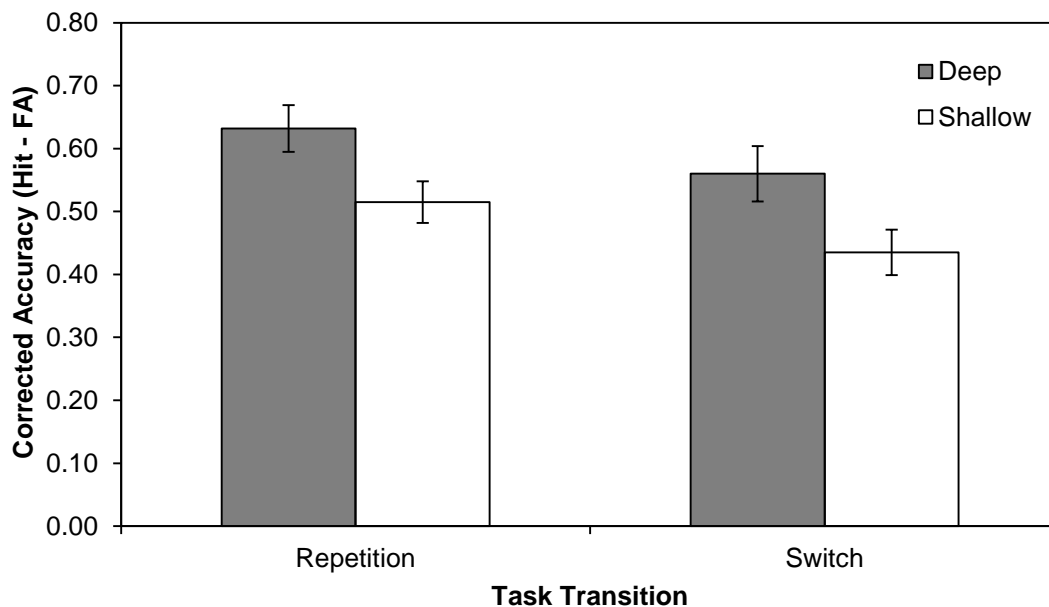


Figure 6.1. Mean corrected accuracy of recognition decisions in Experiment 2 as a function of task transition and level of processing at study. Error bars represent S.E.

($F = 0.04$, $p = 0.849$). If switching tasks influenced familiarity and recollection processes of recognition memory differently, an interaction of transition and LOP would be expected. Specifically, it was hypothesised that while an LOP effect would be observed for repetition trials it would be reduced or eliminated for switch trials. Contrary to this hypothesis, although a task switch reduced accuracy of recognition decisions, the level of processing effect remained unchanged.

A significant effect of test block [$F(1, 31) = 10.15$, $MSE = 0.07$, $p = 0.003$, $\eta_p^2 = 0.25$] indicated that participants mean corrected accuracy improved from the first test ($M = 0.47$, $SD = 0.24$) to the second test ($M = 0.58$, $SD = 0.16$). Performance on the recognition memory task improved as participants had more experience with the task and the experimental setting. This improvement was general in nature, and equivalent in size for repetition and switch trials. That is, test block and task transition did not interact significantly ($F = 0.01$, $p = 0.924$). No other effects or interactions were significant ($F_s < 1.17$).

Recognition accuracy was also analysed using measure A' . These analyses produced the same results as corrected accuracy and are included for the sake of completeness in Appendix C.

Reaction Times

RT analyses used correct recognition memory responses, resulting in the exclusion of 24% of the data. As was noted in Experiment 1, given the relative difficulty of the recognition task compared to typical tasks used to investigate task switching effects, this exclusion rate was not concerning. The percentage of timeouts was very low (1%), indicating that participants were capable of making their recognition decision within the response window used in this experiment. The mean RTs for correct recognition decisions are reported in Table 6.2 as a function of test block, item type (old-deep encoding, old-shallow encoding, new), and task transition. Reported RTs do not include the 400-ms response delay that occurred after stimulus presentation. Across all item types and test blocks, correct RTs for switch trials were longer than those for repetition trials, showing RT switch effects for the recognition task. For repetition trials in particular, RTs were very similar across item status. That is, participants took a similar amount of time to correctly identify words studied in the deep and shallow encoding conditions as well as new words. Performance on switch trials showed more variability across item types. Finally, an obvious decrease in correct

Table 6.2

Mean and Standard Deviation for Correct Recognition Decisions in Experiment 2 as a Function of Test Block, Item Type, and Task Transition.

	Test Block			
	Test 1		Test 2	
	M	SD	M	SD
Deep (old)				
Repetition	404	105	330	99
Switch	499	166	511	172
Shallow (old)				
Repetition	396	147	360	93
Switch	552	177	497	157
New				
Repetition	406	123	360	118
Switch	517	162	474	174

RTs from Test 1 to Test 2 occurred, particularly for repetition trials. Correct recognition memory decisions were made more quickly when participants had gained experience with the task.

Switch effects and LOP effects on correct recognition RTs were assessed using a 2 (test block) x 3 (item type) x 2 (transition) repeated measures ANOVA. Participants were slower to make correct recognition decisions when the preceding trial was a magnitude task ($M = 508.10$, $SD = 113.98$) than when it was a recognition task ($M = 376.01$, $SD = 84.69$), $F(1, 31) = 73.49$, $MSE = 22789.42$, $p < 0.001$, $\eta_p^2 = 0.70$. That is, an RT switch effect was observed. The main effect of item type was not significant ($F = 0.29$, $p = 0.747$), nor was the interaction of item type and transition ($F = 0.18$, $p = 0.836$). Depth of encoding at study did not affect the time to make a correct recognition decision or the RT cost of switching tasks.

Mean RTs for correct recognition responses showed a benefit of practice, reflected in a significant main effect of test block, $F(1, 31) = 7.27$, $MSE = 21572.06$, $p = 0.011$, $\eta_p^2 = 0.19$. Collapsed across repetition and switch trials, participants made

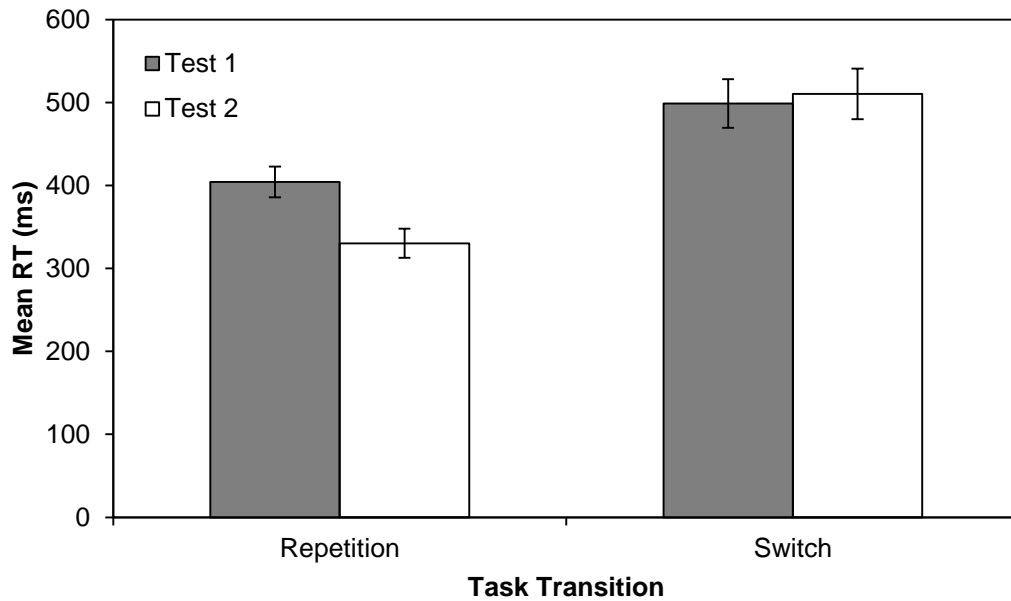


Figure 6.2. Mean RTs for correct recognition memory decisions for deeply encoded words in Experiment 2 as a function of task transition and test block. Error bars represent S.E.

correct recognition decisions more quickly in the second test ($M = 421.85$, $SD = 105.17$) than that in the first test ($M = 462.26$, $SD = 108.41$). The main effect was qualified by a significant 3-way interaction of test block, item type, and transition, $F(1, 31) = 4.43$, $MSE = 744.35$, $p = 0.016$, $\eta_p^2 = 0.13$. The 3-way interaction was driven by a significant interaction of transition and test block for deeply encoded words [$F(1, 31) = 6.91$, $MSE = 8478.32$, $p = 0.013$, $\eta_p^2 = 0.18$] but not shallowly encoded words [$F = 1.25$, $p = 0.273$] or new words [$F = 0.64$, $p = 0.430$]. The interaction of test block and transition for deeply encoded words is illustrated in Figure 6.2. Deeply encoded old words showed an effect of test block for repetition trials [$t(31) = 3.28$, $MSE = 22.58$, $p = 0.003$] but not for switch trials [$t(31) = -0.39$, $MSE = 30.00$, $p = 0.701$]. Most curiously, the time taken to make a correct recognition decision on switch trials showed no improvement from the first to second test block for deeply encoded old words. This result, in stark opposition to traditional practice effects (e.g., Jersild, 1927; Merian, 1996; Rogers & Monell, 1995), is difficult to explain from a theoretical standpoint. No other effects or interactions were significant ($F_s < 0.92$).

Discussion

Experiment 2 examined task switching effects on recognition memory performance when words were studied under deep or shallow encoding conditions.

There were three critical results in this experiment: a LOP effect, task switching effects, and a (non-significant) interaction of LOP and task transition. First, the level of processing manipulation was effective despite using intentional learning conditions. Consistent with prior research (Boldini et al. 2004; Craik & Tulving, 1975; Hogan et al., 2006; Postman & Kruesi, 1977; Rajaram, 1993), deeply encoded words were identified as ‘old’ more accurately during the test phase than shallowly encoded words.⁵ The LOP effect in this experiment was smaller than those reported for recognition performance in Craik and Tulving. This difference may have been the consequence of the different methods used to test recognition memory. Craik and Tulving used an incidental learning paradigm and allowed participants unlimited time to respond to each test word, whereas the current experiment used an intentional learning paradigm and used a response delay and a response deadline (i.e., a response window). When considering the size of the LOP effects reported in this thesis, perhaps the most valid point of comparison is to the results of Boldini et al. The LOP effects reported in this thesis (400-ms RSD, range 0.10 – 0.15) were very similar in size to the LOP effect reported for the 300-ms RSD condition in Boldini et al. (0.12). Thus despite the frequent task switches in the current experiment, the reported LOP effects were consistent with previous literature.

There were significant task switching effects on recognition memory RTs and accuracy. Participants’ recognition memory responses were slower and less accurate on task switch trials than they were for task repetition trials. The reported RT switch effects on recognition memory were similar to those reported in Jones et al. (2001), but smaller than those reported in Experiment 1 of this thesis. This difference could in part reflect the change from using an unlimited response time to using a response window that constrained responding (and thus, maximum switch RTs). Unlike in Experiment 1, the current experiment reported significantly lower corrected accuracy for switch trials than that for repetition trials. That is, participants’ ability to correctly identify old words and reject new words was impaired when they had just switched tasks. Increasing the number of critical trials and changing to a within-participant comparison of task switch and task repetition performance likely increased experimental power and sensitivity to detect switch effects on recognition accuracy. The accuracy switch costs reported in Experiment 2 (range 0.07 – 0.09) were within the range reported in Jones et al. (0.07 –

⁵ The mean RTs for correct recognition responses (repetition and switch trials) were very similar for deeply and shallowly encoded words, indicating that differences in accuracy were not the result of a speed/accuracy trade-off.

0.14), albeit at the lower end. The smaller accuracy switch costs reported in the current experiment may reflect the advantage of frequent and predictable task switches over the relatively infrequent and unpredictable interruptions in Jones et al. It is also worth noting that recognition trials in Jones et al. also served as encoding events (i.e., it was a continuous recognition paradigm), which may have inflated costs to recognition performance.

The hypothesised interaction of LOP at study and task transition at test was based on the assumption that, like divided attention at retrieval and short RSDs, task switching would affect the availability and use of recollection processes. As previous research has shown that, relative to shallow encoding, deep encoding produces a higher proportion of recognition decisions based on recollection (e.g., Gardiner, 1988; Rajaram, 1993; Yonelinas et al., 1996), deep encoding should also show a greater task switching effect. That is, it was hypothesised that the LOP effect would be reduced or eliminated for task switch trials. However, the interaction was not significant—the size of the switch effect on corrected recognition accuracy was equivalent for deeply encoded and shallowly encoded words. Phrased in terms of LOP effects, the size of the performance advantage for deeply encoded words relative to shallowly encoded words was the same regardless of task transition. Thus, there was no evidence to support the hypothesised interaction.

As Experiment 2 produced a LOP effect, a dual process account of recognition memory would suggest that recollection contributed to recognition accuracy, particularly for deeply encoded words (see Boldini et al., 2004). In dual process signal detection models of recognition memory, the probability of correctly recognising a studied item as old is the sum of the probability of the item being recollected and the probability of the item's familiarity exceeding a decision criterion (Jacoby, 1991; Yonelinas, 1994; 2000b). Thus, if the size of the LOP effect is used as an approximate index of recollection, then the contribution of recollection to recognition accuracy did not change between task switch and task repetition trials.

It is worth noting that although dual process theorists emphasise the larger role of recollection in producing LOP effects, they acknowledge that changes in familiarity also play a (relatively small) part in increasing recognition accuracy (see Yonelinas, 2002). Thus, the LOP effect observed in this experiment may reflect differences in memory trace strength (i.e., familiarity) between the two encoding conditions, without any contribution from recollection. The limited (or lack of) contribution from

recollection could explain why the LOP effect reported in this experiment was smaller than those reported in Craik and Tulving (1975) or at the longer RSD conditions in Boldini et al. (2004).⁶

From either of these perspectives (recollection not affected, recollection not contributing), the lower corrected accuracy for switch trials compared to that for repetition trials suggests that familiarity (or memory strength) was affected when switching tasks. While not congruent with the divided attention literature (Jacoby, 1991; Hicks & Marsh, 2000; Jones & Jacoby, 2001; Knott & Dewhurst, 2007a), the isolation of switch effects to familiarity could be reconciled theoretically. Familiarity is a fast and automatic process that is initiated with the onset of a target stimulus (Yonelinas, 2002). Thus, evidence of familiarity/memory trace strength is being accrued and evaluated during a period of time where a participant is likely to be devoting significant resources to retrieving task sets from long term memory (see Mayr & Kliegl, 2000) and resolving instances of interference (Allport et al., 1994; Allport & Wylie, 2000). That is, the process(es) of switching tasks may have created a bottleneck (see Pashler, 1994) that stalled or disrupted the evaluation of a word's familiarity. As a consequence, participants' ability to discriminate between signal plus noise ('old') and noise ('new') levels of familiarity was impeded by a recent task switch.

Summary

Experiment 2 manipulated level of processing at study and task transition at test to further investigate task switching effects on recognition memory performance. A level of processing effect was observed for recognition accuracy, and task switching effects were reported for recognition accuracy and RTs. The size of the level of processing effect was unchanged by task transition. Contrary to the hypothesis that a task switch would affect recollection processes, the equal size of the accuracy switch effects for deeply and shallowly encoded words suggested that task switching affected the familiarity process.

⁶ Boldini et al. (2004) reported an LOP effect of 0.25 for the 3000-ms RSD condition—considerably larger than the size of the LOP effect reported in their 300-ms RSD condition or the LOP effects reported in this experiment.

Chapter Seven

Experiment 3: Recovery from a Task Switch and Transition Cuing

The experimental method developed to assess task switching effects on recognition memory in Experiments 1 and 2 has some unique characteristics that warrant further investigation. Because external cues were not provided, to optimise performance participants needed to monitor task sequences internally by maintaining a representation of the task sequence in working memory, which arguably required considerable cognitive resources. In this case, providing a visual cue to scaffold sequence monitoring (a ‘reminder’) may aid participants in monitoring the task sequence and decrease the size of the RT switch effect.

To this end, Experiment 3 used a 4-trial alternating runs paradigm with a between participant manipulation of sequence cue. For some participants, the standard focal stimulus presented prior to target onset (+++) was replaced with an alternative, novel focal stimulus (\$\$\$ presented in red) on a single trial within the run. This novel focal stimulus occurred at Position 4 in a run (i.e., the last trial in a run) or Position 1 in a run (i.e., the first trial in a run). A final group of participants always saw the standard focal stimulus. Unlike cues in the explicit task cuing paradigm, the novel focal stimulus did not cue the task to be performed (i.e., the cues were not task specific). Instead, it provided an external cue to remind participants of their position within the run of a task and cue an upcoming change in task. That is, it utilised a modified transition cuing paradigm (e.g., Schneider & Logan, 2007). Beyond providing a reminder of one’s position within a run of a task, the sequence cue (particularly when presented at Position 4) could encourage participants to monitor the task sequence and to prepare for a task switch before it occurs.

The use of a 4-trial alternating runs design also allowed for detailed analysis of performance on repetition trials of a run. Of specific interest were the time course of recovery from a task switch (i.e., Monsell et al., 2003; Jones et al., 2011) and evidence for slowing within a run (i.e., Altmann, 2002; Altmann & Gray, 2002). With regard to recovery from a task switch, Monsell et al. (2003) reported that when run length was predictable, the cost of switching tasks was isolated to Position 1 (switch trial) of a run, and subsequent repetition trials showed no further decrease (or increase) in RT. When

run length was unpredictable, RTs showed a more gradual decrease from the switch trial across the first few repetition trials before RTs stabilised for the remaining positions in the run. Jones et al. reported RT costs that persisted to the second recognition memory trial following an interruption/task switch, even though participants knew that the interruption task would last only one trial and many trials would pass before the next interruption. With regard to within-run slowing, Altmann (2002) and Altmann and Gray (2002) have reported that repetition trial RTs showed a significant positive linear trend within a run, which was argued to reflect decay of the online task set activation. That is, the RT for Position 2 was significantly faster than that for Position 8.

In Experiment 3, participants performed the recognition memory and magnitude judgement tasks in 4-trial alternating runs. As the experiment manipulated variables that were equally applicable to the recognition and magnitude tasks, performance on both tasks was analysed. There were three key aims: (a) to examine task switching performance in the recognition and magnitude tasks, (b) to evaluate evidence of recovery and/or slowing within a run for each task, and (c) to investigate whether a visual reminder could decrease the size of task switching effects or influence the pattern of recovery from a task switch.

Consistent with Experiments 1 and 2, it was hypothesised that significant RT and accuracy switch effects would be observed for the recognition task. Furthermore, performance on the magnitude task should show similar switch effects. With regard to performance within runs of a task, although Jones et al. (2011) showed that RT switch costs could persist past the first recognition trial in a run, considerable differences in design may limit its applicability in the current experiment. Specifically, there were fewer interruptions and as participants did not know precisely when the next interruption would occur, the run length of the recognition task was unpredictable. Given that this experiment used a design with frequent and predictable task switches, Monsell et al.'s (2003) findings were used to develop hypotheses. First, RT and accuracy switch effects in Experiment 3 should be isolated to the first position in the run, with performance on the remaining trials showing no differences in performance. Second, providing a sequence cue should help participants monitor task sequences and prepare for task switches. Thus, RT switch effects should be smaller for conditions that use a sequence cue than that for the control condition that did not use a sequence cue.

Method

Participants

One-hundred-and-seventy undergraduates from Victoria University of Wellington participated in the experiment as part of an undergraduate cognitive psychology course, or for course credit in an introductory psychology course. All participants in the final data set: (1) had normal or corrected-to-normal vision, (2) spoke English from early childhood, and (3) met minimum performance requirements outlined in the results section. The performance requirements were: (a) above chance accuracy for repetition trials on the magnitude and recognition tasks, (b) mean RTs for correct trials within 3 *SD* of the group mean; (c) fewer than 20% timeouts for either task. Cases that met inclusion requirements were added to the final dataset chronologically until the maximum number of counterbalances was filled for each of the three sequence cue conditions (i.e., an equal number of participants and counterbalances for each sequence cue condition). Using this process, 120 participants' data (40 per sequence cue condition) were included in the final data set.

Materials

Nouns ($N = 147$) from the English Lexicon Project (Balota et al., 2007) were used to construct 16 lists of 8 words (see Appendix A). The 16 lists of critical words were separated into 2 groups of 8 lists (lists A-H, and lists I-P). Each group was used for one study-test block (i.e., lists A-H for study-test 1, lists I-P for study-test 2). To be clear, while word lists cycled through conditions *within* a study-test block, they did not cycle between study-test blocks. The allocation of list group (A-H, I-P) to test block was not balanced. Lists (and list groups) were balanced word frequency (Kucera & Francis, 1967; $M = 97.92$), word length ($M = 6.54$) and syllables ($M = 2.11$). Buffer words ($N = 20$) had similar linguistic properties. Word stimuli were presented centrally on a white screen in black bold 18-pt Courier New font. The sequence cue was three red dollar signs (\$\$\$) presented in red bold 18-pt Courier New font. Dot stimuli were black circles (diameter = 8 mm). The experiment was conducted using E-Prime software (Schneider et al., 2002).

Design

The critical manipulation in Experiment 3 was the between-participant manipulation of sequence cue. A novel focal point was used on Position 1 (P1 cue

condition), Position 4 (P4 cue condition), or not at all (control condition). The recognition memory task was manipulated in a 3 (sequence cue position: P1, P4, control) x 4 (run position: P1, P2, P3, P4) x 2 (item type: old, new) mixed design. The magnitude task was manipulated in a 3 (sequence cue position: P1, P4, control) x 4 (run position: P1, P2, P3, P4) x 2 (item type: less, more) mixed design.

Procedure

Prior to participants entering the testing room each computer was assigned to a sequence cue condition using latin square counterbalancing. Upon entering the room, individual participants sat at a computer station of their choice approximately 0.5 meters from the monitor. The opening screen of the computer programme was the same for all sequence cue conditions so that participants could not choose which cue condition to complete. The experiment was conducted in a single session that comprised two study-test blocks. In each study phase, participants studied a list of words for a later memory test. In each test phase, participants switched between the recognition memory task and the magnitude judgement task in 4-trial alternating runs.

Study phases.

Each study phase consisted of 40 trials: 4 primacy buffer words, 32 words selected in an alternating pattern from 4 lists, and 4 recency buffer words. Participants were instructed to study each word for a later memory test. A study trial began with the presentation of a brief (100 ms) focal stimulus (+++) in the centre of the monitor. Immediately after the focal stimulus a word was presented for 3000 ms in the same location. After the 3000-ms presentation was complete, an intertrial interval (ITI) with a blank screen occurred for 500 ms.

Test Phases.

Each of the two test phases consisted of a practice test and an actual test. Participants were instructed that they would perform a recognition memory task (R) and a magnitude judgement task (M) in a predictable, 4-trial alternating runs sequence (MMMM RRRR MMMM...). The practice test consisted of 12 trials, starting with the magnitude task. Recognition trials in the practice test used buffer words from the study phase or new words otherwise not presented in the experiment. The actual test consisted of 132 trials, again starting with magnitude task. Each of the four positions within a run of the recognition task used one 'old' list (presented during the immediately preceding

study phase) and one ‘new’ list (not previously presented in the experiment). As was mentioned in the materials subsection, the same eight lists were associated with a study-test block for all participants. However, the allocation of lists to item type (old or new) and position (P1, P2, P3, or P4) *within* a study-test block was counterbalanced across participants. For each position in a run of the magnitude task, half of the trials contained eight dots (‘more’ trials), and half of the trials contained six dots (‘less’ trials). All iterations of ‘more’ and ‘less’ trials within a 4-trial run were presented pseudo-randomly throughout the actual test such that correct magnitude judgements were not predictable within or across runs and were not useful to predict the status of an upcoming recognition trial.

A recognition memory trial began with the brief (100 ms) presentation of a focal stimulus (+++ or \$\$\$) in the centre of the monitor. Immediately after the focal stimulus the target word was presented in the same location. A response delay of 400 ms passed before 8 asterisks (*****) appeared above and below the target word, indicating that old/new judgements could be made. The response delay prevented participants from responding prematurely. Participants responded using the *f* (old) and *j* (new) keys on the keyboard. As the interest of this experiment was not differences between old and new items, and neither Experiment 1 nor 2 reported significant main effects of item type or interactions of item type and transition, the allocation of key to recognition decision was not balanced. After 600 ms, exclamation marks (!!!!!!!) replaced the asterisks for 400 ms, warning participants that they needed to make an immediate response as they were about to run out of time. The trial terminated when a response was made or the warning period ended. The next trial began after a 500-ms ITI of a blank screen. Note that the total length of a recognition memory trial was shorter (1400 ms) in the current experiment than that in Experiment 2 (2000 ms). Given that the mean RTs reported in Experiment 2 were considerably shorter than 2000 ms, this was not expected to be problematic.

The magnitude judgement task began with a brief (100 ms) presentation of a focal stimulus (+++ or \$\$\$) in the centre of the monitor. Immediately after the focal stimulus the dot targets (6 or 8 circles) were presented randomly within a central 100 mm (length) x 70 mm (height) section of the display with the proviso that dots did not overlap. Participants were instructed to judge whether there were more or less than seven dots and to respond using the *f* (more) and *j* (less) keys on the keyboard. As with earlier experiments in this thesis, the allocation of key to magnitude judgement was not

balanced. When a response was made (or 1400 ms had elapsed), the trial was terminated. The next trial began after a 500-ms ITI of a blank screen.

All participants were told that they could use the predictable pattern of tasks (i.e., the 4-trial runs) to know what task they would be performing next. In addition to this sequence information, participants in the P1 sequence cue condition and the P4 sequence cue condition were instructed that, for a specific run position, the normal focal stimulus (+++) would be replaced with a novel focal stimulus (\$\$\$) presented in red. They were informed that the novel focal stimulus could be used to remind them that they were on the first trial of a run (P1 cue condition) or the last trial of a run (P4 cue condition). The control sequence cue condition received no additional information.

Full instructions were presented prior to each test phase. After receiving instructions about the two tasks and the structure of the test phase, participants completed a practice test. Participants were then reminded about the response allocations for the magnitude and recognition tasks prior to starting the actual test. Participants were debriefed and thanked for their participation after completing the experiment.

Results

Data from the practice tests and the first four trials of the actual tests were excluded from analyses. Statistical analyses of RT performance used participants' median RTs for correct trials in each condition. An alpha level of 5% was used for all statistical tests.

Accuracy

Recognition Task.

Overall corrected accuracy (hit rate – false alarm rate) was moderate ($M = 0.48$, $SD = 0.19$), but lower than the accuracy reported in Experiment 2. Experiment 3 did not include any encoding instructions beyond studying words and used a shorter response window for recognition trials at test, thus this difference was not unexpected. Table 7.1 reports the mean corrected accuracy of recognition decisions as a function of run position and sequence cue condition. To reiterate, Position 1 was a task switch trial, and Positions 2, 3, and 4 were task repetition trials. For all sequence cue conditions, the corrected accuracy was lowest on the switch trial (Position 1) and steadily improved across the remaining positions of the run. That is, recognition accuracy performance showed a typical switch effect. The largest improvement in performance was between

Table 7.1.

Means and Standard Deviations for Corrected Recognition Accuracy in Experiment 3 as a Function of Position in Run and Sequence Cue Condition.

	P1		P2		P3		P4	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Sequence Cue								
Control	0.39	0.23	0.45	0.21	0.47	0.22	0.48	0.22
P1 Cue	0.46	0.23	0.50	0.21	0.52	0.22	0.54	0.18
P4 Cue	0.43	0.26	0.46	0.27	0.54	0.25	0.51	0.25

Note. P1 = position 1, P2 = position 2, P3 = position 3, P4 = position 4. Control = no sequence cue. P1 cue = sequence cue at position 1 of a run. P4 cue = sequence cue at position 4 of a run.

Positions 1 and 2, suggesting a rapid recovery from a task switch. The accuracy of the control condition (no external cue) was lowest for all four positions, suggesting that monitoring task sequence may have affected overall accuracy, and the presentation of a sequence cue may have made it easier to perform the recognition task.

Switch effects on recognition corrected accuracy were analysed using a 3 (sequence cue) x 2 (test block) x 4 (run position) mixed ANOVA. The Huynh-Feldt correction was used for statistical tests that included position because the assumption of sphericity was violated. There was a significant main effect of position [$F(2.96, 346.60) = 7.81$, $MSE = 0.05$, $p < 0.001$, $\eta_p^2 = 0.06$], illustrated in Figure 7.1. Pair-wise comparisons using Bonferroni corrections showed that corrected accuracy for Position 1 ($M = 0.43$, $SD = 0.24$) was significantly lower than that for Position 3 ($M = 0.51$, $SD = 0.23$) and Position 4 ($M = 0.51$, $SD = 0.22$) but did not differ from Position 2 ($M = 0.47$, $SD = 0.23$). Although accuracy appeared to trend upwards across the repetition positions of the run, the differences in recognition accuracy between Positions 2, 3, and 4 were not significant. That is, the task switch effect to recognition performance was largely isolated to the first (switch) trial in a run, with perhaps some slight persistence of the switch effect into Position 2. Both the main effect of test block ($F = 0.82$, $p = 0.367$) and the interaction of test block and run position ($F = 1.06$, $p = 0.366$) were not significant. Participants' ability to discriminate between old and new words was no

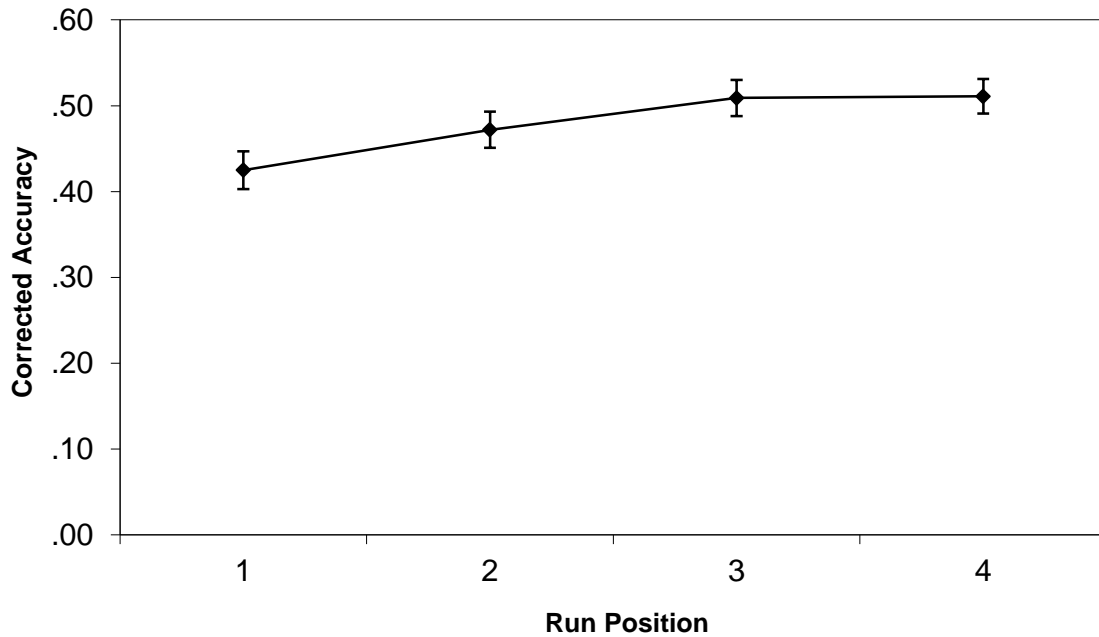


Figure 7.1. Mean corrected accuracy of recognition memory decisions in Experiment 3 as a function of run position. Error bars represent S.E.

different for the first and second test blocks.

Although corrected accuracy data hinted that the use of a sequence cue may have improved recognition memory performance at Position 1, the main effect of sequence cue was not significant ($F = 0.84$, $p = 0.434$), nor was its interaction with position ($F = 0.30$, $p = 0.933$). A planned Helmert contrast comparing the corrected accuracy of the control sequence cue condition to the combined accuracy of the P1 and P4 sequence cue conditions also confirmed that there was no significant difference between the sequence cue conditions ($p = 0.226$). Thus, the recognition accuracy data produced no statistically significant evidence to support the hypothesis that presenting a sequence cue reminder improved performance overall or on the switch trial.

Magnitude task.

Collapsed across all conditions, magnitude judgement accuracy was moderately high, but not near ceiling ($M = 0.83$, $SD = 0.11$). Participants were able to correctly judge whether more or less than seven dots presented. Table 7.2 reports the mean accuracy of magnitude judgements as a function of run position and sequence cue condition. Mean accuracy across the three sequence cue conditions was similar,

Table 7.2.

Means and Standard Deviations for Magnitude Judgement Accuracy in Experiment 3, as a Function of Position in Run and Sequence Cue Condition.

	P1		P2		P3		P4	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Cue Condition								
Control	0.76	0.15	0.84	0.14	0.83	0.13	0.83	0.12
P1 Cue	0.80	0.14	0.88	0.09	0.85	0.10	0.85	0.11
P4 Cue	0.80	0.13	0.84	0.12	0.87	0.09	0.87	0.11

Note. P1 = position 1, P2 = position 2, P3 = position 3, P4 = position 4. Control = no sequence cue. P1 cue = sequence cue at position 1 of a run. P4 cue = sequence cue at position 4 of a run.

suggesting that sequence cues did not improve the accuracy of magnitude judgements. For all cue conditions the mean accuracy was lower at Position 1 than that for remaining (repetition) positions in the run, demonstrating an accuracy switch effect. The accuracy of magnitude judgements for the three repetition positions were very similar, except in the P4 cue condition where Position 2 accuracy fell midway between Position 1 and Position 4, suggesting a more gradual recovery.

Switch effects on magnitude task accuracy were analysed using a 3 (sequence cue) x 2 (test block) x 4 (position) x 2 (item type) mixed design ANOVA. As the key allocation of dot decisions was not counterbalanced, effects and interactions of item type cannot be interpreted and therefore are not reported. The Huynh-Feldt correction was used for statistical tests that included position as the assumption of sphericity was violated. The significant main effect of position, [$F(2.64, 308.46) = 29.32$, $MSE = 0.02$, $p < 0.001$, $\eta_p^2 = 0.20$] is illustrated in Figure 7.2. Pair-wise comparisons using Bonferroni corrections showed that the mean magnitude task accuracy at Position 1 ($M = 0.79$, $SD = 0.14$) was significantly lower than that of Position 2 ($M = 0.85$, $SD = 0.12$), Position 3 ($M = 0.85$, $SD = 0.11$), and Position 4 ($M = 0.85$, $SD = 0.12$). Positions 2, 3, and 4 were not significantly different from one another.

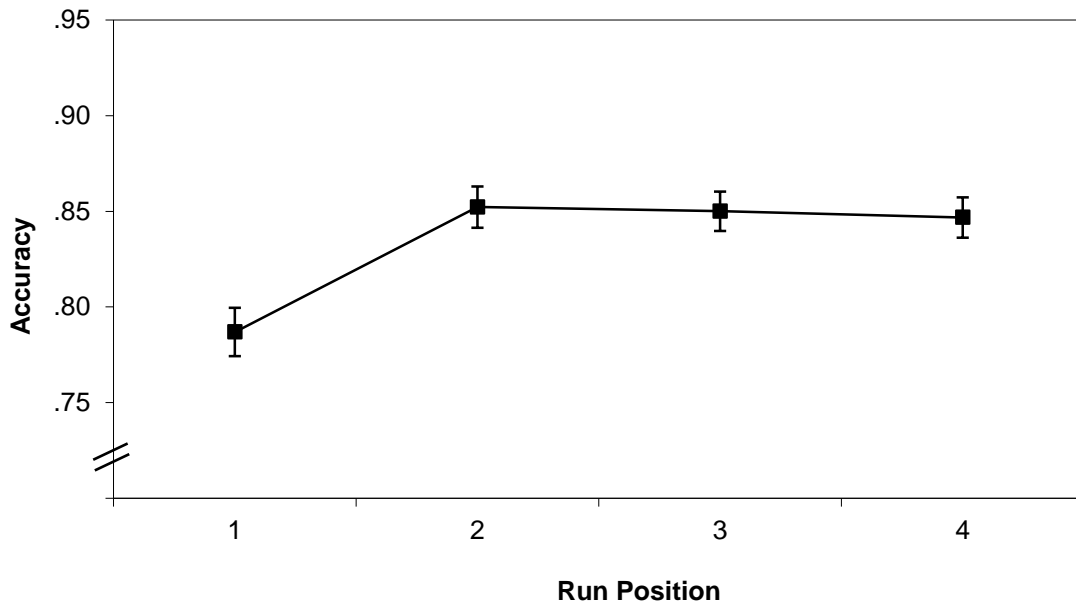


Figure 7.2. Mean accuracy of magnitude judgements in Experiment 3 as a function of run position. Error bars represent S.E.

The main effect of test block was significant, $F(1,117) = 14.52$, $MSE = 0.12$, $p < 0.001$, $\eta_p^2 = 0.11$. The mean accuracy of magnitude judgements in the first test ($M = 0.80$, $SD = 0.16$) was significantly lower than that of the second test ($M = 0.86$, $SD = 0.11$). The accuracy of participants' magnitude judgements improved as they gained experience with the task. The interaction of test block and position was also significant, $F(2.82, 329.70) = 3.46$, $MSE = 0.02$, $p = 0.021$, $\eta_p^2 = 0.03$. The mean magnitude judgement accuracy is shown in Table 7.3 as a function of run position and test block. A 3 (sequence cue) \times 2 (test block) \times 2 (position: P1, P2) \times 2 (item type) mixed ANOVA confirmed that the interaction was driven by a decrease in the size of the accuracy switch effect from the first to second test block, $F(1, 117) = 4.70$, $MSE = 0.02$, $p = 0.032$, $\eta_p^2 = 0.04$. The main effect of position was significant for the first [$F(1, 117) = 45.10$, $MSE = 0.02$, $p < 0.001$, $\eta_p^2 = 0.28$] and second [$F(1, 117) = 17.40$, $MSE = 0.02$, $p < 0.001$, $\eta_p^2 = 0.13$] test blocks, but the size of this difference between Position 1 and Position 2 accuracy was substantially smaller in the second test (0.02) than that in the first test (0.08). By contrast, A 3 (sequence cue) by 2 (test block) by 3 (position: P2, P3, P4) by 2 (item type) mixed ANOVA showed that the pattern of magnitude judgement accuracy within repetition trials of the run was no different for the first and second test blocks, $F(2, 234) = 0.38$, $MSE = 0.01$, $p = 0.684$, $\eta_p^2 = 0.00$. Accuracy of magnitude

Table 7.3

Means and Standard Deviations for Magnitude Judgement Accuracy in Experiment 3 as a Function of Position in Run and Test Block.

	P1		P2		P3		P4	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Test Block								
Test 1	0.74	0.19	0.83	0.19	0.82	0.17	0.83	0.17
Test 2	0.83	0.15	0.88	0.12	0.88	0.12	0.87	0.14

Note. P1 = position 1, P2 = position 2, P3 = position 3, P4 = position 4.

judgements showed a significant practice effect—the size of the accuracy switch effect was smaller after participants had completed the first test block.

The effect of sequence cue ($F = 0.99$, $p = 0.374$) and interaction of sequence cue and position ($F = 1.64$, $p = 0.149$) on magnitude judgement accuracy were not significant. Consistent with the recognition accuracy data, a planned Helmert contrast showed no significant difference in magnitude judgement accuracy between the control and experimental (P1, P4) cue conditions ($p = 0.162$). Consistent with the recognition accuracy data, the magnitude accuracy data showed no indication that the presentation of a sequence cue reminder aided performance.

Reaction Times

The RT analyses used correct recognition RTs and correct magnitude RTs, resulting in the exclusion of 27% and 17% of the data for the recognition and magnitude tasks respectively. The exclusion rate for the recognition task is similar to that reported in Experiments 1 and 2. The smaller exclusion rate for the magnitude task likely reflects the relative ease of performing this task compared to the recognition task. The percentage of timeouts for the recognition and magnitude tasks (2% for both tasks) was higher than that for Experiment 2 and was likely the consequence of the shorter response deadline used in Experiment 3. However, the rate of timeouts was still low.

Recognition Task.

Table 7.4 shows mean RTs for correct trials as a function of test, position in run, and sequence cue condition. Note that the reported recognition RTs do not include the

400-ms response delay. The mean RTs were very similar for all three sequence cue conditions and showed a large drop in RT from Position 1 to Position 2, demonstrating a RT switch effect. There was a very small decrease in RT from Position 2 to Position 3 for both tests, but more-so for the first test block. Relative to the drop in RT from Position 1 to Position 2, this decrease was almost negligible. Participants tended to make correct recognition decisions more quickly in the second test block. This decrease from the first to second test block was most sizable for Position 1 (the switch trial).

A 3 (sequence cue) x 2 (test block) x 4 (position) x 2 (item type) mixed ANOVA was conducted on correct recognition RTs. As the key allocation of recognition decisions was not balanced, effects and interactions of item type cannot be interpreted and therefore are not reported. The Greenhouse-Geisser correction was used for statistical tests that included position because the assumption of sphericity was

Table 7.4.

Means and Standard Deviations for Correct Recognition RTs in Experiment 3 as a Function of Position in Run, Test Block, and Sequence Cue Condition.

	P1		P2		P3		P4	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Test 1								
Control	448	198	313	145	289	137	298	138
P1 Cue	443	204	296	149	300	140	288	142
P4 Cue	455	201	313	147	290	138	287	140
Test 2								
Control	406	145	289	118	278	131	278	122
P1 Cue	385	149	264	121	263	134	275	125
P4 Cue	384	147	285	120	276	132	269	124

Note. P1 = position 1, P2 = position 2, P3 = position 3, P4 = position 4. Control = no sequence cue. P1 cue = sequence cue at position 1 of a run. P4 cue = sequence cue at position 4 of a run. Reported RTs do not include the 400 ms response delay.

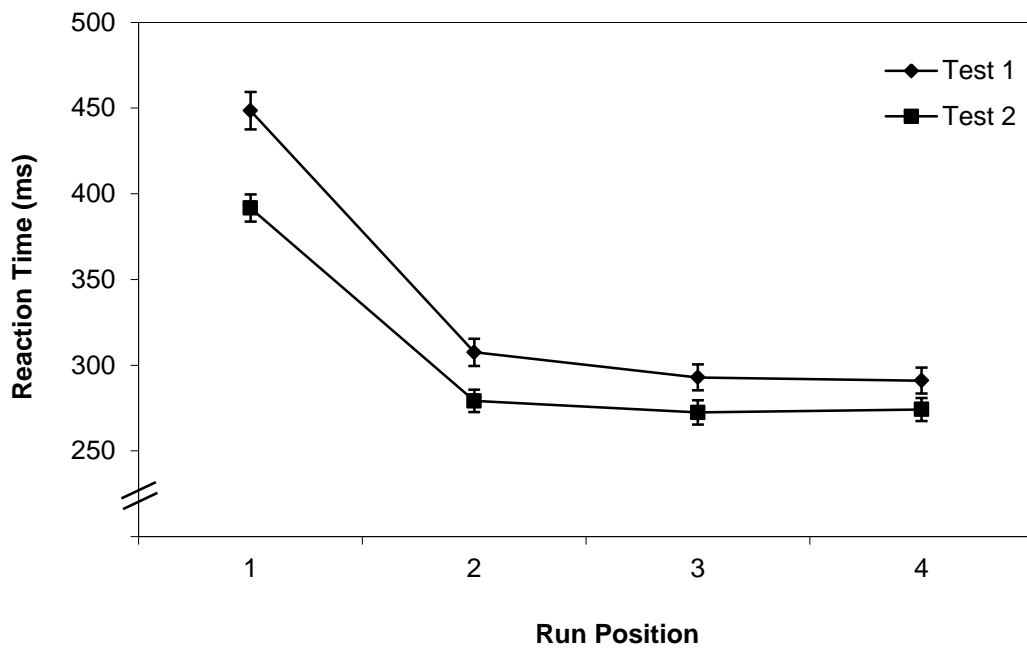


Figure 7.3. Mean RTs for correct recognition memory decisions in Experiment 3 as a function of run position and test block. Error bars represent S.E.

violated.⁷ There was a significant main effect of position, $F(1.85, 205.23) = 332.54$, $MSE = 10022.20$, $p < 0.001$, $\eta_p^2 = 0.75$. The mean RT for recognition decisions at Position 1 ($M = 420.11$, $SD = 87.10$) was significantly slower than that for Position 2 ($M = 293.36$, $SD = 64.38$), Position 3 ($M = 282.71$, $SD = 65.42$) and Position 4 ($M = 282.62$, $SD = 66.28$). The mean RT for correct recognition decisions at Position 2 was also significantly slower than that for Positions 3 and 4. Mean RTs for Position 3 and 4 were not significantly different from each other. That is, recognition RTs recovered from a task switch gradually over the next two repetition trials rather than immediately after a task switch.

The main effect of test block was significant, $F(1, 111) = 24.76$, $MSE = 17247.98$, $p < 0.001$, $\eta_p^2 = 0.18$. Participants' correct recognition RTs were faster in the second test ($M = 304.39$, $SD = 60.65$) than those in the first test ($M = 335.01$, $SD = 79.77$). As is illustrated in Figure 7.3, the interaction of position and test block was also significant, $F(2.72, 301.76) = 7.89$, $MSE = 5212.62$, $p < 0.01$, $\eta_p^2 = 0.07$. A 3 (sequence cue) \times 2 (test block) \times 2 (position: P1, P2) \times 2 (item type) mixed ANOVA confirmed

⁷ Following the recommendations of Field (2009), because the Mauchly's W was below 0.75 a Greenhouse-Geisser correction was used. The earlier analyses of position on recognition and magnitude accuracy used a Huynh-Feldt correction, which Field recommends when the Mauchly's W is above 0.75.

that the interaction was driven by a decrease in the size of the RT switch effect from the first to second test block, $F(1, 111) = 7.28$, $MSE = 6324.35$, $p = 0.008$, $\eta_p^2 = 0.06$. The main effect of position was significant for the first [$F(1, 115) = 304.03$, $MSE = 8114.66$, $p < 0.001$, $\eta_p^2 = 0.73$] and second [$F(1, 113) = 196.77$, $MSE = 7659.53$, $p < 0.001$, $\eta_p^2 = 0.64$] test blocks, but the size of this RT switch effect (P1 - P2) was smaller in the second test (113.99) than that in the first test (144.61). By contrast, A 3 (sequence cue) by 2 (test block) by 3 (position: P2, P3, P4) by 2 (item type) mixed ANOVA showed that the pattern of correct recognition RTs within repetition trials of the run was no different for the first and second test blocks, $F(1.95, 228.75) = 0.84$, $MSE = 3464.98$, $p = 0.432$, $\eta_p^2 = 0.01$. Thus the interaction of test block and position reflected a practice effect—a decrease in the RT switch effect from the first to second test block. It is worth noting that as the first and second test blocks used different list groups, the effects and interactions of test block for the recognition task should be treated with some caution. However, as considerable effort was made to balance linguistic characteristics both within and across list groups, these effects likely represent changes in performance due to experience with the task rather than due to stimulus differences.

Neither the main effect of sequence cue ($F = 0.27$, $p = 0.762$) nor the interaction of sequence cue and position ($F = 0.76$, $p = 0.546$) was significant. Furthermore, a planned Helmert contrast showed no significant difference in correct recognition RTs between the control and experimental (P1, P4) cue conditions ($p = 0.530$). There was no evidence that the presentation of a sequence cue reminder improved overall RT performance or switch trial performance in particular.

Magnitude task.

Table 7.5 shows the mean RTs for correct magnitude judgements as a function of test block, position in run, and sequence cue condition. RTs for magnitude judgements were fast, particularly for repetition trials. Mean RTs for switch trials were longer than all repetition trials, demonstrating a RT switch effect. Performance on the repetition trials (Position 2 – 4) was more variable for the magnitude task than for the recognition task, with some conditions showing a slight slowing at Position 4. Finally, correct magnitude RTs were faster in the second test block compared to the first test block, indicating that experience with the magnitude task allowed participants to make correct magnitude judgements more quickly.

Table 7.5

Means and Standard Deviations for Correct Magnitude Judgement RTs in Experiment 3 as a Function of Position in Run, Test Block, and Sequence Cue Condition.

	P1		P2		P3		P4	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Test 1								
Control	711	191	660	153	653	148	673	125
P1 Cue	693	193	651	155	642	150	641	127
P4 Cue	735	188	681	151	675	146	699	123
Test 2								
Control	651	156	601	140	605	133	621	135
P1 Cue	634	158	583	141	588	135	596	137
P4 Cue	673	154	620	138	642	131	651	134

Note. P1 = position 1, P2 = position 2, P3 = position 3, P4 = position 4. Control = no sequence cue. P1 cue = sequence cue at position 1 of a run. P4 cue = sequence cue at position 4 of a run.

A 3 (sequence cue) by 2 (test block) by 4 (position) by 2 (item type) mixed ANOVA was conducted on correct magnitude RTs. As the key allocation for magnitude judgements was not balanced, effects and interactions of item type are not reported. The Greenhouse-Geisser correction was used for statistical tests that included position as the assumption of sphericity was violated. The significant main effect of position [$F(2.19, 243.11) = 34.66, MSE = 9828.37, p < 0.001, \eta_p^2 = 0.24$] is illustrated in Figure 7.4. Pairwise comparisons using Bonferroni corrections showed that RTs for correct magnitude judgements at Position 1 ($M = 682.91, SD = 83.67$) were slower than for that Position 2 ($M = 632.62, SD = 74.70$), Position 3 ($M = 634.37, SD = 65.82$) and Position 4 ($M = 646.86, SD = 65.53$). However, the mean RT for correct magnitude judgements at Position 4 was also significantly slower than Position 2 and Position 3 (which did not differ significantly from each other). Although the switch effect was isolated to Position

1, this result shows that participants' responses slowed at the end of the run, suggestive of within run slowing.

There was a significant main effect of half, $F(1, 111) = 69.84$, $MSE = 19044.26$, $p < 0.001$, $\eta_p^2 = 0.39$. The mean RT for the first half ($M = 676.20$, $SD = 71.91$) was significantly slower than that for the second half ($M = 622.17$, $SD = 71.56$). Participants made magnitude judgements more quickly after they had gained experience with the task. However, unlike the recognition RT data, the interaction of position and test block was not significant ($F = 1.87$, $p = 0.140$). That is, the size of the RT switch effect and the pattern of performance within a run of a task did not change from the first to the second test. Unlike the recognition task, there was a significant main effect of sequence cue on correct magnitude task RTs, $F(2, 111) = 4.66$, $MSE = 63257.064$, $p = 0.011$, $\eta_p^2 = 0.08$. Post-hoc tests using a Bonferroni correction showed that the mean RT for the P4 sequence cue condition ($M = 672.16$, $SD = 107.30$) was significantly slower than that for the P1 sequence cue condition ($M = 628.33$, $SD = 110.37$). However, neither group differed significantly from the control group ($M = 647.07$, $SD = 108.91$), where no visual sequence cue was presented. Neither sequence cue provided an advantage over the control condition. Critically, there was no interaction of external cue and position ($F = 0.58$, $p = 0.693$), indicating that the use of a sequence cue reminder did not change the size of the RT switch effect or the pattern of recovery from a task switch.

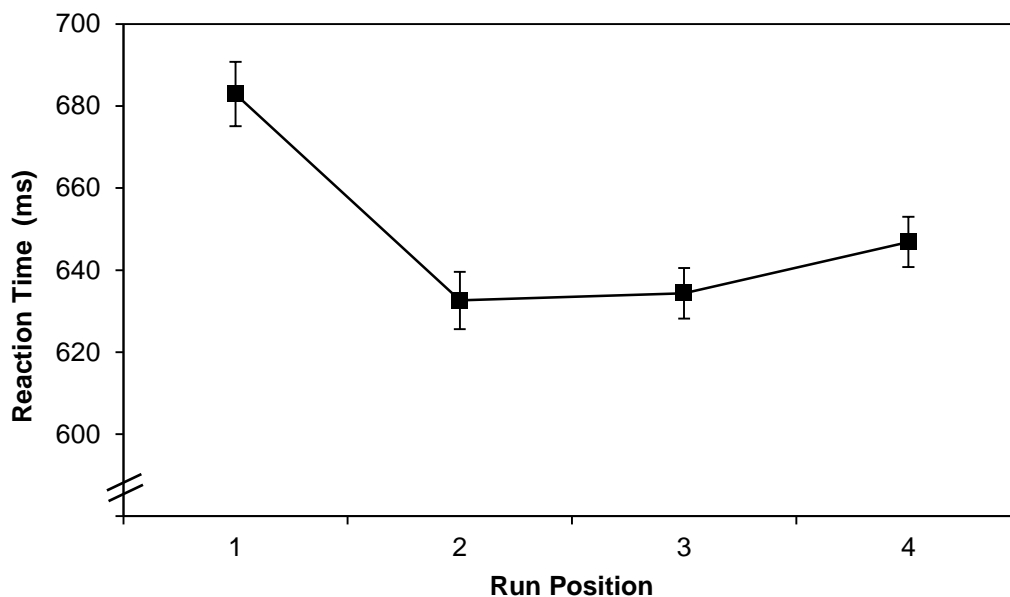


Figure 7.4. Mean RTs for correct magnitude judgements in Experiment 3 as a function of run position. Error bars represent S.E.

Discussion

In Experiment 3, participants switched between a recognition memory task and a magnitude task in 4-trial alternating runs. In addition to examining RT and accuracy performance within a run of a task, the current experiment examined the effectiveness of providing a sequence cue to aid participants in monitoring the task sequence and anticipating task switches. Significant RT and accuracy switch effects were observed for the recognition and magnitude tasks. Participants' performance on switch trials (Position 1) was significantly slower and less accurate than that on subsequent trials of a run. Interestingly, the accuracy switch effect for the recognition task was very similar in size to those reported in Experiments 2 and 3. Even though the current experiment used a longer run length and did not use an encoding manipulation, the cost to recognition accuracy was largely unchanged. Finally, it is particularly important to note that the significant RT and accuracy switch effects for the magnitude task confirm that the novel task switching paradigm used in this thesis can produce task switching effects that are not isolated to the (arguably more complex) recognition task.

Effects of Providing a Sequence Cue

Providing a visual cue to aid task sequence monitoring had no impact on task switching performance relative to a condition where no visual cue was used. For both the recognition and magnitude tasks, supplying a sequence cue did not change RT and accuracy switch effects. Furthermore, the patterns within repetition trials of a run showed no difference across sequence cue conditions. This failure to produce decreases in RT and accuracy switch effects was particularly interesting for the P4 sequence cue condition. In this condition, participants were presented with a visual cue that reminded them of an upcoming task switch and were provided with plenty of time to use this information to prepare. However, this was not the case. The main effect of run position did not differ between sequence cue conditions, indicating that participants in the P4 cue condition were no more prepared for an upcoming task switch than those in the other two cue conditions.

If the sequence cue was perceived, which seems to be a safe assumption, why did the cue have no effect on performance? One explanation could be that, while participants perceived the sequence cue, they may not have remembered the meaning of it. For example, participants in the P4 cue condition may have not remembered that the novel focal stimulus indicated that they were completing the last trial in a run of a task.

However, the instructions presented prior to each test phase placed great emphasis on the meaning of the novel sequence cue and its use as a reminder of their position within a run. Alternatively, participants may have evaluated the visual sequence cue and decided not to use it. Given that the novel focal stimulus (\$\$\$) occurred on only one of four trials, participants still needed to actively monitor the task sequence for the majority of a run. Although presenting a visual sequence cue at Position 4 may have aided participants on the switch trial, it was not particularly helpful when a participant was attempting to remember if they were currently completing the second or third trial in a run. As such, participants may have decided that the additional load associated with monitoring focal stimuli and holding the meaning of the sequence cue in working memory was not worth the bother.

Performance Changes Within a Run

Experiment 3 tested the hypothesis that a constant run length and predictable task switches would produce switch effects isolated to Position 1 (the switch trial). The RT and accuracy data for the magnitude task supported this hypothesis. For the magnitude task, the accuracy switch effect was isolated to the first trial. Magnitude accuracy increased from the first to second trial in the run, but showed no further improvement for later positions. Likewise, the magnitude RTs were very consistent with the hypothesised immediate recovery from a task switch. Correct magnitude judgement RTs were faster for Positions 2 and 3 (which were equal) than that for Position 1. With regard to the recognition task, corrected accuracy was lower on the switch trial than that on Positions 3 and 4 (which were equivalent). However, corrected accuracy of Position 2 fell midway between the Position 1 and 3 and was not significantly different from either position. While these data suggest that *some* of the cost of switching tasks may persist beyond the switch trial, they would have been stronger if Position 2 had been significantly lower than Position 3. Thus it is difficult to definitively conclude if the accuracy data illustrated immediate or gradual recovery. The recognition RT data were more clear-cut. Correct recognition RTs showed that the mean RT for Position 3 was significantly faster than that for Position 2, which was in turn faster than that for Position 1. The switch effect persisted to the second trial in the run.

Why did the recognition task data (but not the magnitude task data) produce evidence of gradual recovery? Monsell et al. (2003) proposed that when run length was unpredictable, participants in their experiment strategically modulated the activation of

the online task set following the first trial of the run, in anticipation of an upcoming (unknown) task switch. This strategic modulation produced a gradual recovery over the first few trials in a run until the benefits of task set priming overcame this strategic modulation. By contrast, when run length was predictable, no such modulation occurred and switch effects were isolated to the first trial in a run. The relative difficulty and high cognitive demands of the recognition memory task (compared to that of the magnitude task) perhaps made it challenging for participants to monitor their position within a run. Participants may have occasionally lost track of when the next task switch would occur and in such cases adopted a strategy similar to that supported by Monsell et al.'s for unpredictable run length data. In essence, despite the run length being consistent and predictable, when participants lost track of their position within the run they behaved as if the run length was unpredictable. That is, they modulated the activation of the online task set in case the next trial was a task switch.

By contrast, the magnitude task may have shown an immediate recovery pattern because it was relatively easy to perform. The mean accuracy of the magnitude task was higher than that for the recognition task and the size of the RT switch effect was considerably smaller. Combined with the limited long term memory retrieval demand (see Mayr & Kliegl, 2003), the relative ease of performing the magnitude task likely meant that participants were more successful at tracking their position within a run of the magnitude task. Participants performed this task as if they were aware of the consistent and predictable run length and largely capable of keeping track of their position within a run. Thus performance on the magnitude task may have produced a pattern similar to that reported by Monsell et al. (2003) for predictable run lengths.

Finally, it is worth noting that the size of the difference between Positions 2 and 3 RTs for the recognition task was very small (10.65 ms) compared to the difference between Positions 1 and 2 (126.75 ms). Occasions where participants lost track of their position within a run of the recognition task—and thus modulated the online task set activation—may have been relatively few in number. It may be that the component of the switch effect that persisted past Position 2 was an artefact of the considerable statistical power in this experiment.

With regard to evidence of within run slowing, that the magnitude task showed significant slowing on the last trial of a run, whereas the recognition task did not. Altmann and Gray (2002) explained within-run slowing as the result of decay in activation of the online task set. That is, performance on the fourth trial of the

magnitude task was slowed because the task set had decayed and thus needed additional activation before it could direct behaviour. However, confirming this account would require a similar slowing for the recognition task, and no such slowing was observed. The run length used in the current experiment was relatively short compared to those used in Altmann and Gray (2002). Thus, the run length may not have been sufficiently long to observe within-run slowing for the recognition task. Alternatively, the slowing of Position 4 in the magnitude task could be explained as a change task switching strategy. A unique characteristic of the experiments in this thesis is that, contrary to most task switching experiments, one task may appear more important than the other. In the study phases prior to each test, participants were instructed to study the words for a later recognition memory test. At test, participants were informed that they would be performing a recognition memory task with an additional magnitude task. In response to this information, participants may have placed more attention and importance on the recognition task. Thus the slowing of performance on Position 4 of the magnitude task may have reflected participants' anticipation of the upcoming switch to the more important recognition task. By contrast, the last trial in a run of recognition tasks is afforded the same importance (and similar performance) as other repetition trials in a run and thus shows no slowing.

Summary

In Experiment 3, participants switched between the recognition memory task and the magnitude task in predictable 4-trial alternating runs. Both tasks showed significant accuracy and RT switch effects. Providing a sequence cue on either Position 1 or Position 4 of a run did not decrease the size of the RT or accuracy switch effects or aid participants in monitoring task sequence. While magnitude task RTs and accuracy showed an immediate recovery pattern, the recognition task RTs showed a more gradual recovery pattern. This difference in recovery patterns likely reflected the different strategies of switching and maintaining task sets used by participants for each task.

An interesting result in Experiment 3 was the decrease in the size of the recognition task RT switch effect and the magnitude task accuracy switch effect from the first to second tests test blocks. The decrease in RT switch effect, called a practice effect, has had relatively limited coverage in the literature (Jersild, 1927; Koch, 2005; Merian, 1996; Minear & Shah, 2008; Rogers & Monsell, 1995) but could provide further information about the processes of switching between the magnitude and recognition tasks. As such, Chapter 8 aims to explore this effect in more detail.

Chapter Eight

Experiments 4A – 5B: Preparation and Practice

As previously noted in this thesis, participants may adjust or refine their task switching processes/strategies in response to methodological characteristics of an experiment (e.g., Altmann, 2002; Altmann & Gray, 2002; Mayr, 2007; Monsell et al., 2003; Monsell & Mizon, 2006). A popular example of this type of change is the preparation effect. A preparation effect is defined as the decrease in the size of the switch effect (commonly RTs, but sometimes accuracy) with longer cue-to-stimulus intervals (CSIs, Altmann, 2004a, 2004b; Mayr & Kliegl, 2003; Meiran, 1996; Monsell & Mizon, 2006), or in the alternating runs paradigm, with longer response-to-stimulus intervals (RSIs, Arrington & Logan, 2004, 2005; de Jong, 2000; Rogers & Monsell, 1995; Monsell et al., 2003). As was illustrated in Experiment 3, the speed or use of specific task switching processes can also change as participants gain experience with the tasks and with switching between them. One consequence of this change, called a practice effect, is the decrease in the size of switch costs or mixing costs across the course of an experiment (Jersild, 1927; Koch, 2005; Kramer et al., 1999; Kray & Lindenberger, 2000; Meiran, 1996; Minear & Shah, 2008; Monsell et al., 2003; Rogers & Monsell, 1995).

Over the span of the last two decades, the preparation effect has become a well-established phenomenon in task switching research (see Monsell, 2003; Vandierendonck et al., 2010). The effect has been used in the development and validation of task switching theories (Altmann & Gray, 2008; Rogers & Monsell, 1995) and as a tool through which to examine the effects of other manipulations (Mayr & Keele, 2000; Miyake et al., 2004; Philipp & Koch, 2006). Proponents of two-stage reconfiguration models of task switching (Meiran, 1996, 2000a; Rogers & Monsell, 1995; Rubinstein et al. 2001) have argued that longer RSIs and CSIs allow participants additional time to reconfigure elements of the task set that are not dependent on stimulus presentation. By contrast, the failure-to-engage (FTE) hypothesis (deJong, 2000, 2001; Lindsen & de Jong, 2010) posits that longer RSIs/CSIs increase the probability that task sets will be completely reconfigured prior to stimulus onset. The Cognitive Control Model (Altmann & Gray, 2008) explains the preparation effect as the

product of additional cycles (iterations) of activation applied to a cued task set with longer CSIs/RSIs.

Conflicting ideas about the methods that will or will not produce a preparation effect have been reported, particularly with the alternating runs paradigm. For example, there are inconsistencies regarding the appropriate way to manipulate RSI. When Rogers and Monsell (1995) presented multiple RSIs within a single block, they did not observe preparation effects. However, other researchers have presented multiple RSIs within an alternating runs block and reported significant preparation effects (e.g., de Jong, 2000). Despite this inconsistency, one generally accepted requirement for observing preparation effects is that the RSI should be manipulated within participants. That is, the effect seems to rely on exposure to more than one RSI (Altmann, 2004b). Another case of inconsistency regards the necessity of task cues in the alternating runs paradigm. While Koch (2003) reported that preparation effects were not observed when a task cue was not used, Kray and Lindenberger (2000) reported significant preparation effects in the absence of task cues. Thus although investigations of preparation effects are incredibly common, there remains an uncertainty about when they will be observed.

Practice effects were first reported by Jersild (1927), who showed that completion times for mixed blocks (e.g., ABABAB) decreased over the course of an experiment to a greater extent than that for pure blocks (AAAAAA). Jersild attributed this differential improvement to participants developing efficient task and instructional sets that allowed them to switch tasks more quickly. A more detailed and theoretically-driven account of practice effects was proposed by Meiran (1996), who identified the speeding of task set reconfiguration as the source of practice effects. Practice with tasks within an experiment strengthened the association between task cues and task sets, allowing those task sets to be retrieved more quickly. As a consequence, task set reconfiguration was completed more quickly. Interestingly, Meiran reported that practice interacted with a manipulation of CSI. That is, significant practice effects were observed for trials using a short (132 ms) CSI, but no practice effects were observed for trials using a long (1632 ms) CSI. This interaction was interpreted as reflecting a difference in the benefit of fast reconfiguration at short and long CSIs. At short CSIs, fast reconfiguration increased the likelihood that this reconfiguration would be completed prior to stimulus onset, thus decreasing the RT switch cost. At long CSIs, reconfiguration is more likely to be completed prior to stimulus onset regardless of the

speed of reconfiguration, and as such practice provides little additional benefit (Meiran, 1996).⁸

Relative to preparation effects, practice effects have a longer, but leaner legacy. With the exception of Meiran's (1996) examination of practice effects, the task switching literature has largely referred to practice effects in passing (e.g., Rogers & Monsell, 1995) or acknowledged them by way of methodology—excluding many trials at the beginning of each experiment. There are, however, some notable exceptions. Task switching costs for older and younger adults have been shown to decrease in size across multiple sessions (Kray & Lindenberger, 2000). Furthermore, the size of older adults' RT switch cost can decrease over the course of an experiment until it is equivalent to younger adults' (Kramer et al., 1999). The benefit of practice for task switching performance does not transfer to different tasks (Minear & Shah, 2008). By contrast, the benefit of practice to task switching performance can remain intact when the predictability of task sequences changes. Koch (2005) showed that a practice effect that developed across multiple blocks of predictable task switching was maintained when a block using random task switches (with the same tasks) was introduced. Although overall RTs increased for this random sequence transfer block, the size of the switch cost was equivalent to the last predictable sequence block. In combination with the findings of Jersild (1927) and Meiran (1996), these studies show that practice effects make an important contribution to task switching performance and support the claim that practice strengthens the associations between task cues and task sets.

Experiment 3 of this thesis reported a significant decrease in the size of the RT switch effect (Position 1 – Position 2) from the first to second test for the recognition memory task. With practice, participants became more efficient at switching tasks and thus made their recognition memory decisions more quickly after a task switch. The current chapter reports results from four experiments that aimed to further examine these practice effects whilst also exploring the role of RSI length on task switching performance. In each experiment, participants completed two study-test blocks. During the study phases, participants studied and made semantic judgements to words. During the test phases, participants switched between the recognition and magnitude tasks in predictable, two-trial runs. Two RSIs were used in each experiment—one for each test

⁸ Koch (2005) reported significant practice effects with a CSI of 900 ms—a CSI clearly beyond the optimal preparation duration posited by Rogers and Monsell (1995). Thus, it may be incorrect to say that practice effects can never be observed in conjunction with preparation effects.

phase. Assessments of preparation and practice were made by comparing performance across these two test blocks.

It is important to acknowledge that the method of examining task switching and preparation used in this chapter departed from usual methods. Typically, experiments that examine preparation effects have used very simple tasks and have provided some type of external task cue. When an alternating runs paradigm has been used participants have often been explicitly instructed that they should use the time between trials to prepare for the upcoming task. The choice to deviate from this standard paradigm was made in response to the specific research questions this thesis sought to address. Critically, the current experiments sought to examine evidence for preparation effects using a more conservative and naturalistic task switching method. Participants switched between the higher order recognition task and the relatively simple magnitude task that have been used throughout this thesis. Switching between these two different tasks with different stimulus sets is arguably more reflective of everyday life than switching between two simple tasks that use the same stimulus set. To encourage participants to monitor the task sequence and keep track of upcoming changes in task (as they would likely do in everyday life), explicit task cues were not provided. Finally, to establish if participants naturally prepare for upcoming task switches (as opposed to if they can be instructed to prepare), participants were made aware of the task progression and the length of RSIs but were not explicitly told to use the RSI to prepare. This desire to ascertain if and how participants naturally prepare is not unheard of, and is a valid point of scientific inquiry. Indeed, research in the related field of interruptions conducted by Trafton et al. (2003) did not instruct participants to prepare during a ‘warning’ interval between the primary task and the start of the interruption task for the specific purpose of “measuring people’s natural propensity to prepare...” (p. 589).

Although these changes in method are fairly substantial, there is evidence to suggest that preparation effects will still be observed. For example, in Kray and Lindenberger (2000), younger and older adults switched between two tasks (e.g., shape or colour discrimination for shape stimuli) in two-trial alternating runs over multiple blocks. Each block used a long (1200 ms) or a short (200 ms) RSI. External task cues were not provided and participants were not explicitly instructed to use the RSI to prepare for upcoming trials. Kray and Lindenberger reported a typical preparation effect; significantly smaller RT switch costs for long RSI blocks than that for short RSI blocks. Kray (personal communication, February 2011) stated that by not using task

cues, participants were forced to keep track of the task sequence and consequentially prepare for task switches.⁹ As has been noted earlier in this chapter, Koch (2003) reported that without task cues preparation effects were not observed, and he argued that internally generated cues (i.e., cues generated as a product of monitoring a task sequence) provided weak support for reconfiguration. Clearly, there is a lack of agreement about the exact parameters for which preparation effects will (or will not) be observed. The current research should aid in clarifying this debate.

There were three key hypotheses for the experiments presented in this chapter. First, it was hypothesised that, consistent with earlier experiments in this thesis, switching between the recognition and magnitude task would produce significant RT and accuracy switch effects. That is, recognition and magnitude task responses should be slower and less accurate for switch trials compared to those for repetition trials. Second, it was hypothesised that significant practice effects would be observed. The size of RT switch effects should decrease from the first to the second test block. Given that previous research has shown that the benefits of practice are specific to task as opposed to other experimental factors (Koch, 2005; Meiran, 1996), it seems reasonable to assume that changes in the RSI across test blocks should not affect the observation of practice effects. Finally, it was hypothesised that a preparation effect would be observed for recognition and magnitude tasks. The size of RT switch effects (and perhaps accuracy switch effects) should be significantly smaller for long RSI test blocks than for short RSI test blocks. If, as suggested by Altmann (2004b), advanced preparation is only enacted after exposure to multiple RSIs, preparation effects may only be observed when examining performance on Test 2 (i.e., between participants). For the second test, the RT switch cost should be smaller for participants completing a long RSI test than that for participants completing a short RSI test.

General Method

Participants

Undergraduate psychology students at Victoria University of Wellington participated in the studies for credit toward a research requirement in an introductory psychology class. All participants in the final data set: (1) had normal or corrected-to-normal vision, (2) spoke English from early childhood, and (3) met minimum

⁹ Kray also stated that it was this property that made it unnecessary to explicitly instruct participants to use the RSI to prepare—as task cues were not provided participants had to keep track of the task sequence internally.

performance requirements. These performance requirements were: (a) above chance accuracy for repetition trials on the magnitude and recognition tasks, (b) mean RTs for correct trials within 3 *SD* of the group mean, and (c) fewer than 20% timeouts for either task. Participants who failed to meet the above criteria were replaced to achieve proper counter-balancing.

Materials

Concrete nouns ($N = 150$) from the Toronto Word Pool (Friendly et al. 1982) were used to construct 8 lists of 16 words (see Appendix A). Lists were closely matched for syllables ($M = 2.00$), word length ($M = 6.27$), frequency (Kucera & Francis, 1967, $M = 11.52$), concreteness ($M = 6.03$) and imagery ($M = 5.66$). Buffer and filler words ($N = 22$) had similar linguistic properties. Words stimuli were presented centrally on a white screen in black bold 24-pt Courier New font. Dot stimuli were black circles (diameter = 8 mm). The experiments were conducted using E-Prime software (Schneider et al., 2002).

Design

The critical manipulation in the current experiments was a within-participant manipulation of RSI. Each experiment consisted of two study-test blocks, and each study-test block was associated with a single RSI at test. The recognition memory task was manipulated in a 2 (RSI: short, long) \times 2 (test block: test 1, test 2) \times 2 (transition: repetition, switch) \times 2 (item type: old, new) repeated measures design. The magnitude task was manipulated in a 2 (RSI: short, long) \times 2 (test block: test 1, test 2) \times 2 (transition: repetition, switch) \times 2 (item type: less, more) repeated measures design. The order of the two RSIs (RSI order: long-short, short-long) was balanced across participants and was included as a variable in the statistical analyses.

Procedure

The experiment was conducted in a single session that comprised two study-test blocks. Participants (one to four in a session) sat at individual computer stations approximately 0.5 meters from the monitor. In each study phase, participants studied a list of words for a later memory test. In each test phase, participants switched between the recognition memory task and the magnitude task in 2-trial alternating runs. The RSIs used in each experiment are presented in Table 8.1.

Study Phases.

Each study phase consisted of 42 trials: 5 primacy buffer words, 32 words from 2 critical lists (one for each position at test), and 5 recency buffer words. Participants were instructed to try to decide whether each word represented a living or non-living thing and to study the word for a later test. The living/non-living encoding task was used to ensure that participants attended to the stimuli during study (evidence from Experiment 2 suggested that it also improved memory performance at test). A study trial began with the presentation of a brief (100 ms) focal stimulus (+++) in the centre of the monitor. Immediately after the focal stimulus a word was presented for 3000 ms in the same location. Once a word was presented, participants entered living/non-living decisions on the keyboard using the *f* and *j* keys respectively. After the 3000-ms presentation was complete, an intertrial interval (ITI) of a blank screen occurred for 500 ms.

Test Phases.

Each of the two test phases comprised a practice test and an actual test. Participants performed a recognition memory task (R) and a magnitude task (M) in a predictable, 2-trial alternating runs sequence (RR MM RR...). The practice test contained 10 trials, starting with the magnitude task. Words for recognition trials in the practice test were either buffer words from the study phase or new words otherwise not presented in the experiment. The actual test consisted of 130 trials, again starting with

Table 8.1

Characteristics of Experiments in Chapter 8.

Experiment	RSI		N	Notes
	Short (ms)	Long (ms)		
4A	200	600	32	
4B	600	1300	32	
5A	200	2600	32	
5B	200	2600	64	practice tests used both RSIs

Note. RSI = Response-to-stimulus interval. N = number of participants in the experiment.

the magnitude task. ‘Old’ words were from the two lists presented in the associated study phase and ‘new’ words were from two lists not previously presented. One ‘old’ list and one ‘new’ list were used for each position in the run of recognition trials. The allocation of lists to test block, position and item status was counterbalanced across participants. For each position in the magnitude task run, half the trials required ‘more’ decisions, and half ‘less’ decisions. All iterations of ‘more’ and ‘less’ decisions across the two run positions were presented pseudo-randomly throughout the actual test such that correct magnitude judgements were not predictable within or across runs and were not useful to predict the status of an upcoming recognition trial.

Each test phase was associated with a long RSI or a short RSI. The order of RSI presentation was balanced across participants so that half experienced the long RSI first and half experienced the short RSI first. The RSI itself was made up of a variable length ITI of a blank screen plus the 100-ms focal stimulus prior to target onset (i.e., the long 600-ms RSI in Experiment 4A consisted of a 500-ms blank ITI plus the focal point). RSI was manipulated in a naturalistic manner—participants were made aware of the time between trials, but not explicitly instructed to use this time to prepare.

A recognition memory trial began with the brief (100 ms) presentation of a focal stimulus (+++) in the centre of the monitor. Immediately after the focal stimulus the target word was presented in the same location. A response delay of 400 ms passed before 8 asterisks (*****) appeared above and below the target word, indicating that old/new judgements could be entered using the *f* and *j* keys on the keyboard. The response delay prevented participants from producing responses pre-emptively or based on anticipatory processes. After 1000 ms, exclamation marks (!!!!!!!) replaced the asterisks for 600 ms, warning participants they needed to make an immediate response as they were about to run out of time. The response deadline was used to constrain responding and reduce the variance of RTs. The trial terminated when a response was made or the warning period ended. The next trial began after the ITI. The allocation of keys to recognition responses was balanced across participants.

A magnitude task began with the brief (100 ms) presentation of a focal stimulus (+++) in the centre of the monitor. Immediately after the focal stimulus the dot targets (6 or 8 circles) were presented randomly within a central 100 mm (length) x 70 mm (height) section of the display with the proviso that dots did not overlap. Participants were instructed to judge whether there were more than or less than seven dots and to respond using the *f* (more) and *j* (less) keys on the keyboard. When a response was

made (or 2000 ms had elapsed), the trial was terminated. The next trial began after the ITI. The allocation of keys to magnitude responses was not balanced as there was no interest in comparing performance for ‘more’ and ‘less’ decisions.

Each test phase provided instructions about the two tasks, the structure (i.e., task sequence) of the test phase, and the RSI. After completing the practice test, participants were reminded about the response allocations for the magnitude and recognition tasks prior to starting the actual test. Upon completion of the first study-test block, the instructions were repeated for the second study-test block (with a change to the RSI information). When both study-test blocks were completed, participants were debriefed and thanked for their participation.

Data from the practice tests and the first two trials (both magnitude trials) of the actual tests were excluded from analyses. Statistical analyses of RT performance used participants’ median RTs for correct trials in each condition. An alpha level of 5% was used in all statistical analyses. Although accuracy data will be reported, the key interest in this chapter is the RT data, where practice and preparation effects are more consistently observed. For each task, the data were analysed with regard to three key interests: (1) the effect of task transition, (2) the effect of test block, and (3) the effect of RSI length. Changes in the size of the accuracy switch effect or the RT switch cost as a function of test block or RSI length were of particular interest as they should represent practice effects and preparation effects respectively. It is important to note that by analysing each performance measure with respect to these effects of interest there is a certain level of redundancy in the analyses. For example, a main effect of test block can also be reported as an interaction of RSI and order of RSI presentation (the between-participant manipulation). As such, interactions that are redundant (i.e., covered elsewhere in the results for that task) will not be reported. Toward this goal of conciseness, it is also important to note that for the RT analyses, an interaction of transition and any other variable (say, test block) is *the same* as a main effect of this variable on RT switch cost. Thus, the RT analyses do not report interactions that include task transition as they are covered in detail in the RT switch cost analyses.

Experiment 4A

Experiment 4A used a short RSI of 200 ms and a long RSI of 600 ms. These RSIs were selected based on previous literature that has suggested that preparation effects reach asymptote at around 600 ms (Rogers & Monsell, 1995; Meiran, 1996).

Results and Discussion

Thirty-two participants met the performance requirements and were included in the final data analysis. The use of only correct responses in the RT and switch cost analyses led to the exclusion of 14% of recognition memory trials and 13% of magnitude judgement trials. The percentage of timeouts was very low for both tasks (< 1%).

Accuracy

Recognition task.

Table 8.2 shows the mean corrected accuracy (hit rate – false alarm rate) for recognition decisions as a function of RSI, transition, and test block. Mean corrected accuracy across all conditions was moderately high ($M = 0.73$, $SD = 0.13$) and there was no difference in overall accuracy between the two RSI order groups ($F = 0.39$, $p = 0.536$).

Effects of transition and test block.

As can be seen in Table 8.2, the mean corrected accuracies for repetition conditions were higher than those for switch conditions at both RSIs and test blocks, illustrating an accuracy switch effect. Corrected accuracy showed an increase from the first to second test, suggesting that experience with the tasks improved overall performance. However, the difference between repetition and switch trial corrected accuracy did not consistently decrease from the first to second test. That is, there was

Table 8.2

Means and Standard Deviations for Corrected Recognition Accuracy in Experiment 4A as a Function of RSI, Test Block, and Transition.

Transition	Long RSI (600 ms)				Short RSI (200 ms)			
	Test 1		Test 2		Test 1		Test 2	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Repetition	0.73	0.16	0.81	0.12	0.75	0.17	0.81	0.16
Switch	0.62	0.28	0.73	0.16	0.71	0.13	0.70	0.18

Note. RSI = response-to-stimulus interval. Within-participant comparisons should be made across RSIs from Test 1 to Test 2 (i.e., Test 1 Long RSI, Test 2 Short RSI).

little evidence of a practice effect on recognition accuracy. The effects of transition and test block on corrected recognition accuracy were analysed using a 2 (test block) x 2 (transition) x 2 (RSI order) mixed ANOVA. The main effect of transition was significant, $F(1, 30) = 17.27$, $MSE = 0.01$, $p < 0.001$, $\eta_p^2 = 0.37$. Typical of an accuracy switch effect, participants' ability to discriminate between old and new words was significantly lower for switch trials ($M = 0.69$, $SD = 0.18$) than that for repetition trials ($M = 0.78$, $SD = 0.12$). The main effect of test block was significant, $F(1, 30) = 6.21$, $MSE = 0.02$, $p = 0.018$, $\eta_p^2 = 0.17$. Participants' corrected accuracy improved with experience from Test 1 ($M = 0.70$, $SD = 0.18$) to Test 2 ($M = 0.76$, $SD = 0.13$). Test block did not interact with transition ($F = 0.25$, $p = 0.623$). The improvement to recognition accuracy from Test 1 to Test 2 was equivalent for repetition and switch trials—there was no practice effect.

Effects of RSI.

The corrected accuracy reported in Table 8.2 shows that overall accuracy and the difference between repetition and switch accuracy showed no consistent pattern across the two RSI conditions. A 2 (RSI) x 2 (transition) x 2 (RSI order) mixed ANOVA confirmed that both the main effect of RSI ($F = 0.55$, $p = 0.464$) and the interaction of RSI and transition ($F = 0.58$, $p = 0.451$) were not significant. Providing additional time between trials did not improve memory accuracy overall or for switch trials in particular. That is, the corrected recognition data showed no evidence of a preparation effect. No other effects or interactions were significant ($F_s < 1.28$).

Magnitude task.

Table 8.3 shows the mean accuracy of magnitude judgements as a function of RSI, transition, decision, and test block. As expected, the mean accuracy across conditions was high ($M = 0.87$, $SD = 0.07$) and there was no difference in overall accuracy between the two RSI order groups ($F = 0.27$, $p = 0.608$).

Effects of transition and test block.

Table 8.3 shows that the mean accuracies for the repetition conditions were higher than those for the switch conditions, illustrating an accuracy switch effect. The accuracy of magnitude judgements generally improved from the first to second test. However, the difference between repetition and switch trials did not show a consistent pattern of improvement across tests—there was no evidence of a practice effect. The effects of transition and test block were analysed using a 2 (test block) x 2 (transition)

Table 8.3

Means and Standard Deviations for Magnitude Judgement Accuracy in Experiment 4A as a Function of RSI, Test Block, Transition, and Decision.

Transition and Decision		Long RSI (600 ms)				Short RSI (200 ms)			
		Test 1		Test 2		Test 1		Test 2	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Repetition	More	0.88	0.16	0.94	0.07	0.90	0.12	0.90	0.10
	Less	0.85	0.13	0.91	0.12	0.83	0.14	0.94	0.06
Switch	More	0.77	0.25	0.93	0.05	0.85	0.10	0.89	0.09
	Less	0.84	0.18	0.88	0.13	0.81	0.13	0.86	0.13

Note. RSI = response-to-stimulus interval. Within-participant comparisons should be made across RSIs from Test 1 to Test 2 (i.e., Test 1 Long RSI, Test 2 Short RSI).

x2 (item type) x 2 (RSI order) mixed ANOVA. For this and all other experiments in this chapter ‘less’ and ‘more’ magnitude judgements were analysed together to detect the differences of critical interest. However as the allocation of keys to responses was not balanced for the magnitude task, the effects and interactions of item type cannot be interpreted. As such, they are not reported. The main effect of transition was significant, $F(1, 30) = 9.34$, $MSE = 0.01$, $p = 0.005$, $\eta_p^2 = 0.24$. Participants’ magnitude judgements were significantly less accurate on switch trials ($M = 0.85$, $SD = 0.10$) than those on repetition trials ($M = 0.90$, $SD = 0.07$). The main effect of test block was significant, $F(1, 30) = 11.52$, $MSE = 0.02$, $p = 0.005$, $\eta_p^2 = 0.28$. The accuracy of magnitude judgements improved from Test 1 ($M = 0.84$, $SD = 0.12$) to Test 2 ($M = 0.91$, $SD = 0.06$). The interaction of test block and transition was not significant ($F = 0.63$, $p = 0.43$). Although magnitude task accuracy showed a general benefit of experience, this benefit was equivalent for repetition and switch trials. That is, there was no practice effect for magnitude accuracy.

Effects of RSI.

As can be seen in Table 8.3, the mean accuracy (and the difference between repetition and switch trial accuracy) showed no discernable pattern of differences for long and short RSI conditions, suggesting that the RSI did not contribute to the accuracy of magnitude judgements. A 2 (RSI) x 2 (transition) x 2 (item type) x 2 (RSI order)

mixed ANOVA confirmed that neither the main effect of RSI ($F = 0.01, p = 0.936$), nor the interaction of RSI and transition ($F = 0.02, p = 0.891$) were significant. The length of the RSI had no impact on overall magnitude task accuracy or the size of the accuracy switch effect. That is, there was no evidence of a preparation effect for magnitude task accuracy. No other effects or interactions were significant ($F_s < 0.83$).

Reaction Times

Throughout this chapter, the RT results are separated into several subsections: (1) analyses of correct recognition RTs, (2) analyses of correct magnitude RTs, and (3) analyses of RT switch costs (switch RTs – repetition RTs). For this experiment as well as subsequent experiments in this chapter, the mean correct median RTs for the recognition and magnitude tasks can be found in the supplementary material (Appendix D). The overall RT for correct recognition decisions was similar to that reported in earlier experiments ($M = 390.10, SD = 77.35$). Furthermore, when the 400-ms response signal delay is taken into account, the overall correct recognition RT was relatively similar to the overall correct magnitude RT ($M = 730.72, SD = 82.44$). The order of RSI presentation had no effect on correct recognition RTs ($F = 0.46, p = 0.501$) or correct magnitude RTs ($F = 0.12, p = 0.727$).

Recognition task.

The effects of transition and test block on correct recognition RTs were analysed using a 2 (test block) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. A significant main effect of transition [$F(1, 30) = 54.72, MSE = 12734.66, p < 0.001, \eta_p^2 = 0.65$] confirmed a RT switch effect. Correct recognition decisions were slower for switch trials ($M = 442.58, SD = 108.80$) than those for repetition trials ($M = 338.23, SD = 57.56$). The main effect of item type was also significant, $F(1, 30) = 36.05, MSE = 6016.10, p < 0.001, \eta_p^2 = 0.55$. Participants were slower to correctly identify new words ($M = 419.52, SD = 78.82$) than old words ($M = 361.30, SD = 85.20$). There was a main effect of test block, $F(1, 30) = 19.36, MSE = 8559.94, p < 0.001, \eta_p^2 = 0.39$. Mean RTs for correct recognition trials were slower in Test 1 ($M = 415.85, SD = 96.90$) than in Test 2 ($M = 364.97, SD = 68.69$). Participants became faster at producing correct recognition responses with experience.

The effect of RSI length on correct recognition RTs was analysed using a 2 (RSI) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. Although the mean RT for correct recognition decisions was slightly faster for the long RSI condition

($M = 379.52$, $SD = 88.42$) than that for the short RSI condition ($M = 401.30$, $SD = 79.30$), this difference did not reach statistical significance ($F = 3.55$, $p = 0.069$). No other effects or interactions were significant ($F_s < 1.53$).

Magnitude task.

The effects of transition and test block on correct magnitude RTs were analysed using a 2 (test block) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. Again, the main effects and interactions of item type are not reported as response key allocation was not balanced for the magnitude task. The main effect of task transition was significant, $F(1, 30) = 25.26$, $MSE = 9078.82$, $p < 0.001$, $\eta_p^2 = 0.46$. Participants took significantly longer to make correct magnitude judgements for switch trials ($M = 768.79$, $SD = 101.78$) than that for repetition trials ($M = 708.79$, $SD = 74.19$). Collapsed across both transitions, there was a main effect of test block, $F(1, 30) = 18.71$, $MSE = 14022.81$, $p < 0.001$, $\eta_p^2 = 0.39$. The mean RTs for correct magnitude judgements were slower in Test 1 ($M = 770.73$, $SD = 104.83$) than in Test 2 ($M = 706.71$, $SD = 78.17$), illustrating the benefit of experience with the magnitude task. The main effect of test block was qualified by a significant interaction with RSI order, $F(1, 30) = 6.95$, $MSE = 14022.81$, $p = 0.013$, $\eta_p^2 = 0.19$. When the long RSI was used in the first test, the decrease in mean RT from Test 1 ($M = 756.34$, $SD = 111.29$) to Test 2 ($M = 731.34$, $SD = 89.86$) was not significant ($F = 1.59$, $p = 0.227$). However, when the short RSI was used in the first test, the decrease in mean RTs from Test 1 ($M = 785.11$, $SD = 97.94$) to Test 2 ($M = 682.07$, $SD = 64.39$) was significant, $F(1, 15) = 21.96$, $MSE = 154710.18$, $p < 0.001$, $\eta_p^2 = 0.59$. Faster correct responses in the long RSI test block drove the interaction of test block and RSI order by inflating the benefit of experience when the short RSI was used for the first test.

The effect of RSI length on correct RTs for the magnitude task was analysed using a 2 (RSI) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. The main effect of RSI was significant, $F(1, 30) = 6.95$, $MSE = 14022.81$, $p = 0.013$, $\eta_p^2 = 0.19$. Participants made correct magnitude judgements more quickly in the long RSI condition ($M = 719.21$, $SD = 90.92$) than in the short RSI condition ($M = 758.23$, $SD = 93.99$). No other effects or interactions were significant ($F_s < 0.07$).

RT switch costs.

In the task switching literature, the preparation effect is generally operationalized as a decrease in RT switch cost with longer RSIs/CSIs (Monsell, 2003;

Vandierendonck et al. 2010). The benefit of the RT switch cost measure (over RTs) is that it gives a concrete value showing the extent that performance is affected by a task switch. Mean RT switch costs, presented in Table 8.4, were calculated for each participant by subtracting the median correct repetition RTs from the median correct switch RTs for each RSI, task and item type. The RT switch costs showed considerable variability across conditions and the standard deviations were very large. There was a great deal of variability in participants' ability to switch tasks quickly. The RT switch costs were generally larger for the recognition task than those for the magnitude task, particularly in the first test block. The order of RSI presentation had no effect on the RT switch costs ($F = 0.00$, $p = 0.968$).

The RT switch costs for both tasks and item types ('old', 'new', 'more', 'less') were analysed together to increase the statistical power to detect the differences of critical interest—changes in RT switch cost due to experience switching tasks (practice effects) and the length of the RSI (preparation effects). However, it is important to acknowledge that in this case comparisons are being made across items where some are balanced for key allocation (recognition task) and some are not balanced for key allocation (magnitude task). As such, item type differences in RT switch cost are interpreted with caution.

Table 8.4

Means and Standard Deviations for the RT Switch Costs in Experiment 4A as a Function of RSI, Test Block, Task and Decision.

Task and Decision		Long RSI (600 ms)				Short RSI (200 ms)			
		Test 1		Test 2		Test 1		Test 2	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Recognition	New	110	145	74	72	119	105	116	113
	Old	125	114	92	98	117	127	81	75
Magnitude	More	59	117	63	70	63	93	37	85
	Less	62	89	52	73	73	127	69	81

Note. RSI = response-to-stimulus interval. Within-participant comparisons should be made across RSIs from Test 1 to Test 2 (i.e., Test 1 Long RSI, Test 2 Short RSI).

Effects of test block.

There was a numeric decrease in the size of the recognition task RT switch cost from the first to second test, suggesting that practice may have decreased the cost of switching (see Table 8.4). However, differences in the magnitude task RT switch cost from the first to second test were smaller and less consistent. The effect of test block on RT switch costs was analysed using a 2 (test block) x 4 (item type) x 2 (RSI order) mixed ANOVA. There was a significant main effect of item type [$F(1, 30) = 4.89$, $MSE = 8802.86$, $p = 0.003$, $\eta_p^2 = 0.14$], illustrated in Figure 8.1. Pairwise comparisons using Bonferroni corrections showed that the mean RT switch cost for ‘old’ recognition trials ($M = 103.94$, $SD = 93.06$) was significantly larger than that for ‘more’ dot trials ($M = 55.45$, $SD = 85.56$), but did not differ from ‘new’ recognition trials ($M = 104.78$, $SD = 92.69$) or ‘less’ dot judgement trials ($M = 64.46$, $SD = 65.91$). There were no other significant differences between item types. These results provide some support for a task effect on task switching performance, but as the effect is limited to one recognition decision and one magnitude judgement, this is far from convincing. Although the mean RT switch cost for the first test ($M = 91.08$, $SD = 83.31$) was larger than that for the second test ($M = 73.12$, $SD = 56.66$), the difference was not significant ($F = 2.30$, $p = 0.140$). The interaction of test block and item type was also not significant ($F = 0.44$,

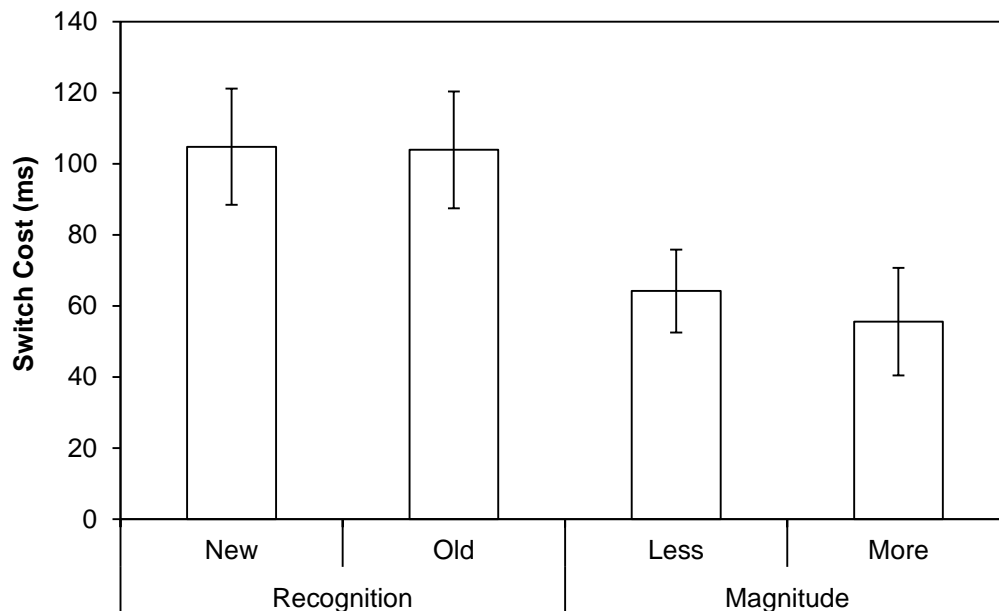


Figure 8.1. Mean RT switch costs for Experiment 4A as a function of task and item type. Error bars represent S.E.

$p = 0.69$). Thus, there was no significant evidence of a practice effect on the RT switch costs.

Effects of RSI.

With regard to evidence for preparation effects, the RT switch costs reported in Table 8.4 showed no systematic differences for RSI length. The effect of RSI on RT switch costs was analysed using a 2 (RSI) x 2 (RSI order) x 4 (item type) mixed ANOVA. The mean switch cost for the long RSI condition ($M = 79.82$, $SD = 74.41$) was numerically lower than that for the short RSI condition ($M = 84.39$, $SD = 67.58$), but the difference was not significant ($F = 0.15$, $p = 0.702$). No other effects or interactions were significant ($F_s < 0.99$). When each test block was analysed separately—with RSI as a between-participant variable—the main effect of RSI was not significant for Test 1 ($F = 0.02$, $p = 0.902$), or after participants had experienced both RSIs (Test 2, $F = 0.08$, $p = 0.785$). Thus the analyses of RSI effects on RT switch costs showed no evidence to support the hypothesised preparation effect.

Cumulative RT distributions.

One proposal regarding task set reconfiguration is that task sets can be reconfigured (or ‘engaged’) prior to stimulus onset but that the process is all or nothing (de Jong, 2000). From this perspective, when pre-stimulus reconfiguration is successful the RT for that switch trial will be indistinguishable from repetition RTs. According to de Jong’s model, increasing the RSI/CSI increases the probability that participants will have successfully engaged the appropriate task set before stimulus onset. To illustrate this idea visually, de Jong (2000) used cumulative RT distributions. When individual (or group) RTs are plotted cumulatively they often produce a sigmoidal function. Functions produced from switch RTs are often displaced to the right of functions produced from repetition RTs (showing the overall slowing of RTs for switch trials). Also, compared to repetition functions, switch functions have a shallower gradient, which reflects the larger variance in switch trial RTs. Figure 8.2 shows prototypical cumulative RT functions that illustrate a preparation effect. There are several characteristics indicative of a preparation effect. First, repetition functions should be the same or similar for both RSIs (although see Monsell et al., 2003 for an exception). Second, switch trial functions for long RSIs should be displaced to the right of the repetition function, but overlap for the first one or two deciles. Finally, switch trial

functions for short RSI should be furthest to the right of the plot and should not overlap with the repetition function at any stage.

Although de Jong (2000) analysed cumulative RT distributions by fitting equations, the current experiment followed Monsell et al.'s (2003) procedure of presenting the distributions and providing a descriptive visual analysis. This procedure was used for the recognition task as it was the task of primary interest in this thesis. In order to construct the cumulative distribution functions, the raw RTs from every correct recognition trial (collapsed across participants) were conditionalised by RSI and transition. These data were then ranked-ordered and split into deciles. The median RT from each decile was used to plot cumulative RT functions. The cumulative distribution functions for correct repetition and switch decisions at the short (200 ms) and long (600 ms) RSIs are presented in Figure 8.3.

Functions for repetition trials were similar at both RSIs, consistent with typical preparation cumulative functions. The long RSI repetition and switch functions were positioned to the left of their respective short RSI functions. Functions for switch trials were displaced to the right, consistent with the slower RTs observed for switch RTs. However, there was very little difference between the long and short RSI switch functions, suggesting that the long RSI provided very little benefit in terms of increasing successful pre-stimulus engaging of task sets. Finally, the long RSI switch function did not overlap the repetition functions at the early deciles, as would be predicted by de Jong (2000) when preparation effects are observed. In sum, cumulative RT distributions showed no evidence of a preparation effect.

Experiment 4B

The 600-ms proposed asymptote for preparation effects (Rogers & Monsell, 1995; Meiran, 1996) has largely been accepted (and unchallenged), without consideration for the role of task complexity, retrieval demand, or experimental design. Given that the experiments in this thesis intentionally deviated from the typical tasks and designs found in the task switching literature, it is possible that Experiment 4A failed to obtain preparation effects because the RSIs were not long enough for advanced preparation to occur. As such, Experiment 4B used a short RSI of 600 ms and a long RSI of 1300 ms.

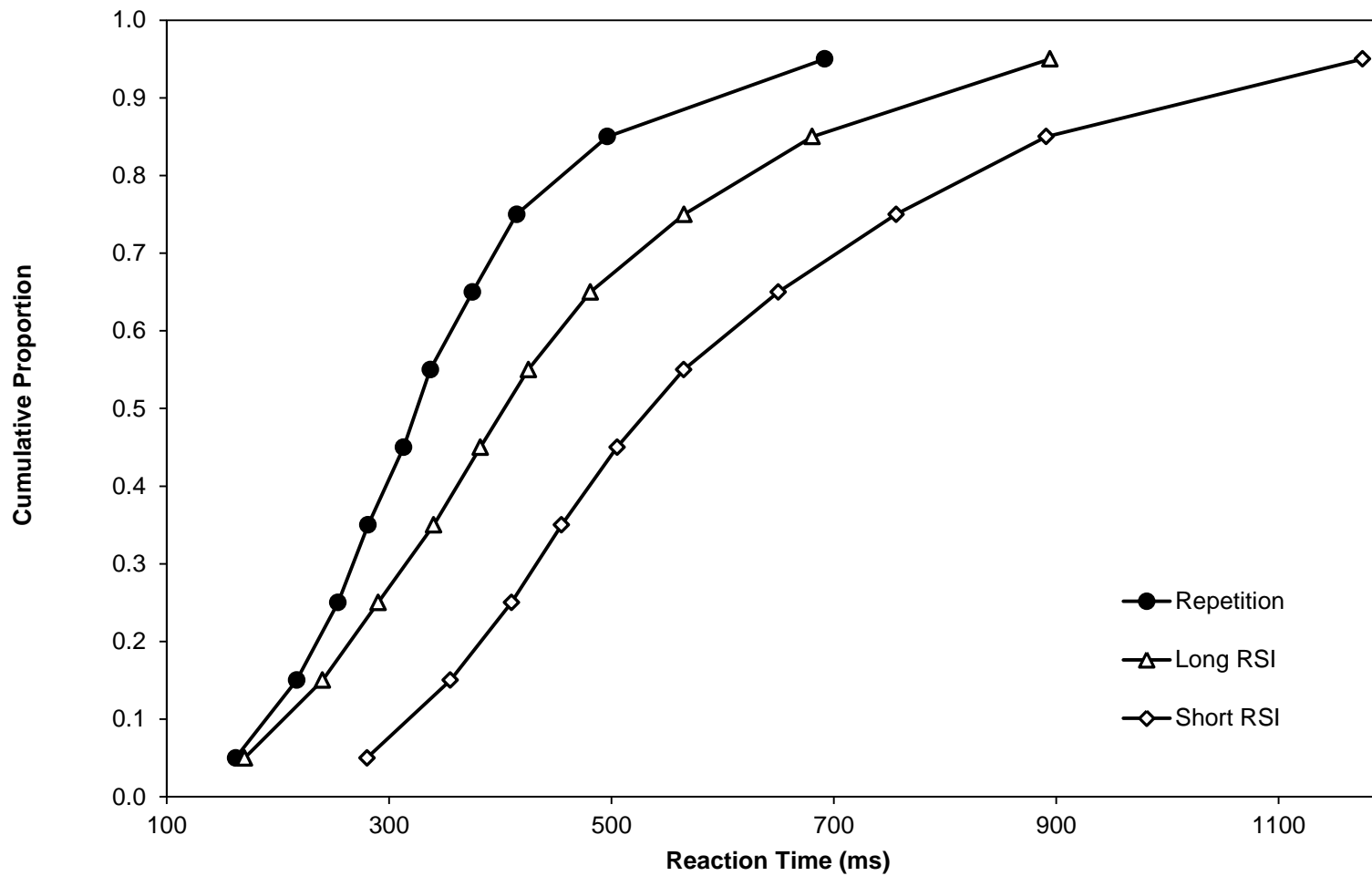


Figure 8.2 A prototypical cumulative RT distribution illustrating a preparation effect.

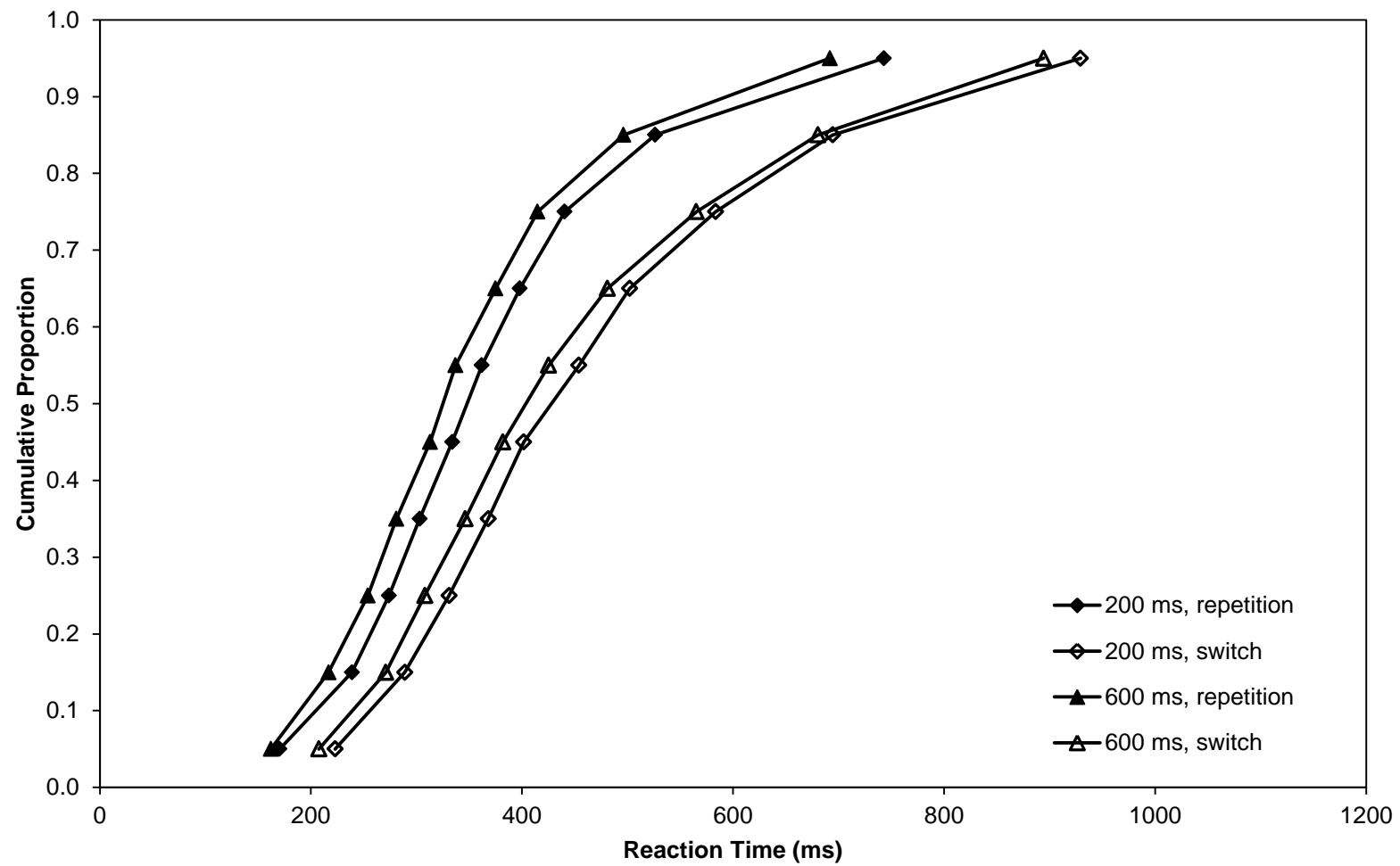


Figure 8.3. Cumulative RT distributions of correct recognition memory trials for Experiment 4A as a function of RSI and task transition.

Results and Discussion

Thirty-two participants met the performance requirements and were included in the final data analysis. The use of only correct responses for RT and switch cost analyses led to the exclusion of 16% of recognition trials and 10% of magnitude trials. The percentage of timeouts was very low ($< 1\%$ for both tasks).

Accuracy

Recognition task.

Table 8.5 shows the mean corrected accuracy (hit rate – false alarm rate) for recognition decisions as a function of RSI, transition, and test block. Mean corrected accuracy across all conditions was moderately high ($M = 0.70$, $SD = 0.17$) and there was no difference in overall accuracy between the two RSI order groups ($F = 0.00$, $p = 0.955$).

Effects of transition and test block.

The mean corrected accuracies reported in Table 8.5 were numerically higher for repetition conditions than those for switch conditions, but this difference was particularly small for the short RSI condition at Test 2. Corrected accuracy improved from the first to second test, suggesting that experience with the task improved participants' ability to perform the recognition task. This increase in corrected accuracy was slightly larger for switch conditions than those for repetition conditions, providing

Table 8.5

Means and Standard Deviations for Corrected Recognition Accuracy in Experiment 4B as a Function of RSI, Test Block, and Transition.

Transition	Long RSI (1300 ms)				Short RSI (600 ms)			
	Test 1		Test 2		Test 1		Test 2	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Repetition	0.70	0.21	0.75	0.12	0.69	0.26	0.74	0.17
Switch	0.64	0.24	0.72	0.1	0.63	0.34	0.73	0.16

Note. RSI = response-to-stimulus interval. Within-participant comparisons should be made across RSIs from Test 1 to Test 2 (i.e., Test 1 Long RSI, Test 2 Short RSI).

some evidence of a practice effect for recognition memory. The effects of transition and test block on corrected recognition accuracy were analysed using a 2 (test block) x 2 (transition) x 2 (RSI order) mixed ANOVA. Although the mean corrected accuracy for switch trials ($M = 0.72$, $SD = 0.18$) was numerically higher than that for repetition trials ($M = 0.68$, $SD = 0.19$), this difference was not significant ($F = 3.03$, $p = 0.092$). That is, there was no accuracy switch effect. Experience with the recognition task numerically increased corrected accuracy from Test 1 ($M = 0.67$, $SD = 0.24$) to Test 2 ($M = 0.73$, $SD = 0.12$), but this difference did not reach statistical significance ($F = 3.92$, $p = 0.057$). Critically, the interaction of test block and transition was not significant ($F = 0.73$, $p = 0.400$). There was no practice effect for corrected recognition accuracy.

Effects of RSI.

As can be seen in Table 8.5, there were no sizable differences in mean corrected accuracy between the long and the short RSIs. A 2 (RSI) x 2 (transition) x 2 (RSI order) mixed ANOVA on corrected recognition accuracy confirmed that the main effect of RSI was not significant ($F = 0.04$, $p = 0.844$). The interaction of RSI and transition was also not significant ($F = 0.04$, $p = 0.838$). Providing additional time prior to stimulus onset did not improve recognition accuracy overall or for switch trials in particular. Corrected recognition accuracy showed no evidence of a preparation effect. No other effects or interactions were significant ($F_s < 0.03$).

Magnitude task.

Table 8.6 shows the mean accuracy of magnitude judgements as a function of RSI, transition, decision, and test block. Overall accuracy was very high ($M = 0.90$, $SD = 0.07$) and there was no difference in overall accuracy for the two RSI order groups ($F = 0.87$, $p = 0.357$).

Effects of transition and test block.

As can be seen in Table 8.6, the accuracy of repetition trials was very similar (if slightly higher) to that of switch trials. There was little evidence of an accuracy switch effect. The accuracy of magnitude judgements was generally higher in Test 2 than that in Test 1, particularly in the case of Short RSIs. However, the benefit of experience to overall magnitude accuracy was quite subtle and was largely equivalent for repetition and switch trials. That is, magnitude task accuracy showed no evidence of a practice effect. The effects of transition and test block were analysed using a 2 (test block) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. As a reminder, the effects

Table 8.6

Means and Standard Deviations for Magnitude Judgement Accuracy in Experiment 4B as a Function of RSI, Test Block, Transition, and Decision.

Transition and Decision		Long RSI (1300 ms)				Short RSI (200 ms)			
		Test 1		Test 2		Test 1		Test 2	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Repetition	More	0.95	0.05	0.96	0.05	0.88	0.16	0.96	0.07
	Less	0.89	0.10	0.89	0.15	0.85	0.13	0.92	0.09
Switch	More	0.91	0.10	0.95	0.08	0.89	0.12	0.93	0.10
	Less	0.88	0.12	0.91	0.11	0.80	0.20	0.87	0.12

Note. RSI = response-to-stimulus interval. Within-participant comparisons should be made across RSIs from Test 1 to Test 2 (i.e., Test 1 Long RSI, Test 2 Short RSI).

and interactions of item type are not reported as response-to-key allocations were not balanced for the magnitude task. Although the mean magnitude judgement accuracy was higher for repetition trials ($M = 0.91$, $SD = 0.06$) than that for switch trials ($M = 0.89$, $SD = 0.08$), this difference did not reach statistical significance ($F = 3.51$, $p = 0.071$). The main effect of test block was significant, $F(1, 30) = 6.28$, $MSE = 0.02$, $p = 0.018$, $\eta_p^2 = 0.17$. The accuracy of participants' magnitude judgements improved from Test 1 ($M = 0.88$, $SD = 0.09$) to Test 2 ($M = 0.92$, $SD = 0.07$), illustrating that experience with the magnitude task improved participants ability to make correct magnitude judgements. The interaction of test block and transition was not significant ($F = 0.07$, $p = 0.793$), confirming that there was no practice effect for magnitude task accuracy.

Effects of RSI.

For the first test, the mean accuracy of the magnitude task was higher for the long RSI condition than that for the short RSI condition (see Table 8.6). However, by the second test performance looked similar for both RSI conditions. Regarding the size of the accuracy switch effect, there was no discernable pattern of differences for short and long RSIs. The effect of RSI was analysed using a 2 (RSI) x 2 (transition) x 2 (RSI order) x 2 (item type) mixed ANOVA. Neither the main effect of RSI ($F = 2.51$, $p = 0.124$) nor the interaction of RSI and transition ($F = 1.57$, $p = 0.220$) were significant.

The length of the RSI had no effect on overall accuracy, and there was no evidence of preparation effects on magnitude accuracy. No other effects or interactions were significant ($F_s < 0.92$).

Reaction Times

As a reminder, the RT results are separated into several subsections: (1) analyses of correct recognition RTs, (2) analyses of correct magnitude RTs, and (3) analyses of RT switch costs (switch RTs – repetition RTs). A comprehensive table of RT performance for both tasks can be found in Appendix D. When the 400-ms RSD is taken into account, the overall mean recognition RT ($M = 402.18$, $SD = 110.17$) was fairly similar to the overall mean magnitude RT ($M = 760.81$, $SD = 25.40$). The order of RSI presentation had no significant effect on overall RTs for the recognition ($F = 0.13$, $p = 0.718$) and magnitude ($F = 0.18$, $p = 0.671$) tasks.

Recognition task.

The effects of transition and test block on correct recognition RTs were analysed using a 2 (test block) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. The main effect of transition was significant, $F(1, 30) = 73.61$, $MSE = 6593.83$, $p < 0.001$, $\eta_p^2 = 0.71$. Correct recognition RTs were significantly slower for switch trials ($M = 445.72$, $SD = 127.85$) than those for repetition trials ($M = 358.64$, $SD = 97.85$), illustrating a typical RT switch effect. The main effect of item type was also significant, $F(1, 30) = 20.06$, $MSE = 7946.66$, $p < 0.001$, $\eta_p^2 = 0.40$. Overall, correct new responses ($M = 427.14$, $SD = 109.06$) were significantly slower than correct old responses ($M = 377.23$, $SD = 119.85$). The main effect of test block was significant, $F(1, 30) = 8.23$, $MSE = 15770.54$, $p = 0.007$, $\eta_p^2 = 0.22$. The mean RT for correct recognition trials was significantly faster for Test 2 ($M = 379.66$, $SD = 129.14$) than that for Test 1 ($M = 424.70$, $SD = 107.42$). Experience with the recognition task allowed participants to produce correct recognition responses more quickly.

The effect of RSI length on correct recognition RTs was analysed using a 2 (RSI) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. Although the main effect of RSI did not reach statistical significance ($F = 3.59$, $p = 0.068$), there was a significant interaction of RSI and item type, $F(1, 30) = 5.74$, $MSE = 2689.44$, $p = 0.023$, $\eta_p^2 = 0.16$. When the effects of RSI were analysed for ‘old’ and ‘new’ words separately, the main effect of RSI was significant for ‘old’ words [$F(1, 30) = 8.00$, $MSE = 8196.34$, $p = 0.008$, $\eta_p^2 = 0.26$] but not ‘new’ words ($F = 0.63$, $p = 0.434$). The mean

RT for ‘old’ decisions was significantly faster for the short RSI condition ($M = 354.60$, $SD = 109.15$) than that for the long RSI condition ($M = 399.86$, $SD = 144.61$). By contrast, the mean RT for ‘new’ decisions was similar for the short ($M = 420.04$, $SD = 102.67$) and long ($M = 434.23$, $SD = 135.58$) RSI conditions. The slower ‘old’ mean RT for the long RSI test block was quite counterintuitive. One possible explanation is that the average delay between encoding and retrieval of ‘old’ words was longer in the long RSI test block. Correctly identifying ‘old’ words may have required a more time-consuming search of memory in the long RSI test block than that in the short RSI test block. No other effects or interactions were significant ($F_s < 0.07$).

Magnitude task.

The effects of transition and test block on correct magnitude RTs were analysed using a 2 (test block) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. Again, the main effects and interactions of item type are not reported as response key allocation was not balanced. The main effect of transition was significant, $F(1, 30) = 8.50$, $MSE = 10454.61$, $p = 0.007$, $\eta_p^2 = 0.22$. The mean correct magnitude RT was significantly slower for switch trials ($M = 779.43$, $SD = 153.82$) than that for repetition trials ($M = 742.18$, $SD = 142.32$), demonstrating a RT switch effect. The main effect of test block was also significant, $F(1, 30) = 27.05$, $MSE = 20046.11$, $p < 0.001$, $\eta_p^2 = 0.47$. The mean magnitude RT was slower in the first test ($M = 806.82$, $SD = 173.08$) than that in the second test ($M = 714.79$, $SD = 127.88$). Experience with the magnitude task allowed participants to correctly respond more quickly.

The effect of RSI length on correct magnitude judgement RTs was analysed using a 2 (RSI) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. The main effect of RSI on correct dot RTs was not significant ($F = 1.81$, $p = 0.189$). The length of the RSI had no impact on the speed of participants’ correct magnitude judgements. No other effects or interactions were significant ($F_s < 0.04$).

RT switch costs.

The mean RT switch costs (switch RT – repetition RT) are reported in Table 8.7 as a function of test block, RSI, task and item type. The RT switch costs were highly variable, particularly for the first test. Furthermore, the reported standard deviations were very large—in some cases over double the size of the mean. The extent to which a task switch negatively affected participants’ performance varied dramatically. The size of the RT switch cost was generally smaller for the magnitude task than that for the

Table 8.7

Means and Standard Deviations for the RT Switch Costs in Experiment 4B as a Function of RSI, Test Block, Task and Decision.

Task and Decision		Long RSI (1300 ms)				Short RSI (200 ms)			
		Test 1		Test 2		Test 1		Test 2	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Recognition	New	86	88	74	76	100	78	71	64
	Old	129	164	81	57	98	80	59	101
Magnitude	More	78	159	70	72	43	148	8	64
	Less	16	96	-4	95	51	138	38	73

Note. RSI = response-to-stimulus interval. Within-participant comparisons should be made across RSIs from Test 1 to Test 2 (i.e., Test 1 Long RSI, Test 2 Short RSI).

recognition task. The order of RSI presentation had no effect on the RT switch costs ($F = 0.05, p = 0.818$).

Effects of test block.

As can be seen in Table 8.7, the size of the RT switch cost generally decreased from the first to second test block, suggesting that practice reduced the effect of a task switch. The effect of test block was analysed using a 2 (test block) x 4 (item type) x 2 (RSI order) mixed ANOVA. A Greenhouse-Geisser correction was used for all analyses of item type as the assumption of sphericity was violated. There was a significant main effect of item type, $F(3, 90) = 4.73, MSE = 16568.51, p = 0.009, \eta_p^2 = 0.14$. Pairwise comparisons using Bonferroni corrections showed that the mean RT switch cost for correct ‘old’ recognition decisions ($M = 91.63, SD = 65.19$) was significantly larger than that for correct ‘less’ dot decisions ($M = 25.06, SD = 93.13$), but did not differ from correct ‘new’ recognition decisions ($M = 81.54, SD = 62.21$) or correct ‘more’ dot decisions ($M = 49.46, SD = 98.72$). No other item types differed significantly from each other. The mean RT switch cost for Test 1 ($M = 74.93, SD = 68.54$) was significantly larger than that for Test 2 ($M = 49.41, SD = 37.99$), $F(1, 30) = 4.32, MSE = 9645, p = 0.050, \eta_p^2 = 0.13$. As can be seen in Figure 8.4, practice with the tasks and switching between them in the first test block (regardless of RSI) produced a decrease in RT switch costs for the second test block. That is, there was a significant practice effect.

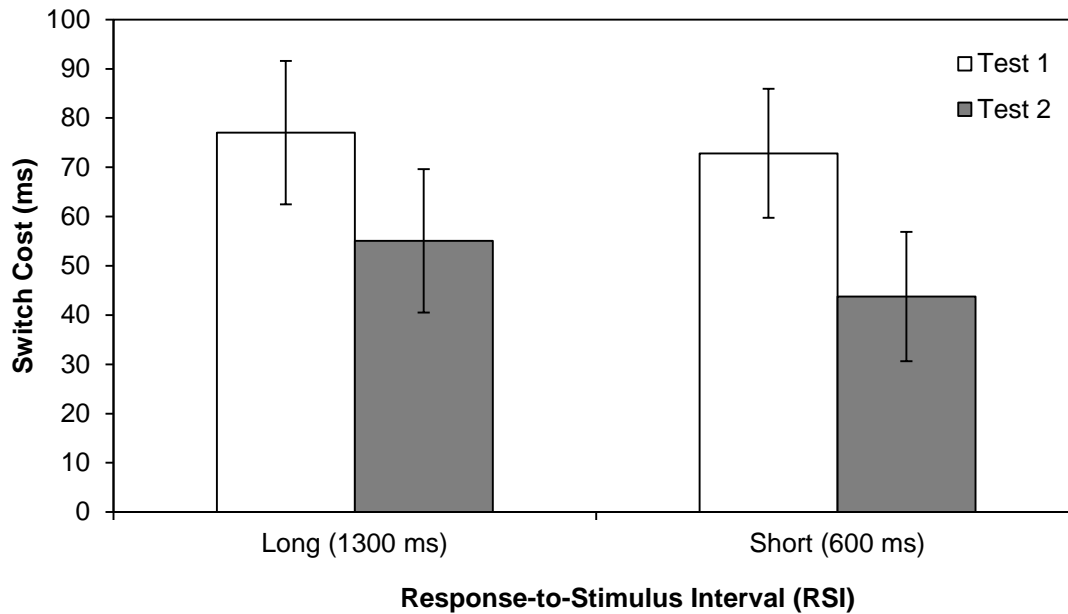


Figure 8.4. Mean RT switch costs for Experiment 4B as a function of RSI and test block. Each test block used a single RSI, and order of RSI presentation was balanced across participants. Error bars represent S.E.

Effects of RSI.

Table 8.7 showed no consistent pattern across RSIs, suggesting that there was limited evidence of a preparation effect. A 2 (RSI) \times 2 (RSI order) \times 4 (item type) mixed ANOVA was conducted on RT switch costs for recognition and magnitude tasks. Although the main effect of RSI was not significant ($F = 0.40, p = 0.532$), there was a significant interaction of RSI and trial type, $F(3, 90) = 3.18, MSE = 9685.14, p = 0.042, \eta_p^2 = 0.10$. The interaction was followed up by conducting paired sample t -tests for RT switch costs at the short and long RSIs for each item type. As can be seen in Figure 8.5, RSI had no significant effect on RT switch cost for ‘new’ recognition memory decisions ($t = -0.34, p = 0.739$) or ‘old’ recognition memory decisions ($t = 0.87, p = 0.391$). The effect of RSI on magnitude judgements was less clear. For ‘less’ magnitude judgements, the mean RT switch cost was larger for the short RSI condition ($M = 44.47, SD = 94.69$) than that for the long RSI condition ($M = 5.64, SD = 109.22$), $t(31) = -2.43, MSE = 16.01, p = 0.021$. By contrast, for ‘more’ magnitude judgements, the mean RT switch cost was smaller for the short RSI condition ($M = 25.23, SD = 113.14$) than that for the long RSI condition ($M = 73.69, SD = 121.46$), $t(31) = 2.09, MSE = 23.17, p = 0.045$. Given that key allocation for magnitude judgements was not balanced across participants and that there was no theoretical reason to hypothesise this interaction, it is

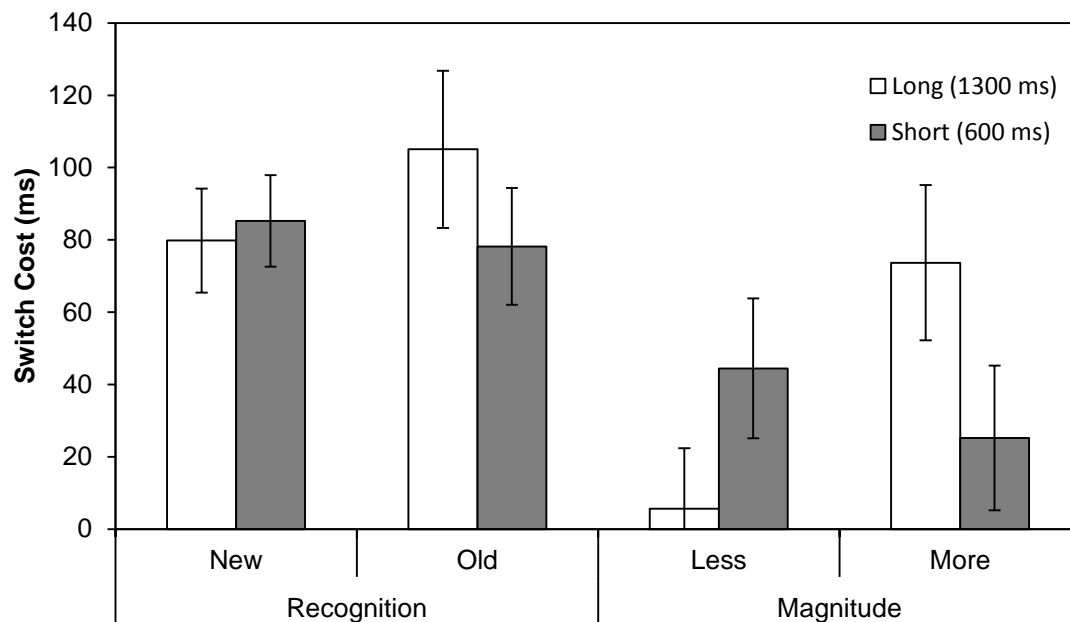


Figure 8.5. Mean RT switch costs for Experiment 4B as a function of RSI, task and item type. Error bars represent S.E.

difficult to interpret these findings as support for preparation effects. No other effects or interactions were significant ($F_s < 0.32$). Finally, the main effect of RSI on the RT switch cost was not significant when Test 1 ($F = 0.03$, $p = 0.863$) and Test 2 ($F = 0.71$, $p = 0.407$) were analysed separately. Thus there was no consistent evidence of a preparation effect on RT switch costs for either task.

Cumulative RT distributions.

The cumulative distribution functions for correct repetition and switch decisions at the short (600 ms) and long (1300 ms) RSIs are presented in Figure 8.6. The repetition trial functions were positioned to the left and the switch trial functions were positioned to the right. However, the displacement of repetition and switch functions (regardless of RSI) was not large, particularly at the earliest deciles. For both transitions, the long and short RSI functions were very close (if not overlapping) for the first few deciles of the distribution. There was a larger spread between the two RSIs at later deciles, but still less so than represented in the prototypical preparation effect distribution illustrated in Figure 8.2 (p. 109). Inconsistent with the predicted pattern for preparation effects, the data points for the short RSI functions were positioned to the left

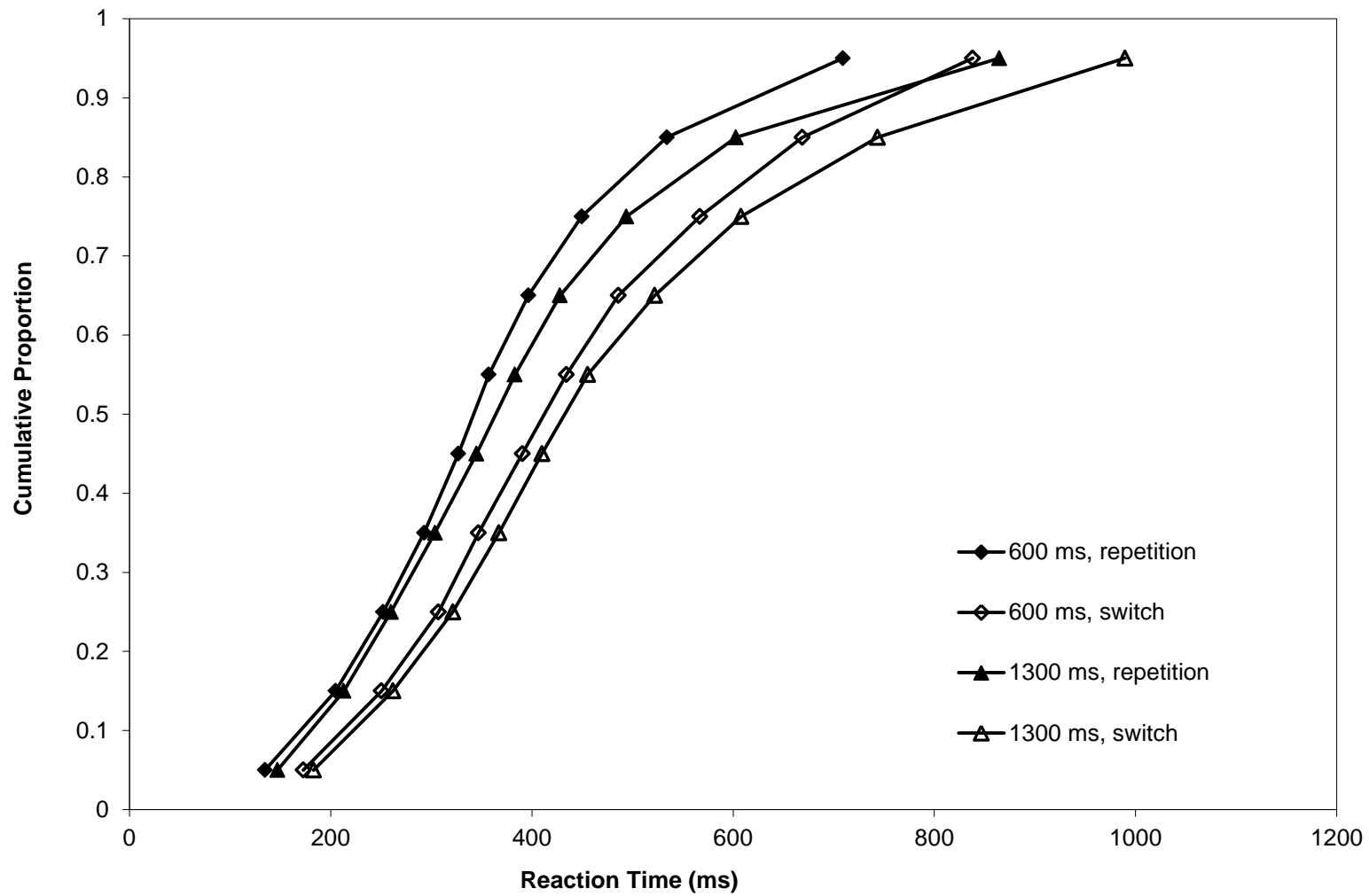


Figure 8.6. Cumulative RT distributions of correct recognition memory trials for Experiment 4B as a function of RSI and task transition.

of the long RSI functions, particularly at the later deciles. Finally, repetition and switch functions for the long RSI condition did not overlap at the first deciles as would be expected in a typical preparation effect distribution. Considered together, the features of the cumulative RT distributions provided no evidence of preparation effects.

Conclusions for Experiments 4A and 4B

In Experiments 4A and 4B, the length of the RSI used at test was manipulated between test blocks (200 ms vs 600 ms; 600 ms vs 1300 ms) to test for evidence of preparation effects—a decrease in the size of switch effects with longer RSIs. Additionally, switch effects were compared between test blocks (Test 1, Test 2) to test for evidence of practice effects—a decrease in the size of switch effects as participants gain experience with the experiment. Both experiments reported significant RT switch effects for the recognition and magnitude tasks. Participants took significantly longer to make correct decisions when they had switched tasks compared to when they had repeated tasks. While both tasks produced means suggestive of an accuracy switch effect, this effect was only significant in Experiment 4A. This result was not wholly unexpected because switch effects on accuracy are less consistently observed than switch effects on RT, which are astoundingly robust (see Monsell, 2003; Vandierendonck, 2010). Considered together, the results of Experiment 4A and 4B showed that switching tasks produced slower and (sometimes) less accurate performance on the recognition and magnitude tasks.

Although the mean RT switch costs were smaller in the second test for both experiments, only Experiment 4B produced a significant practice effect. The inconsistency between the two experiments may be a consequence of the highly variable performance, particularly when switching tasks. The range in the size of RT switch costs clearly shows that there were large differences in participants' ability to switch tasks. At least with regard to Experiment 4B, the results were consistent with previous findings (Jersild, 1927; Rogers & Monsell, 1995; Koch, 2005; Minear & Shah, 2008). The RT switch costs decreased as participants' became more experienced with the tasks and with switching between them, reflecting a strengthening of associations between task cues (in this case, internally produced) and task sets.

Neither Experiment 4A nor Experiment 4B produced a significant main effect of RSI (that is, a preparation effect) on RT switch costs. However, Experiment 4B produced significant interaction of RSI and item type, driven by differences in the RT

switch cost for the magnitude task. One decision ('less') produced results consistent with a preparation effect and one decision ('more') produced results inconsistent with a preparation effect. Given that the response keys were not balanced for the magnitude task and that there is no theoretical reason to predict such an interaction, it is difficult to consider this result as support for the hypothesised preparation effect. Thus, Experiments 4A and 4B produced little evidence that participants used long RSIs to prepare for task switches in advance of stimulus onset.

These experiments covered a typical choice in RSIs (4A: 200 ms vs 600 ms) with an extension to longer RSIs (4B: 600 ms vs 1300 ms). However, preparation effects were not observed. The results of Experiments 4A and 4B suggest either that participants chose not use the RSIs (regardless of length) to prepare or that the long RSIs used in these experiments were not sufficiently long for participants to reconfigure or activate these specific task sets in advance of stimulus onset. Regardless of explanation, this failure to observe preparation effects may be the consequence of the novel (and arguably more naturalistic) experimental characteristics used in these experiments. The tasks used in these experiments had very different task sets and the size of each task set (i.e., the number of items) was considerably larger than is typical of the task switching literature. Furthermore, in the absence of external task cues, participants needed to monitor and generate task cues internally—a process that may take considerably longer than interpreting a visually presented task cue. In light of these and other less-typical characteristics of the current experiments, it may be valuable to extend the length of the long RSI beyond those previously used in order to definitively determine if (and when) participants use an RSI to prepare for known switches between the recognition and magnitude tasks.¹⁰

Experiment 5A and 5B

Experiment 5A and 5B used a short RSI of 200 ms and a long RSI of 2600 ms. As such, the short RSI was well below Rogers and Monsell's (1995) preparation effect asymptote (600 ms), and the long RSI provided plenty of time for participants to anticipate and prepare for upcoming task switches. There was one critical difference between Experiment 5A and 5B. To address Altmann's (2004b) claim that multiple

¹⁰ Meiran et al. (2000) used CSIs of up to 3032 ms when examining the role of advanced preparation and task set decay on task switching performance with simple spatial orientation tasks. It seems reasonable to be equally thorough when investigating evidence for preparation effects for the current recognition and magnitude tasks.

Table 8.8

Means and Standard Deviations for Corrected Recognition Accuracy in Experiment 5A as a Function of RSI, Test Block, and Transition.

Transition	Long RSI (2600 ms)				Short RSI (200 ms)			
	Test 1		Test 2		Test 1		Test 2	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Repetition	0.73	0.19	0.61	0.22	0.73	0.20	0.80	0.14
Switch	0.69	0.21	0.64	0.26	0.66	0.25	0.74	0.13

Note. RSI = response-to-stimulus interval. Within-participant comparisons should be made across RSIs from Test 1 to Test 2 (i.e., Test 1 Long RSI, Test 2 Short RSI).

RSIs need to be experienced in order to observe preparation effects, the practice tests in Experiment 5B used both RSIs. Thus, while participants in Experiment 5A experienced only one RSI in each study-test block, participants in Experiment 5B were experienced both RSIs in the practice test prior to each actual test.

Experiment 5A

Results and Discussion

Thirty-two participants met the performance requirements and were included in the final data analysis.¹¹ The use of only correct responses for RT and Switch cost analyses led to the exclusion of 16% of recognition trials and 15% of magnitude trials. The percentage of timeouts for the recognition task (1%) and the magnitude task (2%) were slightly higher than those for Experiments 4A and 4B but still quite low.

Accuracy

Recognition task.

Table 8.8 shows the mean corrected accuracy (hit rate – false alarm rate) for recognition decisions as a function of RSI, transition, and test block. Mean corrected accuracy across all conditions was moderate ($M = 0.70$, $SD = 0.18$) and there was no difference in overall accuracy between the two RSI order groups ($F = 1.63$, $p = 0.211$).

Effects of transition and test block.

¹¹ Two participants had clearly reversed the response keys for dot decisions in the first test block. These accuracy and RT data were corrected before being included in the analysis.

As can be seen in Table 8.8, the mean corrected recognition accuracies were generally lower for switch conditions than those for repetition conditions, except for the Test 2, long RSI condition where corrected accuracy was slightly higher for the switch trial than that for the repetition trial. That is, the evidence for an accuracy switch effect was mixed. Neither the overall corrected accuracy nor the difference between repetition and switch trial accuracy showed a consistent pattern of improvement from the first to second test blocks. There was no evidence of preparation effects on corrected accuracy. The effects of transition and test block on corrected recognition accuracy were analysed using a 2 (test block) x 2 (transition) x 2 (RSI order) mixed ANOVA. Although the mean corrected accuracy for repetition trials ($M = 0.72$, $SD = 0.17$) was higher than that for switch trials ($M = 0.68$, $SD = 0.20$), the main effect of transition was not significant ($F = 3.69$, $p = 0.064$). Performance on the Test 2, long RSI condition likely played a large role in producing the non-significant accuracy switch effect. While the main effect of test block was not significant ($F = 0.05$, $p = 0.830$), there was a significant interaction of test block and RSI order, $F(1, 30) = 9.18$, $MSE = 0.01$, $p = 0.005$, $\eta_p^2 = 0.23$. When the long RSI was used in the first test, the increase in accuracy from Test 1 ($M = 0.71$, $SD = 0.18$) to Test 2 ($M = 0.77$, $SD = 0.12$) was not significant ($F = 3.69$, $p = 0.074$). Most unusually, when the short RSI was used in the first test, the corrected accuracy of recognition decisions decreased significantly from Test 1 ($M = 0.69$, $SD = 0.21$) to Test 2 ($M = 0.62$, $SD = 0.22$), $F(1,15) = 5.67$, $MSD = 0.01$, $p = 0.031$, $\eta_p^2 = 0.27$. The low corrected accuracy of the Test 2, long RSI condition likely produced this interaction. Neither RSI order condition showed a significant benefit of experience with the recognition task to overall accuracy of recognition decisions. The interaction of test block and transition was not significant ($F = 0.64$, $p = 0.431$), indicating that there was no practice effect on recognition accuracy.

Effects of RSI.

Table 8.8 shows that, for the first test block, the mean corrected accuracy of recognition decisions was very similar for the short and long RSI conditions. However, for the second test block the mean corrected accuracy was slightly higher for the short RSI condition than that for the long RSI condition. The difference between repetition and switch trial corrected accuracy was slightly larger for the short RSI conditions than that for the long RSI conditions, suggestive of a preparation effect for recognition accuracy. The effect of RSI on corrected recognition accuracy was analysed using a 2 (RSI) x 2 (transition) x 2 (RSI order) mixed ANOVA. The main effect of RSI was

Table 8.9

Means and Standard Deviations for Magnitude Judgement Accuracy in Experiment 5A as a Function of RSI, Test Block, Transition, and Decision.

Transition and Decision		Long RSI (2600 ms)				Short RSI (200 ms)			
		1st		2nd		Test 1		Test 2	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Repetition	More	0.92	0.13	0.97	0.03	0.79	0.24	0.92	0.07
	Less	0.89	0.13	0.92	0.10	0.79	0.14	0.90	0.06
Switch	More	0.86	0.16	0.86	0.15	0.71	0.20	0.89	0.11
	Less	0.92	0.11	0.91	0.11	0.74	0.25	0.85	0.15

Note. RSI = response-to-stimulus interval. Within-participant comparisons should be made across RSIs from Test 1 to Test 2 (i.e., Test 1 Long RSI, Test 2 Short RSI).

significant, $F(1, 30) = 9.18$, $MSE = 0.01$, $p = 0.005$, $\eta_p^2 = 0.23$. The mean corrected accuracy for the long RSI conditions ($M = 0.67$, $SD = 0.21$) was significantly lower than that for short RSI conditions ($M = 0.73$, $SD = 0.17$). As was suggested in Experiment 4B, this may reflect the cost of the additional time delay between study and test in the long RSI block. Critically, the interaction of RSI and transition was not statistically significant ($F = 2.29$, $p = 0.141$). The size of the accuracy switch effect showed no difference between long and short RSI conditions. The corrected recognition accuracy data provided no evidence of a preparation effect. No other effects or interactions were significant ($F_s < 0.64$).

Magnitude task.

Table 8.9 shows the mean accuracy of magnitude judgements as a function of RSI, transition, decision and test block. Overall accuracy was high ($M = 0.87$, $SD = 0.08$) and similar to previous experiments. Although the overall magnitude task accuracy was numerically lower for the short-long RSI order condition, the difference was not significant ($F = 3.64$, $p = 0.066$).¹²

Effects of transition and test block.

¹² This was likely driven by poor overall performance in the first test block of the short-long RSI condition that was not isolated to any individual participant.

Table 8.9 shows an overall trend towards an accuracy switch effect, with the exception of ‘less’ decisions for the long RSI condition in Test 1. The accuracy of magnitude judgements showed evidence of improvement from the first to second tests. However, the size of the accuracy switch effect showed no consistent improvement from the first to the second test blocks, indicating that there was no practice effect for magnitude accuracy. The effects of transition and test block were analysed using a 2 (test block) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. As with earlier experiments in this chapter, the effects and interactions of item type are not reported. The main effect of transition was significant, $F(1, 30) = 15.80$, $MSE = 0.01$, $p < 0.001$, $\eta_p^2 = 0.35$. Magnitude judgement accuracy was significantly higher for repetition trials ($M = 0.89$, $SD = 0.08$) than that for switch trials ($M = 0.84$, $SD = 0.11$), demonstrating an accuracy switch effect. A significant main effect of test block [$F(1, 30) = 16.25$, $MSE = 0.02$, $p < 0.001$, $\eta_p^2 = 0.35$] showed that the accuracy of participants’ magnitude judgements improved from Test 1 ($M = 0.83$, $SD = 0.13$) to Test 2 ($M = 0.90$, $SD = 0.06$). This main effect of test block was qualified by a significant interaction with RSI order, $F(1, 30) = 18.37$, $MSE = 0.02$, $p < 0.001$, $\eta_p^2 = 0.35$. When the long RSI was used in the first test, the increase in accuracy from Test 1 ($M = 0.90$, $SD = 0.09$) to Test 2 ($M = 0.89$, $SD = 0.05$) was not significant ($F = 0.04$, $p = 0.836$). However, when the short RSI was used in the first test, the increase in accuracy from Test 1 ($M = 0.76$, $SD = 0.16$) to Test 2 ($M = 0.92$, $SD = 0.08$) was significant, $F(1, 15) = 27.20$, $MSD = 0.03$, $p < 0.001$, $\eta_p^2 = 0.65$. The high mean accuracy of magnitude judgements in the long RSI conditions (and the low accuracy of the short RSI, Test 1 condition) likely drove the interaction of test block and RSI order. With regard to evidence of practice effects on magnitude judgement accuracy, the interaction of test block and transition was not significant ($F = 0.19$, $p = 0.664$).

Effects of RSI.

Although magnitude judgements were more accurate for the long RSI condition than that for the short RSI condition, the data reported in Table 8.9 did not show any difference consistent with a preparation effect. If anything, the accuracy switch effects were smaller for the short RSI conditions than those for the long RSI conditions. The effect of RSI on magnitude task accuracy was analysed using a 2 (RSI) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. The mean magnitude judgement accuracy for the long RSI condition ($M = 0.91$, $SD = 0.08$) were significantly higher than that for the short RSI condition ($M = 0.83$, $SD = 0.12$), $F(1, 30) = 18.37$, $MSE = 0.02$, $p < 0.001$,

$\eta_p^2 = 0.38$. Long RSIs produced more accurate magnitude task performance. However, the interaction of RSI and transition was not significant ($F = 0.58$, $SD = 0.451$). Thus, there was no evidence of a preparation effects on magnitude accuracy. No other effects or interactions were significant ($F_s < 1.89$).

Reaction Times

As with Experiments 4A and 4B, RT data will be reported in several subsections: (1) analyses of correct recognition RTs, (2) analyses of correct magnitude RTs, and (3) analyses of RT switch costs (switch RTs – repetition RTs). A comprehensive table of RTs for both tasks can be found in Appendix D. The overall RT for correct recognition decisions ($M = 481.84$, $SD = 121.41$) was slightly longer and more variable than previous experiments in this chapter. When the 400-ms RSD is taken into account, the recognition RT remained similar to that for the magnitude task ($M = 821.74$, $SD = 141.21$). The order of RSI presentation had no effect on the overall mean correct RTs for the recognition task ($F = 1.29$, $p = 0.265$). However, the overall mean correct RT for the magnitude task was significantly longer when the short RSI was used in the first test ($M = 872.64$, $SD = 245.12$) than that when the long RSI was used in the first test ($M = 770.85$, $SD = 140.26$), $F(1, 15) = 0.12$, $MSE = 159513.21$, $p = 0.050$, $\eta_p^2 = 0.12$. Critically, this slower performance in the short-long RSI order condition was *not* limited to their first test block, where participants' accuracy was also relatively poor. Thus, participants who completed the short RSI test block first consistently struggled to make correct magnitude judgements quickly.

Recognition task.

The effects of transition and test block on correct recognition RTs were analysed using a 2 (test block) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. The main effect of transition was significant, $F(1,30) = 70.69$, $MSE = 13370.71$, $p < 0.001$, $\eta_p^2 = 0.70$. Correct recognition RTs were significantly slower on switch trials ($M = 541.60$, $SD = 146.36$) than those on repetition trials ($M = 420.07$, $SD = 106.14$), demonstrating an RT switch effect. Correctly identifying 'new' words ($M = 501.98$, $SD = 121.60$) took significantly longer than correctly identifying 'old' words ($M = 459.70$, $SD = 127.42$), $F(1,30) = 17.05$, $MSE = 6713.83$, $p < 0.001$, $\eta_p^2 = 0.36$. The mean correct RT for the recognition task was slower for Test 1 ($M = 531.94$, $SD = 152.90$) than that for Test 2 ($M = 429.73$, $SD = 113.34$), $F(1,30) = 24.30$, $MSE = 27518.80$, $p < 0.001$, $\eta_p^2 = 0.45$. Experience with the recognition task sped correct responses regardless of task

transition. The main effect was qualified by a significant interaction with RSI order, $F(1,30) = 8.28$, $MSE = 27518.80$, $p = 0.007$, $\eta_p^2 = 0.22$. When the short RSI was used in the first test, the decrease in mean RT from Test 1 ($M = 526.44$, $SD = 209.28$) to Test 2 ($M = 483.42$, $SD = 160.78$) was not significant ($F = 1.79$, $p = 0.201$). However, when the long RSI was used in the first test, the decrease in mean RT from Test 1 ($M = 537.45$, $SD = 222.96$) to Test 2 ($M = 375.59$, $SD = 159.79$) was significant, $F(1, 15) = 37.09$, $MSE = 22603.24$, $p < 0.001$, $\eta_p^2 = 0.71$. Slower performance for the long RSI condition could have driven the interaction of test block and RSI order by inflating the observed practice effect when the long RSI was presented first. Alternatively, the slower overall performance of participants who completed short RSI test block first may have masked any general improvement due to experience.

The effect of RSI length on correct recognition RTs was analysed using a 2 (RSI) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. Inconsistent with earlier experiments in this chapter, the main effect of RSI was significant, $F(1, 30) = 8.28$, $MSE = 27518.80$, $p = 0.007$, $\eta_p^2 = 0.22$. The mean correct recognition RT was *slower* for the long RSI condition ($M = 510.66$, $SD = 137.44$) than that for the short RSI condition ($M = 451.01$, $SD = 131.66$). However, the main effect was qualified by an interaction with item type, $F(1, 30) = 7.66$, $MSE = 5109.86$, $p = 0.010$, $\eta_p^2 = 0.20$. Follow-up analyses showed that correct recognition RTs for ‘old’ words were slower for the long RSI conditions ($M = 501.89$, $SD = 149.69$) than those for the short RSI conditions ($M = 417.50$, $SD = 137.22$), $F(1, 30) = 13.00$, $MSE = 17530.60$, $p < 0.001$, $\eta_p^2 = 0.30$. By contrast, correct recognition RTs for ‘new’ words were no different for the long RSI ($M = 519.44$, $SD = 135.40$) and the short RSI ($M = 484.52$, $SD = 137.07$) conditions ($F = 2.58$, $p = 0.12$). Contrary to what is normally observed in experiments that manipulate RSI (e.g., Altmann, 2002) and Experiments 4A and 4B, the long RSI actually slowed recognition memory responses, particularly for ‘old’ decisions. No other effects or interactions were significant ($F_s < 0.64$).

Magnitude task.

The effects of transition and test block on correct magnitude RTs were analysed using a 2 (test block) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. Again, the main effects and interactions of item type are not reported. The main effect of transition was significant, $F(1, 30) = 18.14$, $MSE = 8681.86$, $p < 0.001$, $\eta_p^2 = 0.38$. Typical of a RT switch effect, the mean RT for correct magnitude judgements was

faster for repetition trials ($M = 796.94$, $SD = 132.89$) than that for switch trials ($M = 846.55$, $SD = 156.17$). Participants' mean correct RT decreased from Test 1 ($M = 864.12$, $SD = 170.08$) to Test 2 ($M = 779.37$, $SD = 139.55$) as they gained experience with the task, $F(1, 30) = 13.49$, $MSE = 34068.03$, $p = 0.001$, $\eta_p^2 = 0.31$. The interaction of test block and order of RSI was also significant, $F(1, 30) = 4.18$, $MSE = 34068.03$, $p = 0.050$, $\eta_p^2 = 0.12$. When the short RSI was used in the first test, the decrease in mean RT from Test 1 ($M = 891.43$, $SD = 197.70$) to Test 2 ($M = 853.84$, $SD = 183.82$) was not significant ($F = 1.05$, $p = 0.32$). However, when the long RSI was used in the first test, the decrease in mean RT from Test 1 ($M = 836.81$, $SD = 138.00$) to Test 2 ($M = 704.89$, $SD = 71.84$) was significant, $F(1, 15) = 30.725$, $MSE = 18125.45$, $p < 0.01$, $\eta_p^2 = 0.67$. Consistent with the recognition task RTs in this experiment, slow performance for the long RSIs (and particularly poor performance in the short-long RSI order condition) contributed to the interaction of test block and RSI order.

The effect of RSI length on correct magnitude judgement RTs was analysed using a 2 (RSI) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. The main effect of RSI was significant, $F(1, 30) = 4.18$, $MSE = 34068.03$, $p = 0.050$, $\eta_p^2 = 0.12$. Consistent with the recognition memory data, the mean RTs were *slower* for long RSI conditions ($M = 845.33$, $SD = 162.52$) than those for short RSI conditions ($M = 798.16$, $SD = 148.25$). Given that the mean accuracy of magnitude judgements was lower in short RSI condition, this result may in part reflect a speed-accuracy trade-off. No other effects or interactions were significant ($F_s < 1.26$).

RT switch costs.

The mean RT switch costs (switch RT – repetition RT) are reported in Table 8.10 as a function of test block, RSI, task and item type. The standard deviations of the RT switch costs, for the magnitude task in particular, were very large. There was a large range in participants' ability to quickly switch tasks. The size of the RT switch cost was generally smaller for the magnitude task than that for the recognition task. Finally, the order of RSI presentation had no effect on the RT switch costs ($F = 0.18$, $p = 0.672$).

Effects of test block.

The RT switch costs reported in Table 8.10 generally decreased in size from the first to second tests, providing evidence of a practice effect. This decrease was particularly salient for the recognition memory task. The effect of test block on RT switch costs was analysed using a 2 (test block) x 4 (item type) x 2 (RSI order) mixed

Table 8.10

Means and Standard Deviations for the RT Switch Costs in Experiment 5A as a Function of RSI, Test Block, Task and Decision.

Task and Decision		Long RSI (2600 ms)				Short RSI (200 ms)			
		1st		2nd		Test 1		Test 2	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Recognition	New	135	115	68	111	155	96	117	117
	Old	152	125	84	100	167	175	95	84
Magnitude	More	71	153	58	125	48	93	20	98
	Less	15	110	44	143	100	120	48	78

Note. RSI = response-to-stimulus interval. Within-participant comparisons should be made across RSIs from Test 1 to Test 2 (i.e., Test 1 Long RSI, Test 2 Short RSI).

ANOVA. The main effect of item type was significant, $F(3, 90) = 10.48$, $MSE = 10618.33$, $p < 0.001$, $\eta_p^2 = 0.26$. Pairwise comparisons using Bonferroni corrections showed that the RT switch cost for ‘new’ recognition decisions ($M = 118.59$, $SD = 90.32$) was no different to that for ‘old’ recognition decisions ($M = 124.47$, $SD = 102.53$). Both recognition RT switch costs were significantly larger than those for ‘less’ magnitude judgements ($M = 53.22$, $SD = 91.12$) and ‘more’ magnitude judgements ($M = 45.99$, $SD = 78.62$). The RT switch costs for both magnitude judgements did not differ significantly from each other. The significant main effect of test block [$F(1,30) = 7.98$, $MSE = 12717.13$, $p = 0.008$, $\eta_p^2 = 0.21$] is illustrated in Figure 8.7. The size of the RT switch cost decreased from Test 1 ($M = 105.48$, $SD = 85.46$) to Test 2 ($M = 66.71$, $SD = 67.43$), producing evidence of a RT practice effect.

Effects of RSI.

As can be seen in Table 8.10, the RT switch costs showed no pattern representative of a preparation effect when comparisons were made within participants. However, within each test block, the size of the RT switch cost was slightly smaller for the long RSI condition than that for the short RSI condition, which is suggestive of a preparation effect. The effect of RSI on RT switch cost was analysed using a 2 (RSI) x 2 (RSI order) x 4 (item type) mixed ANOVA. Although the mean RT switch cost for the

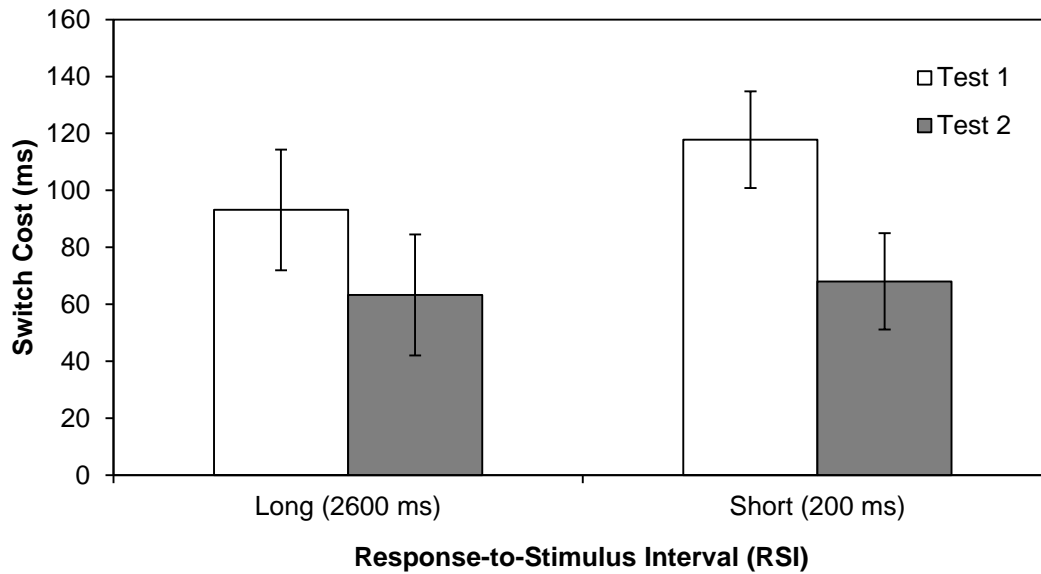


Figure 8.7. Mean RT switch costs for Experiment 5A as a function of RSI and test block. Each test block used a single RSI, and order of RSI presentation was balanced across participants. Error bars represent S.E.

long RSI condition ($M = 78.22$, $SD = 84.84$) was numerically smaller than that for the short RSI condition ($M = 92.91$, $SD = 67.79$), this difference was not significant ($F = 1.09$, $p = 0.306$). No other effects or interactions were significant ($F_s < 2.19$).

Furthermore, when the effect of RSI was analysed for each test block separately, neither the first ($F = 0.66$, $p = 0.421$) nor the second ($F = 0.04$, $p = 0.842$) test block produced significant evidence of a preparation effect. The length of the RSI had no significant impact on the size of the RT switch cost—participants did not use the additional time prior to stimulus onset to prepare.

Cumulative RT distributions.

The cumulative distribution functions for correct repetition and switch trial recognition decisions at the short (200 ms) and long (2600 ms) RSIs are presented in Figure 8.8. The repetition functions were positioned further to the left of the plot and showed steeper gradients than the switch functions. The two repetition functions did not overlap, and contrary to the expected superiority of long RSIs, the short RSI functions were positioned left-most on the plot. Interestingly, and quite counter to the expected preparation effect distributions, the long RSI repetition function overlapped with the short RSI *switch* function at early deciles. de Jong's (2000) failure to engage hypothesis predicts that long RSIs shift the earliest deciles of the switch trial functions to the left, so that they overlap with the repetition trial function. The functions for Experiment 5B

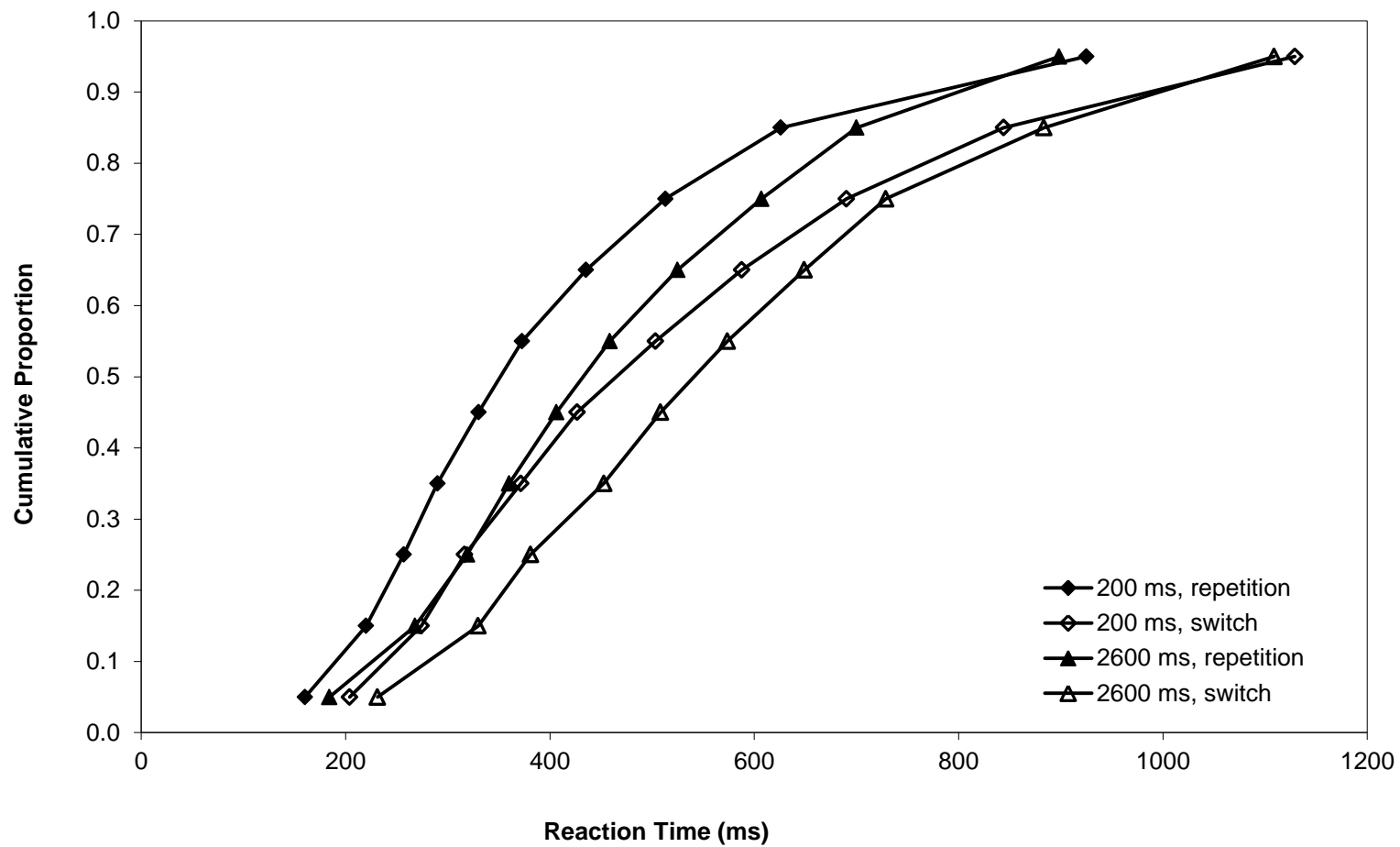


Figure 8.8. Cumulative RT distributions of correct recognition memory trials for Experiment 5A as a function of RSI and task transition.

showed the opposite pattern and thus did not show any evidence of the hypothesised preparation effect.

Experiment 5B

Results and Discussion

Sixty-four participants met the performance requirements and were included in the final data analysis. The use of only correct responses for RT and Switch Cost analyses led to the exclusion of 17% of recognition memory trials and 11% of magnitude judgement trials. The percentage of timeouts for both tasks was low ($< 1\%$).

Accuracy

Recognition task.

Table 8.11 shows the mean corrected accuracy (hit rate – false alarm rate) for recognition decisions as a function of RSI, transition, and test block. The mean corrected accuracy across all conditions ($M = 0.66$, $SD = 0.18$) was lower than previous experiments in this chapter but not overly concerning. There was no difference in overall accuracy between the two RSI order groups ($F = 1.12$, $p = 0.294$).

Effects of transition and test block.

Table 8.11 shows that the corrected accuracies were lower for switch conditions than those for repetition conditions, particularly for the short RSI tests. There was evidence of an accuracy switch effect. Although overall accuracy tended to increase from the first to second tests, the difference between repetition and switch accuracy

Table 8.11.

Means and Standard Deviations for Corrected Recognition Accuracy in Experiment 5B as a Function of RSI, Test Block, and Transition.

Transition	Long RSI (2600 ms)				Short RSI (200 ms)			
	Test 1		Test 2		Test 1		Test 2	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Repetition	0.69	0.19	0.67	0.16	0.65	0.25	0.73	0.20
Switch	0.66	0.26	0.64	0.20	0.59	0.33	0.68	0.26

Note. RSI = response-to-stimulus interval. Within-participant comparisons should be made across RSIs from Test 1 to Test 2 (i.e., Test 1 Long RSI, Test 2 Short RSI).

remained relatively stable. Thus there was little evidence of practice effects for recognition accuracy data. The effects of transition and test block on corrected recognition accuracy were analysed using a 2 (test block) x 2 (transition) x 2 (RSI order) mixed ANOVA. The mean corrected accuracy of repetition trials ($M = 0.69$, $SD = 0.16$) was significantly higher than that for switch trials ($M = 0.64$, $SD = 0.24$), $F(1, 62) = 5.58$, $MSE = 0.02$, $p = 0.021$, $\eta_p^2 = 0.08$. The recognition data showed a typical accuracy switch effect. Neither the main effect of test block ($F = 1.03$, $p = 0.313$) nor the interaction of test block and transition ($F = 0.05$, $p = 0.829$) were significant. Overall corrected recognition accuracy did not increase with experience. More importantly, the size of the accuracy switch effect did not decrease with experience—there was no evidence of a practice effect.

Effects of RSI.

The mean corrected accuracies reported in Table 8.11 showed no consistent difference between the long and short RSI. The difference between repetition and switch accuracy was slightly smaller for the long RSI conditions than that for the short RSI conditions, suggesting that recognition accuracy may show a preparation effect. However, when the effect of RSI on recognition corrected accuracy was analysed using a 2 (RSI) x 2 (transition) x 2 (RSI order) mixed ANOVA, neither the main effect of RSI ($F = 0.01$, $p = 0.930$), nor the interaction of RSI and transition ($F = 0.40$, $p = 0.532$) were significant. The length of the RSI had no significant effect on the overall accuracy of recognition decisions and the size of the accuracy switch effect was not decreased by the long RSI. That is, there was no preparation effect for recognition accuracy. No other effects or interactions were significant ($F = 0.00$).

Magnitude task.

Table 8.12 shows the mean accuracy of magnitude judgements as a function of RSI, transition, decision, and test block. The mean accuracy across conditions was high ($M = 0.89$, $SD = 0.07$) and there was no difference in overall accuracy for the two RSI order groups ($F = 0.43$, $p = 0.516$).

Effects of transition and test block.

The magnitude data reported in Table 8.12 showed a typical accuracy switch effect. The mean accuracy of magnitude judgements was generally higher for repetition conditions than that for switch conditions. Overall accuracy showed a moderate increase from the first to second test blocks, indicating that experience with the magnitude task

Table 8.12.

Means and Standard Deviations for Magnitude Judgement Accuracy in Experiment 5B as a Function of RSI, Test Block, Transition, and Decision.

Transition and Decision		Long RSI (2600 ms)				Short RSI (200 ms)			
		Test 1		Test 2		Test 1		Test 2	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Repetition	More	0.93	0.11	0.98	0.03	0.90	0.11	0.93	0.11
	Less	0.89	0.10	0.91	0.12	0.82	0.18	0.91	0.11
Switch	More	0.87	0.12	0.96	0.06	0.87	0.14	0.91	0.12
	Less	0.88	0.13	0.92	0.10	0.73	0.21	0.85	0.12

Note. RSI = response-to-stimulus interval. Within-participant comparisons should be made across RSIs from Test 1 to Test 2 (i.e., Test 1 Long RSI, Test 2 Short RSI).

improved participants' ability to accurately judge the number of dots displayed on the screen. There was also evidence of practice effects on magnitude accuracy—the difference between repetition and switch trial accuracy was generally smaller in the second test block. The effects of transition and test block were analysed using a 2 (test block) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. Again, the effects of item type are not reported as response keys were not balanced. There was a significant accuracy switch effect. The mean accuracy of repetition trials ($M = 0.91$, $SD = 0.07$) was significantly higher than that for switch trials ($M = 0.88$, $SD = 0.08$), $F(1, 62) = 17.25$, $MSE = 0.01$, $p < 0.001$, $\eta_p^2 = 0.22$. The overall accuracy of magnitude judgements increased from the first ($M = 0.86$, $SD = 0.10$) to second ($M = 0.92$, $SD = 0.06$) test block, [$F(1, 62) = 28.56$, $p < 0.001$, $MSE = 0.02$, $\eta_p^2 = 0.32$], illustrating the general benefit of experience to magnitude judgement accuracy. The main effect of test block was qualified by a significant interaction with RSI order, $F(1, 62) = 22.01$, $MSE = 0.02$, $\eta_p^2 = 0.26$, $p < 0.001$. When the long RSI was used in the first test, the increase in accuracy from Test 1 ($M = 0.89$, $SD = 0.11$) to Test 2 ($M = 0.90$, $SD = 0.10$) was not significant ($F = 0.31$, $p = 0.579$). However, when the short RSI was used in the first test, the increase from Test 1 ($M = 0.83$, $SD = 0.16$) to Test 2 ($M = 0.94$, $SD = 0.07$) was

significant, $F(1, 31) = 38.16$, $MSE = 0.02$, $p < 0.001$, $\eta_p^2 = 0.55$. Highly accurate performance in the Long RSI block likely inflated the effect of test block when the short RSI was presented first, but suppressed the effect of test block when the long RSI was presented first. Critically, the interaction of test block and transition failed to reach statistical significance ($F = 2.71$, $p = 0.105$). The decrease in accuracy switch effect was not significant—there was no evidence of a practice effect for magnitude task accuracy.

Effects of RSI.

Table 8.12 shows that the mean magnitude task accuracies were higher for the long RSI conditions than those for the short RSI conditions. Furthermore, the accuracy switch effects were generally smaller for the long RSI conditions than those for the short RSI conditions. Thus there was evidence of a preparation effect on magnitude accuracy. The effect of RSI on magnitude task accuracy was analysed using a 2 (RSI) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. The mean accuracy of magnitude judgements was significantly higher for the long RSI condition ($M = 0.92$, $SD = 0.06$) than that for the short RSI condition ($M = 0.87$, $SD = 0.10$), $F(1, 62) = 22.01$, $MSE = 0.02$, $p < 0.001$, $\eta_p^2 = 0.26$. Critically, the interaction of RSI and transition was significant, $F(1, 62) = 4.78$, $MSE = 0.01$, $p = 0.033$, $\eta_p^2 = 0.07$. As can be seen in Figure 8.9, the main effect of transition was significant at short RSIs [$F(1, 62) = 4.78$, $MSE = 0.01$, $p = 0.048$, $\eta_p^2 = 0.06$] and at long RSIs [$F(1, 62) = 15.91$, $MSE =$

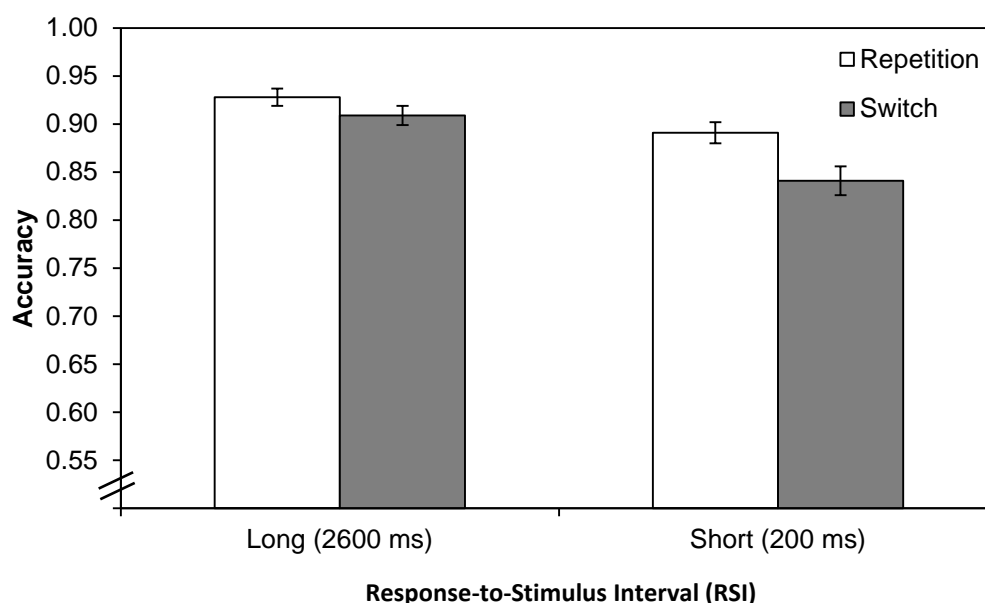


Figure 8.9. Mean accuracy of magnitude judgements in Experiment 5B as a function of RSI and transition. Error bars represent S.E.

0.01, $p < 0.001$, $\eta_p^2 = 0.20$]. However, the size of the accuracy switch effect was considerably larger for the short RSI condition than that for the long RSI condition. There was a significant preparation effect on magnitude judgement accuracy. No other effects or interactions were significant ($F = 0.00$).

Reaction Times

The RT data are reported in the same format as earlier experiments in this chapter: (1) analyses of correct recognition RTs, (2) analyses of correct magnitude RTs, and (3) analyses of RT switch costs (switch RTs – repetition RTs). A comprehensive table of RTs for both tasks can be found in Appendix D. The mean overall RT for the recognition task ($M = 479.50$, $SD = 119.83$) was similar Experiment 5A and, when the 400-ms RSD is added, was slightly slower than that for the magnitude task ($M = 802.76$, $SD = 102.14$). The main effect of RSI order was not significant for the recognition ($F = 1.93$, $p = 0.170$) or magnitude ($F = 0.01$, $p = 0.932$) tasks.

Recognition task.

The effects of transition and test block on correct recognition RTs were analysed using a 2 (test block) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. An RT switch effect was confirmed by the significant main effect of transition, $F(1, 62) = 120.95$, $MSE = 14030.62$, $p < 0.001$, $\eta_p^2 = 0.66$. The mean correct recognition RT for repetition trials ($M = 422.13$, $SD = 96.71$) was significantly faster than that for switch trials ($M = 537.27$, $SD = 151.24$). Correct responses were longer for ‘new’ words ($M = 504.86$, $SD = 131.92$) than those for ‘old’ words ($M = 454.54$, $SD = 120.77$), $F(1, 62) = 24.80$, $MSE = 13069.59$, $p < 0.001$, $\eta_p^2 = 0.29$. Experience with the tasks allowed participants to make correct recognition decisions more quickly. Compared to performance on Test 1 ($M = 503.36$, $SD = 153.00$), correct recognition RTs were faster in Test 2 ($M = 456.03$, $SD = 106.42$), $F(1, 62) = 11.92$, $MSE = 24063.66$, $p = 0.001$, $\eta_p^2 = 0.16$. Consistent with Experiment 5A, the interaction of test block and RSI order was also significant, $F(1, 62) = 27.45$, $MSE = 24063.66$, $p < 0.001$, $\eta_p^2 = 0.31$. When the long RSI was used in the first test, the decrease in mean correct RT from Test 1 ($M = 518.47$, $SD = 213.82$) to Test 2 ($M = 399.31$, $SD = 149.49$) was significant, $F(1, 31) = 39.56$, $MSE = 22970.99$, $p < 0.001$, $\eta_p^2 = 0.56$. By contrast, when the short RSI was used in the first test, the mean correct RTs for Test 1 ($M = 488.26$, $SD = 218.90$) and Test 2 ($M = 512.75$, $SD = 151.52$) were not significantly different ($F = 1.53$, $p = 0.226$). Slow RTs for the long RSI conditions (which were not the product of a speed-accuracy trade-off)

inflated the main effect of test block when Test 1 used a long RSI and suppressed the main effect of test block when Test 1 used the short RSI.

The effect of RSI length on correct recognition RTs was analysed using a 2 (RSI) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. Contrary to what would be expected in the preparation effect literature but consistent with Experiment 5A, the mean correct RT was slower for the long RSI condition ($M = 515.61$, $SD = 131.02$) than that for the short RSI condition ($M = 443.78$, $SD = 132.54$), $F(1, 62) = 27.45$, $MSE = 24063.66$, $p < 0.001$, $\eta_p^2 = 0.31$. This effect of RSI was qualified by an interaction with item type, $F(1, 62) = 19.91$, $MSE = 5280.36$, $p < 0.001$, $\eta_p^2 = 0.24$. At long RSIs, the mean correct RT for ‘new’ decisions ($M = 526.44$, $SD = 141.40$) was no different to that for ‘old’ decisions ($M = 504.78$, $SD = 138.18$), $F = 3.16$, $p = 0.080$. By contrast, at short RSIs the mean correct RT for ‘new’ decisions ($M = 483.27$, $SD = 153.43$) was significantly slower than that for ‘old’ decisions ($M = 404.29$, $SD = 126.55$), $F(1, 62) = 45.13$, $MSE = 8846.80$, $p < 0.001$, $\eta_p^2 = 0.42$. Long RSIs were particularly damaging when a recognition trial required an ‘old’ response, to the point that they eliminated the RT advantage of a self-terminating search of memory that was observed at short RSIs. No other effects or interactions were significant ($F_s < 1.57$).

Magnitude task.

The effects of transition and test block on correct magnitude RTs were analysed using a 2 (test block) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. Again, the main effects and interactions of item type are not reported as response key allocation was not balanced. The RT switch effect was confirmed by the significant main effect of transition, $F(1, 62) = 25.77$, $MSE = 14751.80$, $p < 0.001$, $\eta_p^2 = 0.29$. The mean correct RT for the magnitude task was significantly faster for repetition trials ($M = 775.51$, $SD = 101.29$) than that for switch trials ($M = 830.00$, $SD = 119.95$). The main effect of test block was also significant, $F(1, 62) = 22.63$, $MSE = 38635.53$, $p < 0.001$, $\eta_p^2 = 0.27$. Participants’ mean correct RTs decreased from Test 1 ($M = 884.08$, $SD = 134.75$) to Test 2 ($M = 761.43$, $SD = 111.20$), suggesting that experience with the magnitude task allowed participants to make correct decisions more quickly.

The effect RSI length on correct RTs for the magnitude task was analysed using a 2 (RSI) x 2 (transition) x 2 (item type) x 2 (RSI order) mixed ANOVA. Contrary to the recognition RT data, the main effect of RSI was not significant ($F = 0.93$, $p = 0.339$). The length of the RSI had no impact on the speed of participants’ correct

Table 8.13.

Means and Standard Deviations for the RT Switch Costs in Experiment 5B as a Function of RSI, Test Block, Task and Decision.

Task and Decision		Long RSI (2600 ms)				Short RSI (200 ms)			
		1st		2nd		Test 1		Test 2	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Recognition	New	119	141	65	75	147	182	122	111
	Old	130	126	138	116	103	121	98	90
Magnitude	More	135	162	46	85	47	85	-4	117
	Less	50	127	-2	104	93	113	70	126

Note. RSI = response-to-stimulus interval. Within-participant comparisons should be made across RSIs from Test 1 to Test 2 (i.e., Test 1 Long RSI, Test 2 Short RSI).

magnitude judgements. No other effects or interactions were significant ($F_s < 0.60$).

RT switch costs.

RT switch costs (switch RT – repetition RT) are reported in Table 8.13 as a function of test block, RSI, task and item type. As with the earlier experiments in this chapter, there was large variation in the size of the RT switch costs. Interestingly, two of the magnitude task RT switch costs were less than zero (i.e., no RT switch cost), suggesting that the cost of switching tasks could be eliminated. The large standard deviations indicated that there was a fair amount of individual difference in people's ability to quickly switch tasks. The mean RT switch costs were generally smaller for the magnitude task than those for the recognition task. The order of RSI presentation had no effect on the RT switch costs ($F = 0.33, p = 0.567$).

Effects of test block.

The mean RT switch costs reported in Table 8.14 generally decreased from the first to the second test block, illustrating a practice effect. However, this decrease was inconsistent across different task and item types. The effect of test block on RT switch cost was examined using a 2 (test block) x 4 (item type) x 2 (RSI order) mixed ANOVA. A Greenhouse-Geisser correction was used for the item type main effect as the assumption of sphericity was violated. The main effect of item type was significant,

$F(2.61, 161.73) = 13.53$, $MSE = 12578.99$, $p < 0.001$, $\eta_p^2 = 0.18$. Pairwise comparisons using Bonferroni corrections showed that the mean RT switch cost for ‘new’ decisions ($M = 113.06$, $SD = 103.74$) and ‘old’ decisions ($M = 117.23$, $SD = 92.49$) did not differ from each other. However, both recognition task RT switch costs were significantly larger than those for ‘more’ ($M = 56.07$, $SD = 90.51$) and ‘less’ ($M = 52.92$, $SD = 100.78$) magnitude RT switch costs (which did not differ from each other). The cost of switching to the recognition task was larger than that of switching to the magnitude task.

A practice effect was confirmed by a significant main effect of test block, $F(1, 62) = 12.64$, $MSE = 13357.18$, $p = 0.001$, $\eta_p^2 = 0.17$. Collapsed across tasks, the size of the RT switch cost decreased from Test 1 ($M = 102.98$, $SD = 95.42$) to Test 2 ($M = 66.66$, $SD = 65.80$). However, this main effect of test block was qualified by an interaction with item type [$F(3, 186) = 2.94$, $MSE = 9250.46$, $p = 0.038$, $\eta_p^2 = 0.05$], illustrated in Figure 8.10. Paired sample t -tests showed that the interaction was driven by a non-significant effect of test block for ‘old’ ($t = -0.07$, $p = 0.943$) and ‘new’ ($t = 1.83$, $p = 0.071$) recognition decisions. By contrast, the decrease in RT switch cost from the first to the second test block was significant for both ‘less’ [$t(63) = 2.23$, $MSE = 16.76$, $p = 0.029$] and ‘more’ [$t(63) = 3.45$, $MSE = 20.25$, $p = 0.001$] magnitude judgements. That is, while the magnitude task showed evidence of practice effects, the recognition task did not.

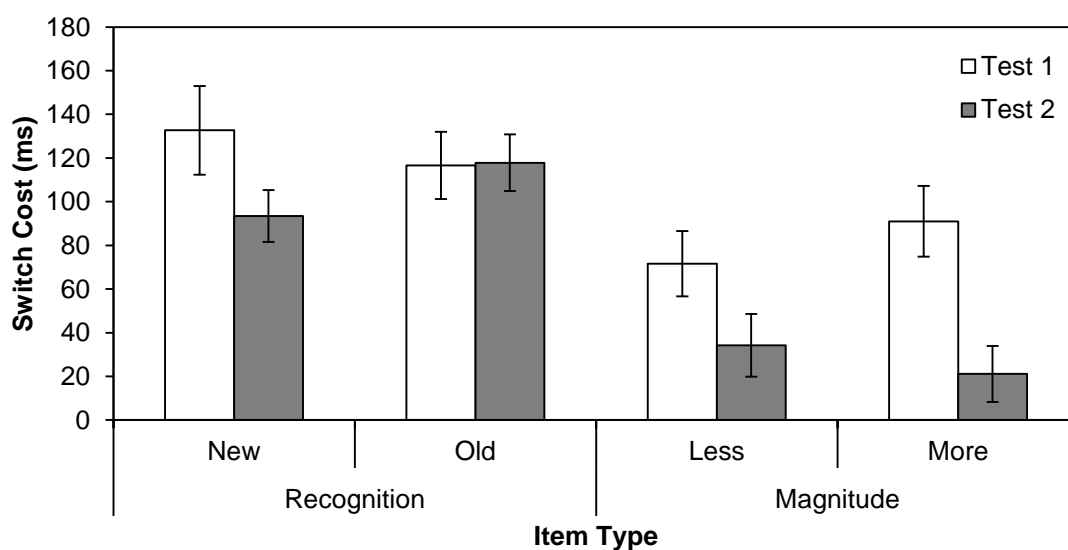


Figure 8.10. Mean RT switch costs for Experiment 5B as a function of test block and item type. Error bars represent S.E.

Effects of RSI.

The size of the RT switch costs reported in Table 8.14 (p. 138) did not show a pattern across RSIs representative of a preparation effect, despite the fact that participants were exposed to both RSIs during each practice test. The effect of RSI on RT switch costs was analysed using a 2 (RSI) x 2 (RSI order) x 4 (item type) mixed ANOVA. While the main effect of RSI was not significant ($F = 0.00$, $p = 0.951$), there was a significant interaction of RSI and item type [$F(3, 186) = 12.71$, $MSE = 9274.90$, $p < 0.001$, $\eta_p^2 = 0.17$], illustrated in Figure 8.11. Paired sample t -tests were used to follow up this interaction. Effects of RSI consistent with a preparation effect (long RSI RT switch cost < short RSI RT switch cost) were found for ‘new’ RT switch costs [$t(63) = -1.99$, $MSE = 21.30$, $p = 0.050$] and ‘less’ RT switch costs [$t(63) = -3.63$, $MSE = 15.83$, $p = 0.001$]. By contrast, RSI effects in the opposite direction (long RSI RT switch cost > short RSI RT switch cost) were found for ‘old’ RT switch costs [$t(63) = 2.00$, $MSE = 16.53$, $p = 0.050$] and ‘more’ RT switch costs [$t(63) = 3.42$, $MSE = 20.28$, $p = 0.001$]. While it is worth noting that the interaction of RSI and item type was consistent with that observed in Experiment 4B for the magnitude task, this difference in RSI effect should be treated cautiously as the response keys for the magnitude task were not balanced. However, even considering only the recognition data, it is difficult to produce a theoretically driven explanation for the opposing direction of RSI effects for ‘old’ and ‘new’ recognition responses. Thus it is difficult to consider this result as strong

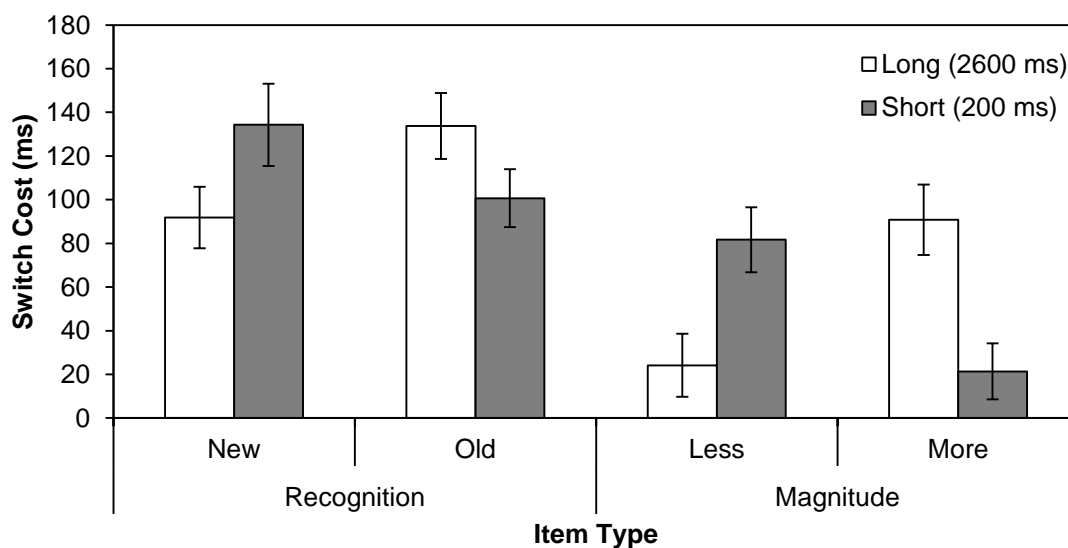


Figure 8.11. Mean RT switch costs for Experiment 5B as a function of RSI and item type. Error bars represent S.E.

evidence of a preparation effect. No other effects or interactions were significant ($F_s < 0.37$).

Cumulative RT distributions.

On account of the significant interaction of RSI and item type for recognition RT switch costs, separate cumulative RT distributions were constructed for each recognition item type. Figures 8.12 and 8.13 show the cumulative distribution functions for correct ‘old’ and ‘new’ decisions as a function of RSI (200 ms vs. 2600 ms) and transition. For ‘new’ words (Figure 8.12), task switch functions were positioned to the right of task repetition functions. For the earliest deciles, short and long RSI functions overlapped for each transition but began to separate at around the third decile. Consistent with Experiment 5A and contrary to what would be expected for preparation effect functions, the short RSI repetition and switch functions were positioned to the *left* of their respective long RSI functions. The switch and repetition functions were positioned close together at the first deciles, but this closeness was equally so for both RSIs. Thus while the main effect of RSI on ‘new’ RT switch costs was consistent with a preparation effect, the cumulative RT functions did not show the typical preparation effect characteristics.

For ‘old’ words (Figure 8.13), the cumulative distribution functions were more spread across the plot. Unlike distributions usually seen for preparation effects, the two repetition functions did not overlap. Interestingly, the long RSI switch condition was more similar to the short RSI switch condition than the short RSI repetition condition. There was little overlap at the first decile for repetition and switch functions of the same RSI. The cumulative RT functions produced for old words in Experiment 5B showed none of the characteristics typical of preparation effect distributions, and if anything suggested that long RSIs were particularly damaging to performance when an ‘old’ recognition decision was required.

Conclusions for Experiments 5A and 5B

In Experiments 5A and 5B, the length of the RSI was manipulated between test blocks (200 ms vs 2600 ms) to test for evidence of preparation effects. Additionally, switch effects were compared between test blocks (Test 1, Test 2) to test for evidence of practice effects. Experiment 5B differed from Experiment 5A in that the practice tests completed prior to each actual test used *both* RSIs to ensure recent exposure to both RSIs. Regardless of this difference, the general findings of both experiments were

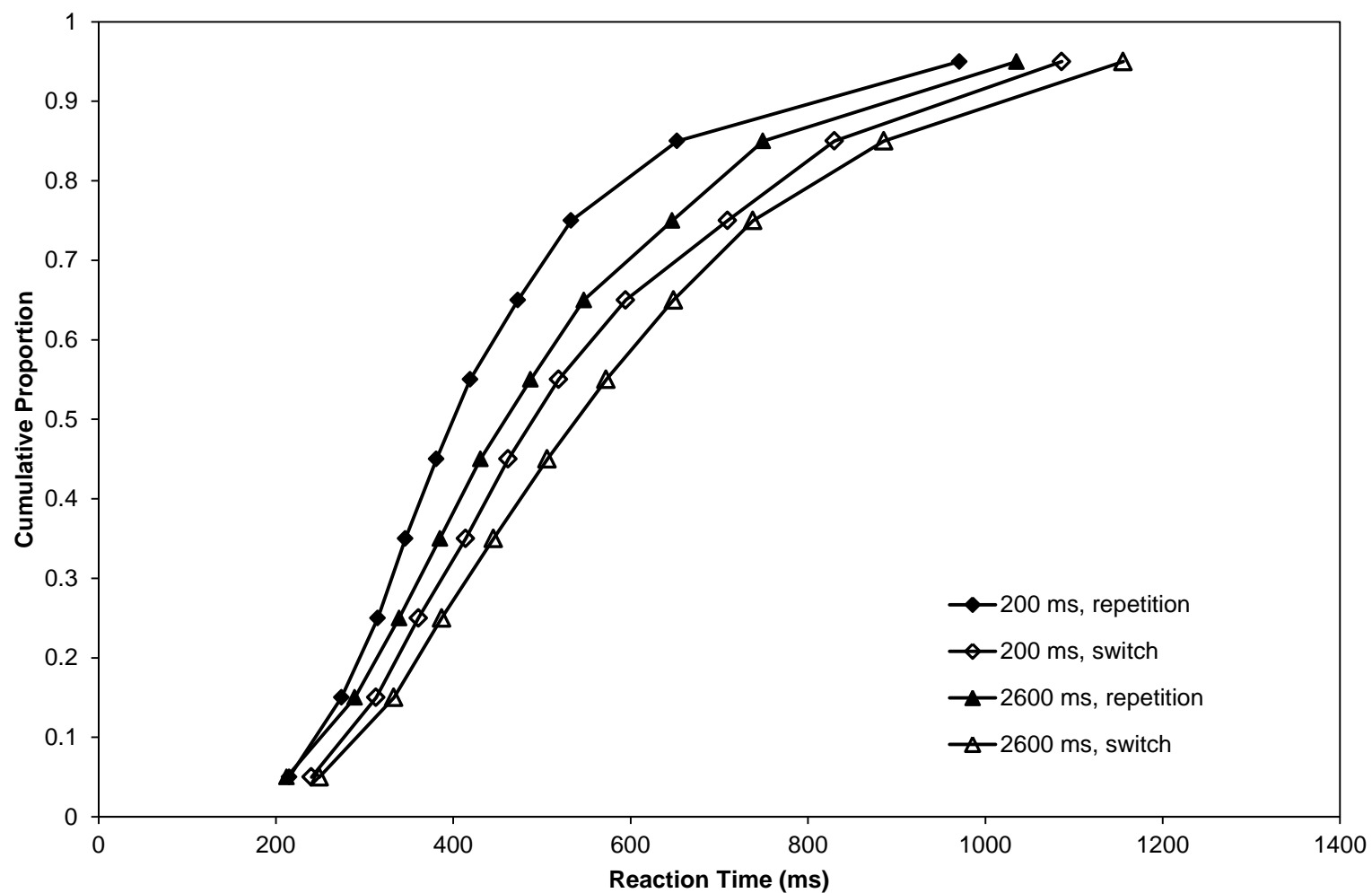


Figure 8.12. Cumulative RT distributions of correct 'new' recognition memory trials for Experiment 5B as a function of RSI and task transition.

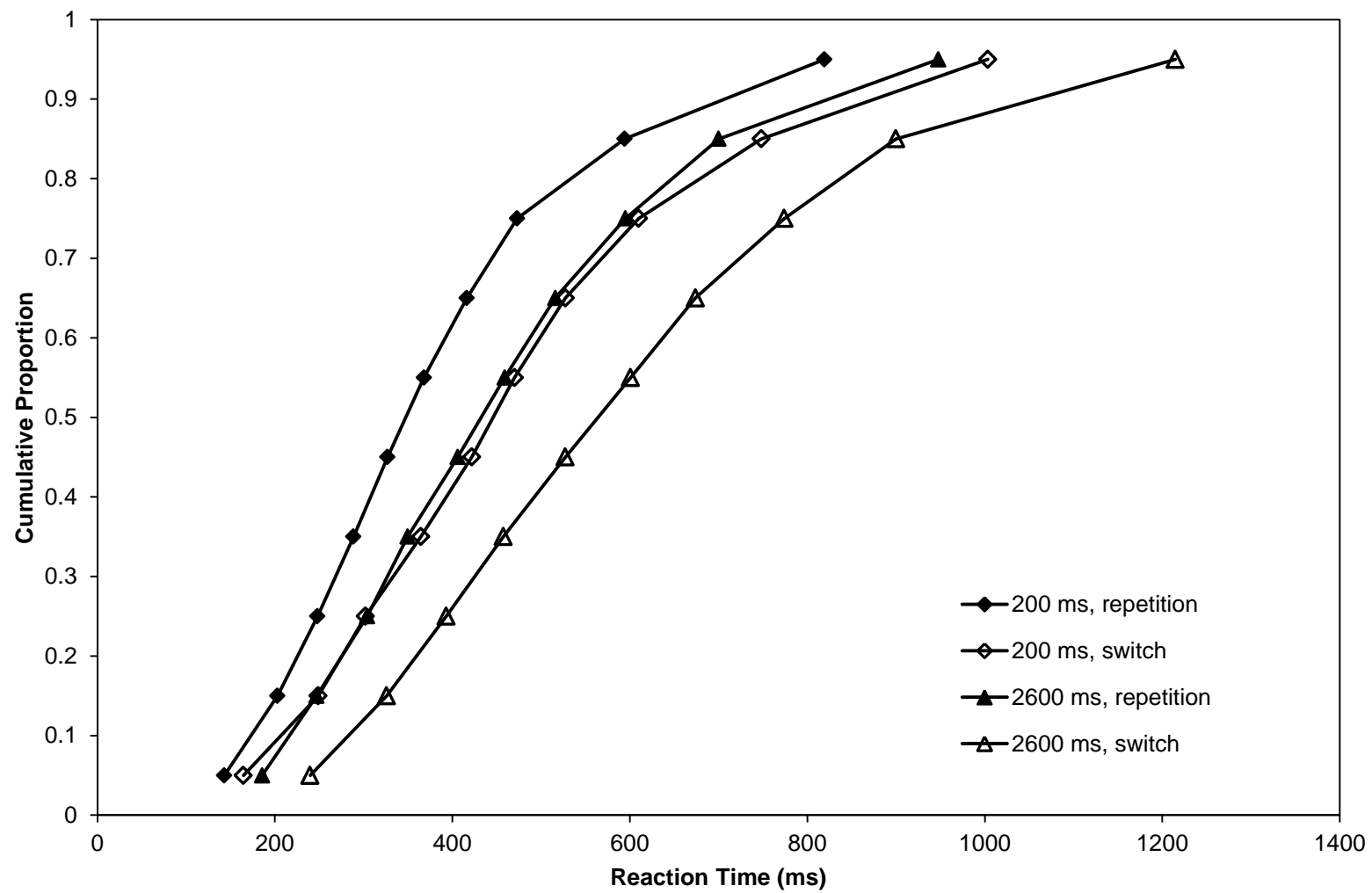


Figure 8.13. Cumulative RT distributions of correct 'old' recognition memory trials for Experiment 5B as a function of RSI and task transition.

quite similar and are therefore discussed together. Consistent with Experiments 4A and 4B, significant RT switch effects were reported for both tasks. Participants' correct recognition and magnitude responses were slower and when they had switched tasks relative to when they had repeated tasks. Both experiments reported significant accuracy switch effects for the magnitude task. However, while both experiments reported numerically lower recognition accuracy for switch trials than that for repetition trials, this difference was only significant for Experiment 5B. The current results are fairly consistent with previous literature, given that accuracy switch effects are less robust than RT switch effects.

Both experiments showed a significant main effect of test block (i.e., a practice effect) on RT switch costs. However, the interaction of item type and test block in Experiment 5B showed that this practice effect did not reach statistical significance for the recognition task. In balance, the results of Experiments 5A and 5B supported the hypothesised practice effect but showed that the presence (or extent) of the effect may depend on the task or other experimental factors.

While Experiment 5A showed no evidence of preparation effects, Experiment 5B produced a peculiar interaction of RSI and item type. The interaction of RSI and item type observed for recognition task RT switch costs is problematic. The cumulative RT distributions for 'new' recognition decisions in Experiment 5B showed none of the key characteristics of a preparation effect distribution and visually illustrated the rather counterintuitive slowing of recognition decisions (particularly 'old') decisions at the long RSI. There is no theoretically-based reason to expect preparation effects for one task decision and not another, particularly as participants did not know what type of recognition decision would occur on any given trial. As such, it is difficult to interpret these results as support for the hypothesised RT preparation effect. This curious interaction of RSI and item type will be discussed in greater detail in the discussion section of this chapter.

Interestingly, Experiment 5B also reported an accuracy preparation effect—the size of the accuracy switch effect for the magnitude task was larger in the short RSI condition than that in the long RSI condition. However, given the one-off nature of this finding and that the RT switch cost analyses provided little corroborating evidence of preparation effects, it is likely unwise to assign too much weight to this finding. In balance, neither Experiment 5A nor Experiment 5B produced convincing evidence of preparation effects. Using two RSIs that were noticeably different from each other

(Experiment 5A and 5B) and providing recent exposure to both RSIs (Experiment 5B) did not change task switching and preparation performance.

Vandierendonck (personal communication, November 2009) suggested that Experiments 5A and 5B may have failed to show preparation effects because of the length of the long RSI. That is, the 2600-ms long RSI may have been so long that participants did not use the RSI to prepare for the upcoming task switch, but rather allowed the time to pass passively. The mean RTs for the recognition task were slower at the long RSI than at the short RSI, which does provide some support for Vandierendonck's suggestion. However, given that Experiments 4A and 4B—which used less extreme long RSIs—also failed to observe preparation effects, I would argue that the length of the RSI is not the main issue driving the null effect. Indeed, the work of Meiran et al. (2001) used a CSI of 3032 ms and reported a smaller RT switch effect for this extremely long CSI than those for shorter CSIs. Clearly, there must be other reasons for the lack of significant preparation effect.

Discussion

With regard to accuracy switch effects, the overall difference between the accuracy of repetition and switch transitions was statistically significant for both tasks in Experiments 4A and 5B, and significant for the magnitude task in Experiment 5A. With the exception of Experiment 4A, accuracy switch effects were smaller than reported in Experiments 2 and 3. These smaller costs, combined with the large variance in performance, likely contributed to the non-significant switch effects in Experiment 4B and 5A. However, it is worth noting that, as task switching effects have a predicted direction (switch accuracy < repetition accuracy), a one-tailed test of significance could be applied (i.e., $p < 0.10$). Using this alpha level, all of the accuracy switch effects reported in Chapter 7 were statistically significant. Furthermore, a Wilcoxon signed ranks test conducted on overall accuracy for each task and experiment produced a significant difference between repetition and switch transitions ($Z = -3.35$, $p = 0.001$), which is unsurprising given 15 of the 16 cells showed numerically higher accuracy for repetition trials than that for switch trials. The experiments reported in this chapter consistently produced evidence that the accuracy of participants' recognition and magnitude responses was impaired by a recent change in tasks.

Significant RT switch effects were consistently reported across both tasks. Participants were slower to make (correct) recognition and magnitude responses when

they had switched tasks compared to when they had repeated tasks. When expressed as RT switch costs (switch RTs – repetition RTs), the size of the RT cost varied across tasks and conditions. The standard deviations for RT switch costs were large, particularly for the magnitude task. Rather than reflecting unstable data, it is likely this variance was the product of large individual differences in performance. Indeed, some participants showed very small RT switch costs or even switch advantages. Clearly some people were very able task switchers (by natural ability or the strategies they utilised) while others had difficulty quickly switching tasks.

Experiments 5A and 5B showed equivalent RT switch costs for both decisions within the recognition and magnitude tasks but significantly larger RT switch costs for recognition decisions than for magnitude judgements. Experiments 4A and 4B showed similar trends, however only the differences between ‘old’ and ‘less’ decisions and ‘old’ and ‘more’ decisions were statistically significant for Experiments 4A and 4B, respectively. These results provided convergent evidence that the size of the RT switch cost to performance was larger for the recognition task than that for the magnitude task. This is consistent with Mayr and Kliegl’s (2000) long term memory retrieval demand hypothesis. Switching to a new task requires that the relevant task set be selected from long term memory and installed in working memory (c.f., Rubinstein et al. 2001). When the task proper also requires retrieval from long term memory—as is the case for the recognition task but not the magnitude task—responses are slowed even further on switch trials. Retrieving a task set and searching episodic memory for a response to a task stimulus must occur sequentially. Subsequent processes are halted until the earlier processes are completed—there is a bottleneck (see Pashler, 1994). When a long RSI is used to retrieve a task set in advance (i.e., a participant has ‘prepared’) there should be no slowing at the bottleneck and thus no difference in RT switch costs for tasks with different long term memory retrieval demands. However, differences in RT switch cost for the recognition and magnitude tasks *were* observed at both short and long RSIs. This persistence in task differences suggests that participants did not use the long RTIs to prepare for tasks switches, a topic which will be revisited later in this discussion.

Practice Effects

With regard to practice effects—the decrease in the size of the RT switch cost from the first to the second test block—the main effect of test block (collapsed across tasks) was significant in all experiments except Experiment 4A. Nevertheless, the mean

RT switch costs for the first test block were numerically larger than those for the second test block in all but one condition ('more' decisions at Long RSIs) of Experiment 4A. In balance, the evidence for practice effects in the current experiments was fairly convincing. Consistent with previous research (Jersild, 1927; Koch, 2005; Kramer et al., 1999; Kray & Lindenberger, 2000; Meiran, 1996; Minear & Shah, 2008; Monsell et al. 2003; Rogers & Monsell, 1995), task switching performance generally showed a reliable improvement from the first to the second test block. Although the size of the RT switch costs reported in this chapter are somewhat smaller than are generally reported in experiments utilising a 2-trial alternating runs procedure, the benefits of practice to task switching performance were similar. Across all four experiments, the RT switch cost decreased between 20 – 38% from the first to second test. This percentage decrease in switch cost was roughly equivalent to those reported by Koch (2005) and Meiran (1996) across a similar number of trials. Interestingly, they were also similar to the decreases reported by Rogers and Monsell (1995) across two days.

As detailed in the introduction, the benefit of practice to task switching performance can be conceptualised as the product of faster task set reconfiguration (Meiran, 1996). Through repeated exposure (i.e., practice) the association between a task cue and a task set is strengthened, and consequentially the task set can be retrieved from long term memory more quickly. In the case of the current experiments, this task cue could be the product of internal monitoring of the predictable pattern of tasks (e.g., Baddeley et al., 2001; Kray & Lindenberger, 2000) or, as each task used different stimulus sets, the stimulus itself. While Meiran (1996) reported that practice effects were observed at short CSIs but not at long CSIs, the current research observed no such difference between short and long RSIs. Participants showed little evidence of using long RSIs to prepare and as a consequence the practice effect remained intact. These results suggest that the practice effects reported in this chapter likely represented the strengthening of associations between stimuli (or stimulus sets) and task sets over the course of the experiment, making task switching processes more efficient.

Preparation Effects

As has been noted, the persistence of task differences for RT switch costs and practice effects at long RSIs suggests that participants did not prepare for task switches in advance of stimulus presentation. Direct comparisons of task switching performance at short and long RSIs produced similar conclusions across the experiments reported in

this chapter. The most basic test of the RT preparation effect—a main effect of RSI on mean RT switch cost—was not significant for any of the experiments in the chapter.¹³ Given that exposure to both RSIs may be considered essential to producing preparation effects (Altmann, 2004), a stronger test for preparation effects in Experiments 4A – 5A was the between-subjects comparison of RT switch cost at Test 2. However, this test also failed to produce a significant effect of RSI. Even when each practice test utilised both RSIs (Experiment 5B) the main effect of RSI remained non-significant. However, there was a significant interaction of RSI and item type for Experiment 5B, which posed a challenge to interpret within task switching theory. In Experiment 5B, the mean RT switch costs for ‘new’ recognition decisions and ‘less’ magnitude judgements showed a significant effect of RSI that was consistent with a preparation effect (long RSI switch cost < short RSI switch cost). By contrast, mean RT switch costs for ‘old’ recognition decisions and ‘more’ magnitude judgements showed a main effect of RSI that was in the *opposite* direction (long RSI switch cost > short RSI switch cost). It is interesting to note that this interaction of RSI and item type occurred for the magnitude task in Experiment 4B. The remaining experiments showed the same pattern of performance, but the interactions did not reach statistical significance. However, the consistency of this curious pattern across experiments with different RSIs suggests that something systematic was driving the interaction.

Although the pattern of RT switch costs was consistent with a preparation effect for ‘new’ and ‘less’ decisions, there are reasons to discount this interaction of RSI and item type as evidence for preparation effects. First, as participants had no way of knowing the status of an upcoming recognition (‘old’ or ‘new’) or magnitude (‘less’ or ‘more’) task, proposing that the interaction was driven by different amounts of preparation prior to the presentation of a word is problematic. Second, as task sets are considered to contain general information and rules about how to process stimuli and respond in a task-appropriate manner (Monsell, 2003; Rubinstein et al., 2001), there is no reason to anticipate that any preparation of a task set (be that through activation, encoding, or reconfiguration) would be biased toward a particular response.

Finally, there was no converging evidence from other sources of information about preparation effects. In Experiment 5B (where the interaction of item type and RSI was significant) the RT switch cost for ‘new’ recognition decisions remained larger than

¹³ When RT switch costs were calculated as a proportion of the repetition RT, the data provided no evidence consistent with preparation effects.

that for ‘less’ magnitude judgements at the long RSI. Likewise, the cumulative RT functions produced for the recognition task were not consistent with those observed when preparation effects are present. In Experiment 5B, despite the main effect of RSI for ‘new’ recognition decisions that was consistent with a preparation effect, the functions for ‘new’ decisions lacked several of the key characteristics of preparation functions. Although switch and repetition functions were close to overlapping at the first decile, this outcome was found for short *and* long RSIs. Also, the two repetition functions deviated at the fourth decile, at which point the short RSI functions were to the left (instead of to the right) of long RSI functions. Thus the cumulative distribution functions do not support a preparation effect interpretation of the RSI effect for ‘new’ recognition decisions. In balance, evidence from the experiments reported in this chapter strongly suggest that participants did not prepare for upcoming task switches any more with long RSIs than they did with short RSIs. Thus, the hypothesis that a long RSI would produce smaller switch effects than those for a short RSI was not supported.

If the interaction of RSI and item type does not reflect preparation for some items and not others, what does it reflect? The interaction does not reflect bias for a particular response-to-key allocation in the recognition task. When key mapping is added as a factor to the mixed ANOVA for RT switch cost in Experiment 5B (the strongest example of the RSI by item type interaction), it does not interact with item type and RSI (Experiment 5B: $F = 0.73$, $p = 0.396$). Likewise, comparisons of hit and false alarm rates (see Appendix B) showed no consistent pattern that would explain the interaction of RSI and item type for RT switch costs.¹⁴ Experiments that showed (or trended toward showing) the interaction of item type and RSI for RT switch costs (Experiments 4B, 5A, and 5B) also reported slower RTs at the long RSI than at the short RSI. By contrast, the experiment that showed the least indication of an item type and RSI interaction (Experiment 4A) reported faster RTs at the long RSI than at the short RSI. It is interesting to note that both the ‘new’ decisions and the ‘less’ decisions were generally the slower of the responses for each task. That is, when the RSI was long and the response was relatively slow, a main effect of RSI consistent with a preparation effect was observed. However, when the RSI was long and the response was relatively fast (e.g., ‘old’ and ‘more’ decisions), the main effect of RSI showed the reverse

¹⁴ Recognition performance in Experiment 5B does show a bias toward responding ‘old’ at short RSIs. However, it is unclear how this would translate to the observed interaction, and none of the other experiments show any hint of bias.

pattern. Determining if (and how) the time needed to respond plays a role in producing patterns of performance consistent with preparation effects in Experiment 4B and 5B will require additional research.

The experiments in this chapter were designed to examine preparation effects in a more naturalistic manner than is typical in the task switching literature. That is, they shifted away from using simple tasks that shared stimulus sets (with sizeable stimulus-response overlap) towards using a more complex task (i.e., the recognition memory task) that used a different stimulus set than the secondary magnitude task. Likewise, instead of investigating how participants can be encouraged to prepare in advance of a task switch, the critical question was whether participants naturally and independently prepared in advance of a task switch when presented with conditions that would support this strategy. External task cues were not provided, both to maintain the naturalistic goals of the experiments and to encourage participants to monitor the task sequence internally. There is evidence in the literature that these types of changes can be made without eliminating preparation effects. For example, Kray and Lindenberger (2000) observed preparation effects without the use of task cues or explicit instructions to use RSIs to prepare. Likewise, Rogers and Monsell (1995) observed preparation effects when using stimuli that were constructed to contain elements relevant to only one task (e.g., ‘A#’ was used for the letter task and ‘&9’ was used for the parity task), in essence producing different stimulus sets for each task. However, these changes have never been made in concert, which could explain why evidence of advanced preparation was not found in Experiments 4A – 5B.

As one of the key interests of this thesis was task switching effects on recognition memory, it was important to work within the limitations of this task. The recognition task required that participants complete separate study phases prior to each test phase, thus limiting the number of study-test blocks that could be administered whilst maintaining an appropriate level of task difficulty and minimising fatigue. The length of each experiment (and number of trials completed) was substantially smaller than is typical for task switching and preparation effect experiments. Most task switching research also excludes the first few blocks from analysis. Thus the failure to observe preparation effects could be interpreted as the consequence of analysing relatively few trials—the very trials others might typically exclude. Indeed, the large variability in RT switch costs, particularly at Test 1, does suggest that excluding initial blocks and/or trials may have aided in the detection of more subtle effects. However,

Meiran's (1996) examination of preparation and practice effects excluded only 50 trials from the beginning of the experiment. Even for the first test block following these excluded trials a sizable preparation effect was reported. While Meiran's (1996) tasks were relatively simple, preparation effects clearly occurred without much practice with the tasks. Thus, it is unlikely that the absence of preparation effects in the current experiments can be wholly explained by the relatively small number of trials and limited practice.

With regard to Rogers and Monsell's (1995) preparation effect with tasks using different stimulus sets, it is important to note that, compared to experiments where tasks used the same stimulus set (i.e., when 'A9' was used for both the letter and parity tasks), the sizes of the RT switch costs were smaller and the subsequent decline with increasing RSI was comparatively subtle. Also important is that Rogers and Monsell very explicitly instructed their participants to prepare for upcoming task switches "as quickly as possible following the previous response" (p. 217) and assessed performance over many blocks spread across two days. One could argue that the parameters of the experiment produced a relatively subtle effect in response to considerable effort to force advanced preparation. With this in mind, perhaps the method used for the current research lacked the necessary power and control over how participants switched tasks (via specific instructions to prepare) to produce preparation effects.

For the current experiments, the hypothesis that preparation effects would be observed rested on the assumption that, as was the case in Kray and Lindenberger (2000), participants would use the predictability of task sequences to anticipate the identity of the upcoming task. That is, participants would generate an internal task cue during the RSI that could be used to prepare. However, Koch (2003) argued that internally generated cues are functionally different from and comparatively weaker to external cues with regard to the processes they can initiate. It may be that in the case of the current experiments these internal cues were either ineffective or, considering the cognitive effort required to monitor task sequences, not even produced. Instead, participants may have relied on the stimulus (a perfect predictor of the task to be performed) to prompt task switches. Using a task set reconfiguration model of task switching, this type of strategy would indicate that on switch trials both the general task processes *and* task set switching (reconfiguration) processes were engaged after stimulus presentation (Meiran, 2000a; Rubinstein et al., 2001). Participants did not take advantage of long RSIs to initiate reconfiguration processes in advance—the RSI was

wasted. This explanation sits well with Altmann's (2004b) claim that *if* the preparation effect was the product of controlled reconfiguration, this process was inherently "lazy" (p. 160).

If, rather than using internally generated task cues based on the predictability of task sequences, participants used the stimuli themselves as task cues, the results reported in this chapter are consistent with what would be predicted using the Cognitive Control Model (CCM, Altmann & Gray, 2008). The CCM model posits that preparation effects occur because longer CSIs allow for more cycles of activation to be applied to the online task set prior to stimulus presentation, decreasing RTs (particularly on switch trials). Thus when stimuli are acting as cues, the length of the RSI has no impact on the time taken to activate a task set to a level where it can direct behaviour to produce appropriate responses.

Summary

Experiments 4A, 4B, 5A, and 5B produced evidence consistent with practice effects but not preparation effects with a recognition task and a magnitude task. The experiments used more naturalistic conditions—tasks with different stimulus sets and no external cues or explicit preparation instructions. The observation of similar sized practice effects for the short and long RSIs provided evidence that task set reconfiguration/activation was stimulus-driven as opposed to endogenously controlled. Likewise, the failure to observe significant preparation effects suggests that participants did not use the time between trials to prepare for upcoming task switches. Future research should attempt to determine what conditions must be met in order for participants to use a RSI to prepare in advance of stimulus presentation when tasks use different stimulus sets.

Chapter Nine

General Discussion

In his review of task switching, Logan (2003) noted that in order to advance task switching theory, new methods and new tasks needed to be used so as to gather “many examples of how tasks are controlled to learn what is general and what is not” (p. 48). Toward achieving this goal, this thesis extended the work of Jones et al. (2011) by examining task switching effects with a recognition memory task. Seven experiments explored the costs associated with switching between a recognition task and a magnitude task. Beyond my interest in the costs to recognition memory performance, the recognition and magnitude tasks were selected because they had limited stimulus-response (S-R) overlap and were trial unique (i.e., no stimulus repetition at test). The unique characteristics of the tasks and the methods used to examine task switching costs produced a series of novel experiments that tested the boundaries of task switching effects.

Key Findings

There were four key findings in this thesis. First, there were RT and accuracy switch effects. When participants switched from one task (e.g., the magnitude task) to another task (e.g., the recognition task) their responses were consistently slower and often less accurate relative to that when no task switch occurred. Second, the switch effect on recognition accuracy did not differ for deeply or shallowly encoded words, suggesting that the accuracy switch effects did not reflect a specific deficit to consciously controlled recollection processes. Third, characteristics of the switch effect changed as a function of experience within the experiment—there was evidence of practice effects. Finally, there was no evidence of advanced preparation. Providing long RSIs or sequence cues did not decrease the size of the RT switch effects.

Task switching effects.

Table 9.1 shows the RT switch costs (switch RT – repetition RT) and accuracy switch costs (repetition accuracy – switch accuracy) of Experiments 1 – 5B. RT switch effects were significant for all experiments. Switching from the recognition task to the magnitude task (or vice versa) required a change in the ‘online’ task set, which took

time to complete. While this result might seem unsurprising to the reader based on their own ‘real life’ experience and is consistent with some previous studies (Rogers & Monsell, 1995), it is important to acknowledge that the observation of switch effects in the absence of S-R overlap is inconsistent with some early experiments (e.g., Allport et al., 1994; Jersild, 1927; Spector & Biederman, 1976). Nonetheless, the finding is consistent with the related interruptions literature, where significant interruption effects have been reported for tasks that use different stimuli (e.g., Altmann & Trafton, 2007; Gillie & Broadbent, 1989, Trafton et al. 2003).

The size of the RT switch cost ranged from 37 ms to 366 ms, and was generally larger for the recognition task than that for the magnitude task. With the exception of Experiment 1, the RT switch cost for each task remained fairly consistent across experiments. The larger RT switch cost for the recognition task in Experiment 1 likely

Table 9.1.

Summary of Accuracy and RT Switch Costs for the Recognition and Magnitude Tasks Across Experiments 1 – 5B.

Experiment	Recognition				Magnitude			
	N	Accuracy		RT	Accuracy		RT	
1	96	0.02		366 **				
2	32	0.07	*	132 **				
3	120	0.08	**	137 **	0.06	**	49	**
4A	32	0.09	**	104 **	0.05	*	37	**
4B	32	0.04	#	87 **	0.02	#	60	**
5A	32	0.04	#	122 **	0.05	*	50	**
5B	64	0.05	*	115 **	0.03	**	54	**

Notes: # $p < 0.10$. * $p < 0.05$. ** $p < 0.01$. The RT switch effect is represented as a cost in ms. (switch RT – repetition RT). The accuracy switch effect is represented as a cost in proportion correct (Repetition – Switch). Experiments 1 and 2 did not analyse magnitude task performance, therefore these data were not included in this table. Experiment 1 is collapsed across all three run lengths as none of these conditions differed significantly. Likewise, Experiment 4 is collapsed across the visual cue conditions.

reflected that this experiment did not use a response delay or response deadline. The task differences in RT switch cost were consistent with a reconfiguration-based task switching account, with the relatively high long-term memory retrieval demands of the recognition task producing larger RT switch costs than that of the magnitude task (e.g., Mayr & Kliegl, 2000). Alternatively, the larger sized RT switch cost for the recognition task could reflect that word stimuli are encountered frequently in everyday life and are associated with other tasks such as reading. By contrast, the dot arrays were relatively novel and were unlikely to be associated with alternative tasks. The magnitude task may have required less activation than the recognition task to overcome any latent competing task sets, thus producing faster RTs on switch trials than that for the recognition task (see Altmann & Gray, 2008).

Given that this thesis used relatively novel tasks with limited S-R overlap, it is difficult to make direct comparisons to earlier research regarding the size of the RT switch effect. One potential point of comparison in the literature is Rogers and Monsell's (1995) fourth experiment, which used stimuli with limited S-R overlap for each task (e.g., '#9' for parity and 'A#' for consonant/vowel decisions).¹⁵ The RT switch costs for the magnitude task used in the current experiments were similar to the costs reported in Rogers and Monsell (43 – 67 ms, depending on RSI). By contrast, the RT switch costs for the recognition task used in the current experiments were larger than the costs reported in Rogers and Monsell. This likely reflects the higher long term memory retrieval demands (and possibly latent interference) for the recognition task than that for the parity and letter tasks in Rogers and Monsell.

With regard to the recognition RT switch effects, comparisons should also be made to previous task switching and interruptions literature that has utilised recognition memory tasks. The recognition task RT switch effects reported in this thesis were considerably smaller than those reported by Mayr and Kliegl (2000) when participants switched between two memory tasks using the same stimulus sets. By contrast, the recognition task RT switch effects reported in this thesis were very similar to those reported by Werkle-Bergner et al. (2005) who also used the same stimulus set for both tasks. When different stimulus sets were used for both tasks, as was the case in Jones et al. (2011), there was again some inconsistency. The recognition task RT switch effects

¹⁵ The mean RTs for Rogers and Monsell's (1995) tasks were slightly faster ($M = 630$ ms for repetition trials) than that for the recognition and magnitude tasks in this thesis. As such, this direct comparison of RT switch cost is made cautiously.

reported in this thesis were similar (albeit slightly larger) in size to those reported in Experiments 1A and 1B, but considerably smaller than those reported in Experiment 2 (~314 ms). Clearly, additional factors beyond S-R overlap must contribute to the size of the RT switch effect.

Switch effects on accuracy.

Although significant accuracy switch effects were less consistently reported, repetition trial accuracy was numerically higher than that for switch trial accuracy for both tasks across all experiments. A Wilcoxon signed ranks test conducted on mean repetition and switch trial accuracy for the recognition task produced a significant difference consistent with a task switching effect ($Z = -2.32, p = 0.020$). Switching tasks had a negative effect on participants' ability to discriminate between 'old' and 'new' words. These task switch effects on recognition accuracy were similar in size to those reported in Mayr and Kliegl (2000). They were also similar (but much more reliable) than the stimulus switch effects reported in Kavcic et al. (1999). This difference in reliability between the current research and Kavcic et al. likely reflects the more extensive/demanding processes required when switching tasks as well as stimulus sets. With regard to the accuracy switch effect on the magnitude task, a Wilcoxon signed ranks test conducted on the mean magnitude task accuracy across Experiments 3 – 5 showed a significant difference between repetition and switch accuracy ($Z = -2.03, p = 0.042$). When asked to respond whether a dot array contained more or less than a specified number of dots, participants' were less accurate following a task switch than following a task repetition. An important property of this research was that accuracy switch effects could not be attributed to applying the incorrect task rules to a stimulus: it was impossible to make a recognition memory response to an array of dots, just as it was impossible to make a magnitude response to a word stimulus. Therefore the errors produced on a task switch reflected occasions where participants failed to process the stimulus and determine the correct response based on the task rules.

Switch effects on recognition processes.

One aim of this thesis was to determine if specific memory processes were impaired when participants switched tasks. In particular, it examined whether task switching had a similar effect on recognition memory as divided attention, which affects the controlled recollection process more than the automatic familiarity process (Hicks & Marsh, 2000; Jacoby, 1991; Jones & Jacoby, 2001; Knott & Dewhurst, 2007a).

However, although Experiment 2 reported a significant level-of-processing effect—higher accuracy for deeply encoded words than that for shallowly encoded words—the size of the effect remained unchanged for repetition and switch trials. That is, there was no evidence that switching tasks differentially affected the controlled recollection processes of recognition memory. A survey of hit and false alarm rates across all experiments (see Appendix B) provided little evidence that switching tasks affected participants' preferences or strategies toward responding old or new. Combined, these results suggest that, at least for the experiments presented in this thesis, the accuracy switch effect on recognition memory performance was more general and reflected an impaired ability to discriminate between the latent familiarity of 'new' words and the familiarity of studied 'old' words. It is worth noting that the response signal delay (RSD) used in this thesis was at the short end of the range that Boldini et al. (2004) reported level of processing effects (and thus a larger role for recollection). Thus it may be valuable to lengthen the RSD in future research to further encourage the use of recollection.

An alternative (albeit unlikely) explanation for the accuracy switch effects was that the decrease in accuracy reflected a failure of memory for response sets/rules (Meiran, 2000a, 2000b). That is, one could argue that the decrease in accuracy on switch trials reflected a systematic failure to retrieve the correct motor responses for each decision (e.g., pressing 'j' for an old response instead of 'f') that was spontaneously corrected on the following repetition trial. However, as no feedback was provided between trials, episodes of spontaneous self-correction were likely rare. Rather, participants who retrieved the wrong task rules probably applied these rules across both trials in the run, producing lower overall accuracy rather than larger differences between switch and repetition accuracy. As persistent cases of response-to-key reversal produced accuracy well below chance, these participants would have been identified and excluded from the dataset based on the strict performance criteria used for these studies. Finally, Jones et al. (2011) reported significant accuracy switch effects of a similar (if slightly larger) size even though spatially congruent response cues were presented below each word stimulus. Based on these observations, an explanation that identifies the periodic retrieval of incorrect response sets as the source of the recognition accuracy switch effect is unlikely to adequately explain the data. A similar argument can be made in the case of magnitude task performance.

Switch effects and practice.

For the experiments reported in this thesis, participants had limited experience with the recognition and magnitude tasks before the first actual test began (both individually and switching between them). Participants' performance changed as they developed the skills necessary to quickly and efficiently switch tasks over the course of the experiment. The role of practice was first highlighted in Experiment 1, where mean RTs were plotted as a function of experiment third (see Figure 5.3, p. 53). A visual comparison of the three experimental conditions (where critical trials were switch trials) showed a larger decrease from the first to last third than that of the control condition (where critical trials were repetition trials). Thus, in addition to practice improving the overall speed of responding, Experiment 1 suggested that this improvement was larger for switch trials than that for repetition trials. This trend was confirmed statistically in Experiment 3, where recognition task RTs changed from the first to second test block. There was a substantial decrease in the difference between Position 1 and Position 2 RTs; the cost of switching tasks decreased with practice.

Practice effects were examined more purposefully in Experiments 4A – 5B, where task switching effects were compared across test blocks using different RSIs. Collapsed across both tasks, RT switch costs decreased from the first to the second test block (although this difference was not statistically significant in Experiment 4A). Consistent with research that has compared task switching performance across test blocks (Jersild, 1927; Koch, 2005; Kramer et al., 1999, Kray & Lindenberger, 2000; Meiran, 1996) or test days (Monsell et al., 2003; Rogers & Monsell, 1995), experience with tasks and with switching between them generally produced more efficient task switching performance. Practice strengthened the association between a task cue (in this case, the dots or words) and the task set/rules, resulting in the faster retrieval and engaging of the online task set on switch trials (Meiran, 1996). With a few exceptions (e.g., Koch, 2005; Minear & Shar, 2008), task switching research has focused on performance after the skill of switching tasks has been developed and performance has stabilised. The results of the current thesis showed that early task switching trials that are usually excluded from analysis provided valuable information about task switching effects and the development of task switching skills. The benefit of experience built up quickly and produced fairly consistent evidence of practice effects.

Switch effects and advanced preparation.

There is an abundance of evidence in the task switching literature that when task switching processes are engaged before the presentation of a stimulus, the size of task switching effects are decreased (see Altmann, 2004b). However, in the experiments reported in this thesis the RT switch costs remained relatively consistent across a number of experimental manipulations conducive to advanced preparation. The absence of evidence for advanced preparation has been discussed in detail in Chapters 7 and 8, and thus this section will focus on critical findings and conclusions. Key experimental properties such as the use of recognition and magnitude tasks, the absence of external cues, and the absence of explicit instructions to prepare were purposefully selected to test the boundaries of the preparation effect in naturalistic conditions. No difference in the size of the RT switch cost was reported across a variety of RSIs from 200 ms to 2600 ms. Even when two RSIs were used in the practice test prior to the actual test (ensuring recent exposure to both RSIs) there was no indication that participants used the RSIs to prepare for a known change in task.

Furthermore, the provision of a visual sequence cue in Experiment 3 had no effect on task switching performance or recovery from this task switch. In the Position 4 sequence cue condition the visual sequence cue reminded participants that they would need to switch tasks on the next trial and there was sufficient time for this information to be used (Monsell, 2003; Rogers & Monsell, 1995). Regardless of these ‘ideal’ conditions, the RT and accuracy switch effects were no different to the control condition where no sequence cue was provided. Considered together, the results of Experiments 3 – 5B produced very convincing evidence that within the context of switching between a recognition task and a magnitude task, participants did not begin to switch tasks in advance of stimulus presentation. This lack of evidence for advanced preparation shows that task switching phenomena that have been established with simple tasks and highly controlled experimental settings may not generalise to less typical situations and tasks. In particular, the preparation effect may be as much (or more) a product of the tasks and methods we use to examine them as they are a product of natural task switching processes or strategies (see Altmann, 2004b for a similar suggestion).

Interpreting the Findings: Task Set Reconfiguration

The results of this thesis can be interpreted using established task switching theories. Of particular interest are reconfiguration-based accounts of task switching

(e.g., de Jong, 2000; Meiran, 2000a, 2000b; Rogers & Monsell, 1995; Rubinstein et al., 2001; Sohn & Anderson, 2001). As was described in Chapter 3, there are a variety of different conceptual and computational reconfiguration models of task switching. For the sake of this discussion, the results will be interpreted within the model outlined by Rubinstein et al., where performance is modelled by two processes. The first—the executive control process—involves changing task goals and rules in working memory, which can be initiated (and sometimes completed) prior to stimulus presentation given suitable conditions. This process only occurs when the task switches. The second—the task process—involves identifying stimuli, selecting and executing a response, which can only be initiated after stimulus presentation. This process occurs on every trial, regardless of task transition. The absence of preparation effects and the null effect of the visual sequence cue suggests that in the case of the current experiments, at least the rule activation stage of the executive control process was engaged upon presentation of a stimulus—it was exogenously triggered. Over time, the rule activation stage was engaged more quickly as the association between task stimuli (or stimulus type) and task sets were strengthened, thus producing the practice effects reported in Experiments 3 – 5B.

Prior research has shown that when task switches are predictable and task cues are absent, participants monitor tasks covertly to produce internally generated task cues (Baddeley et al., 2001; Emerson & Miyake, 2003; Miyake et al., 2004). These types of monitoring and task cue generation could be interpreted as functions of the goal setting stage of the executive control process. Interestingly, researchers have suggested that these types of internally generated cues, although sufficient to engage the goal setting stage, fail to engage the subsequent rule activation stage of reconfiguration that is essential for producing preparation effects (Koch, 2003). With this in mind, the absence of preparation effects in this thesis is insufficient evidence to completely rule out a role for endogenous engaging of the executive control process. When switching between the recognition and magnitude tasks, participants may have engaged the goal setting stage via internally generated task cues, but not the rule activation stage.

Alternatively, both stages of the executive control process may have been initiated after stimulus presentation. Indeed, the novel characteristics of tasks and stimuli in this thesis may have made internal monitoring of task sequences unnecessary or even counterproductive. When the same stimuli are used for both tasks (as is typical in task switching experiments), producing a task-appropriate response requires that a

stimulus be processed and interpreted in conjunction with information about the task (see Arrington & Logan, 2003). With the exception of occasions when participants make use of task sequence predictability—as is possible in the alternating runs paradigm—this information is generally conveyed through the use of an external task cue. In the experiments of this thesis, as the recognition and magnitude tasks used different stimulus sets (i.e., words for the recognition task and dots for the magnitude task) the stimuli themselves could be used as reliable task cues. Very soon after presentation the basic features of a given stimulus could be used to accurately decipher the task to be performed on that trial. By contrast, internally generated cues may have required substantially more cognitive resources (e.g., maintaining task sequence and position within a run) and when these internal cues prompted a task switch they had the potential to produce errors (e.g., initiating a task switch prematurely). Just as performance in Experiment 3 suggested that participants chose to not use the visual sequence cue to prepare for task switches, it may be that participants also chose not to monitor the predictable task sequence for similar reasons. At the prospect of applying additional cognitive effort for very little payoff, participants may have chosen (implicitly or otherwise) to instead rely on the stimuli to guide them to switch or repeat tasks.

Interpreting the Findings: Cognitive Control Model

Despite the absence of external cues and a lack of interference due to the limited S-R overlap in the current experiments, the Cognitive Control Model (CCM) of task switching (Altmann & Gray, 2008) provides an interesting framework for interpreting the results of this thesis. As has been noted previously, the stimuli (or stimulus sets) used in this thesis can perform the role of task cues. Therefore, the process of applying activation to the online task set through successive encoding cycles could be initiated by the presentation of the stimulus. Indeed, the results reported in this thesis are largely congruent with key concepts of the model. From a CCM perspective, the failure to produce preparation effects and visual sequence cue effects indicates that any internally generated task cues that were produced prior to stimulus onset were not sufficient to initiate cycles of encoding/activation. With the stimulus itself acting as an explicit task cue, the duration of the RSI had no impact on the time taken to activate a new task set. The RTs reflected the amount of time between stimulus onset and when task set activation met threshold and an appropriate response could be selected and executed. By

extension, the RT switch cost reflected the difference in the amount of time/activation required for repetition and switch trials to reach this threshold.

Drawing from evidence of run-length effects (e.g., Altmann, 2002; Altmann & Gray, 2002), CCM posits that task sets decay when they are not being used (or sampled). With this in mind, one could hypothesise that the longer it has been since the online task set was used (i.e., the longer the ‘offline’ run length was) the more activation/encoding cycles should be required to attain threshold levels of activation. Thus the size of RT switch effects should decrease as run length increases. However, Monsell et al. (2003) reported no significant differences in the size of RT switch effects for their predictable 2-, 4-, and 8-trial runs, perhaps reflecting the opposing effects of run length on levels of interference from the offline task set and the extent of decay for the online task set on switch trial RTs.¹⁶ By contrast, this thesis represented a case where there would be little interference from the offline task set, thus theoretically allowing one to isolate the contribution of run length to the decay of a task set. Despite this, the main effect of run length in Experiment 1—although in the hypothesised direction—was not significant. Furthermore, a cursory comparison of the size of RT switch effects for Experiment 3 (4-trial runs) and Experiments 2, 4, and 5 (2-trial runs) showed no overwhelming differences. Any effect of run length on the size of the RT switch effect was too subtle to be observed within the experimental conditions (particularly the relatively few trials) used in this thesis. It is worth noting that a 4-trial run length is still considerably shorter than the long run-lengths used in prior research (Altmann, 2002; Altmann & Gray, 2002, 2008; Monsell et al., 2003). Consequentially, further research with the magnitude and recognition tasks using more trials and longer run lengths may increase the likelihood of producing significant run length effects.

When the lack of run length effect is considered in conjunction with the limited evidence of within-run slowing in Experiment 3, one could posit that the rate of decay for task sets may be particularly shallow. Altmann and Gray (2008) have argued that when run-lengths are short, decay functions are adjusted upwards (i.e., decay occurs more quickly) to decrease levels of interference and produce optimal performance. The experiments reported in this thesis may represent cases where decay rates were adjusted in the opposite direction: in response to limited interference between the magnitude and recognition tasks, the rate of decay may have been slowed so as to optimise

¹⁶ Research on run length effects by Altmann (2002) and Altmann and Gray (2002) did not report analyses of the size of the RT switch effect as a function of the length of the preceding run.

performance. As an additional benefit, a slower rate of decay would maintain a higher level of task set activation for longer, decreasing the potential for interference from latent task sets associated with a stimulus type. A higher level of activation (for Task B) going into a task switch (from Task A to Task B) would produce smaller RT switch effects. With this in mind, a slower decay rate may also provide a good theoretical explanation for the smaller (or absent) task switching effects typically reported in experiments where tasks have limited S-R overlap compared to when there is S-R overlap (Allport et al., 1994; Jersild, 1927; Rogers & Monsell, 1995; Spector & Biederman, 1976). However, this interpretation of task switching performance in the current experiments using CCM is clearly post hoc, and would require further evidence for confirmation. Examining performance over longer run lengths and comparing the gradient of slowing (if observed) to that reported when tasks share stimulus sets would be a first step toward confirming the role of changes in decay rates for experiments that use tasks that do not share stimulus sets.

Alternative Sources of Switch Cost

It was noted in the introduction that one of the advantages of using two tasks with little S-R overlap was that it limited interference from the offline task set. As such, any observed RT switch cost was likely to reflect the additional time required to reconfigure or activate a task set on switch trials relative to repetition trials. The contribution of time to resolve task conflicts and/or proactive interference was likely limited. However, the nature of the recognition memory task did create a potential (although unlikely) source of negative priming (see Allport & Wylie, 2000; Wylie & Allport, 2000). In Experiments 2 and 4A – 5B, participants made perceptual and/or semantic judgements to words during the study phases. These task sets (e.g., ‘does this word represent a living or non-living thing?’) may have been un-intentionally elicited when words were presented during the test phase, producing longer RTs, particularly on switch trials. However, there are reasons to discount negative priming as a substantial source of RT switch cost in this thesis. Allport and Wylie (2000) have shown that this type of negative priming is largely item-specific on the switch trial, and therefore should have produced larger RT switch costs for old words than that for new words (which were not encountered during the study phase). However, there were no differences in the size of the RT switch cost for ‘old’ and ‘new’ words in any experiment reported in this thesis, suggesting that negative priming had no impact on

task switching performance for the recognition task. Finally, if negative priming contributed to recognition task RT switch costs, then one would expect that the RT switch cost would be smaller when semantic/perceptual judgements were not made in study phases (e.g., Experiments 1 and 3). This was not the case.

Challenging Definitions of Task Switching

Influenced by early reports that switching between simple tasks with different stimulus sets (and limited S-R overlap) did not produce task switching effects (Allport et al., 1994; Jersild, 1927; Spector & Biederman, 1976), many definitions of task switching (and task switching effects) have explicitly or implicitly stated that this phenomenon is limited to situations where tasks share a stimulus set (e.g., Altmann, 2006; Arbuthnott & Frank, 2000; Dreisbach & Haider, 2008; Monsell, 2003; Wylie & Allport, 2000). As a consequence, some researchers may view the results of this thesis and challenge the applicability of the terms ‘task switching’ and ‘task switching effects.’ In response to this potential claim I present two arguments for the results reported in this thesis representing task switching effects. The first response appeals to the desire for parsimony in explanations of cognitive processes and suggests that the results reported in this thesis illustrate a key component of task switching effects. The second response highlights a contradiction in the literature between the ‘real life’ and experimental definitions/examples of task switching effects, and shows how the current research provides a first step toward bridging these two definitions to produce consistent task switching effects.

As has been established, in order to accurately perform a task, participants must use the appropriate task rules (i.e., a task set) to identify the relevant features of a stimulus and use them to formulate a task-appropriate response. When the task changes the rules used to direct behaviour must do so as well, regardless of whether tasks share stimulus sets. It seems logical (and parsimonious) to propose that the basic processes of changing task set would be the same regardless of S-R overlap. What may differ is the timing of these processes (pre- or post-stimulus) and whether additional mechanisms are needed to resolve instances of conflict or interference. Given the evidence that task switching processes were largely stimulus-driven, the current thesis provides a clear picture of the time and cognitive effort required to switch task sets without the influence of additional costs due to interference (e.g., Allport et al., 1994; Altmann, 2002; Altmann & Gray, 2002, 2008), inhibition (e.g., Arbuthnott & Franks, 2000; Mayr &

Keele, 2000), or negative priming (e.g., Allport & Wylie, 2000). The RT and accuracy performance costs of switching between tasks with limited S-R overlap should not be discounted simply because the contribution of additional components was negligible. Rather, their value as a relatively pure measure of an important component of task switching effects needs to be utilised to further our understanding of task switching. Indeed, the results produced by this type of experimental design may provide an excellent linking point between task switching theories and interruption theories (e.g., Altmann & Trafton, 2002, 2007) which arguably share many components.

An interesting contradiction has developed in the field of task switching research between the ‘real life’ descriptions of task switching used when introducing the concept, and the definitions (and limitations) of task switching as operationalized from an experimental perspective. As noted previously, many definitions of task switching effects have emphasised the use of a single stimulus set (and thus, high S-R overlap) for all tasks in an experiment. Indeed, modern researchers have largely accepted the assumption that tasks that use different stimulus sets represent a “boundary condition” – a point at which task switching effects will not be observed (Wylie & Allport, 2000, p. 213). Yet despite this perceived limitation of the task switching effect, researchers continue to use examples such as switching between writing a manuscript and talking to a student who drops by the office to illustrate the everyday pervasiveness of task switching effects (see Altmann, 2004a; Logan, 2003; Monsell, 2003; Vandierendonck et al., 2010 for examples).¹⁷ Just as there are few people who would deny that this ‘real life’ example describes two tasks with two different stimuli, there are few people that would deny the assertion that performance immediately after switching tasks is slower and often more error-prone. The results reported in this thesis provided convincing evidence that experimental methods of examining task switching *can* produce significant switching effects when two tasks do not share stimulus sets. With this (and early comments about parsimony) in mind, I argue that the experimental definition of task switching should not be limited to situations where there is S-R overlap. Rather, these definitions should be more expansive and reflexive, accepting that many experimental factors can influence the occurrence (and extent) of task switching effects.

Alternative reasons for the null effects reported by early researchers using tasks with different stimulus sets (Jersild, 1927; Spector & Biederman, 1976) should be

¹⁷ It is also worth noting that these types of examples bear more similarities with interruptions than they do with task switching, particularly from an experimental context.

considered. For example, both Jersild and Spector and Biederman used self-paced tests where participants progressed through mixed blocks of relatively simple tasks. It is worth considering that the complexity of the task set and the short run length (1-trial runs) may have made it possible to maintain both task sets in working memory. Furthermore, without control over RSIs (these tests were self-paced) the appropriate task set may have been fully engaged prior to stimulus onset, further reducing mixing costs. Some researchers have maintained that the residual component of the RT switch effect reflects the contribution of task set inhibition when switching away from a task that was likely to interfere with the performance of the next task (Mayr & Keele, 2000). In the absence of such potential interference (and thus, inhibition) and with sufficient time and information to prepare for a known change in tasks, it is not surprising that the size of mixing costs reported in Jersild and in Spector and Biederman were limited. Critically, this does not mean that a change in task sets did not occur or did not require some cognitive resources, but that this process was not effectively captured by the variable chosen to measure it (i.e., RTs and mixing costs).

Future Directions

A question that remains to be answered is whether the stimulus-driven nature of task switching in this thesis was the product of the experimental design, the choice of secondary task, or the unique qualities of the recognition memory task (or any combination of these factors). Clarifying this issue will require additional research focused not only on the tasks utilised, but also the method used to examine task switching effects. One line of enquiry would be to systematically manipulate the level of S-R overlap between the recognition task and the secondary task. A greater level of overlap could be achieved by presenting a word and dot compound for each trial and instructing participants to switch between responding to different elements within this compound stimulus (see Paulitzki, Risko, Oakman, & Stolz, 2008 for an example of this compound stimulus method). In this case, the two tasks (and task sets) will remain the same but the value of an individual stimulus as a task cue will be diminished, perhaps forcing participants to initiate the process(es) of switching tasks in advance of stimulus onset. However, it is worth noting that this type of change would introduce a level of interference that was not present in the experiments of this thesis that could impact not only the size of RT and accuracy switch effects, but the components that make up these effects. Alternatively, switch effects to recognition memory could be examined using an

explicit task cuing paradigm. With this change, the two tasks (and limited S-R overlap) could be maintained, but by removing the predictable task transitions and providing explicit external cues in advance of stimulus onset, participants may be more likely to prepare (Altmann, 2004b; Koch, 2003). Comparing the products of these two extensions would provide valuable insight into the relative roles of interference and strong cues in encouraging advanced preparation when switching tasks, while maintaining many of the novel characteristics of the experiments presented in this thesis.

On a related topic, the results of this thesis have also highlighted a need to clarify the role of internal monitoring of predictable sequences when switching tasks, both in general and when participants are switching between tasks with limited S-R overlap. The arguments for and against the proposal that participants actively monitored the predictable task sequence were presented earlier in this chapter, but remain unresolved. A starting point for clarifying the role of internally generated task cues produced by monitoring predictable task sequences would be to use articulatory suppression (e.g., Baddeley et al. 2001; Emerson & Miyake, 2003) during the RSI to deliberately limit the use of these internal task cues and to compare the size of task switching effects to those reported without articulatory suppression. These types of manipulations, administered across a variety of tasks and S-R mappings, may shed light on the role of internal monitoring and experimental conditions that mediate its use.

Conclusions

In the seven experiments reported in this thesis participants switched between a recognition and magnitude task in predictable 1-, 2-, or 4-trial runs. The use of tasks that do not share S-R overlap, and in particular the use of a recognition memory task, represents a considerable step toward the ‘real life’ tasks used in the interruptions literature. Participants’ performance on task switch trials was consistently slower and often less accurate than that for task repetition trials, illustrating typical task switching effects. Furthermore, the accuracy switch effects on recognition memory shows that frequent switching between tasks can have a negative effect on episodic memory *in addition* to the accepted negative effect on general productivity/speed.

However, this cost of switching tasks was largely unaffected by run length, the use of sequence cues, or the length of the RSI. The results of this thesis provide strong evidence that significant task switching effects *can* be observed when two tasks do not share stimulus sets and have a limited S-R overlap. Participants’ task switching

performance can be readily interpreted within established task switching theories, and represents an occasion where the processes involved in switching tasks were largely engaged after stimulus presentation (i.e., task switching was exogenously driven). The experiments reported in this thesis represent an important step towards expanding the scope of experimental research into task switching. They emphasise the importance of a variety of experimental characteristics (i.e., tasks, S-R overlap, practice) in determining how the processes involved in switching tasks are manifested in performance (particularly RTs). Future research will help clarify the roles of these and other experimental characteristics and will help develop more refined definitions of the mechanics and costs of switching tasks.

Appendix A

Stimuli

Table A.1

Critical Words from Experiment 1.

Word	KF	Letters	Syll.	Word	KF	Letters	Syll.
ADDRESS	77	7	2	FAILURE	89	7	2
ANGLE	51	5	2	KITCHEN	90	7	2
APPEAL	62	6	2	LAGOONS	50	7	2
BALANCE	90	7	2	LEAGUE	69	6	1
BOTTOM	88	6	2	MESSAGE	64	7	2
CAPTAIN	85	7	2	MISSION	78	7	2
CAREER	67	6	2	PLEASURE	62	8	2
COMMAND	72	7	2	PROJECT	93	7	2
COMPLEX	91	7	2	RELIEF	66	6	2
CONSTANT	71	8	2	SESSION	80	7	2
COVER	88	5	2	SILENCE	52	7	2
CREDIT	64	6	2	STATUS	97	6	2
DEAFEN	70	6	2	SYMBOL	54	6	2
DETAIL	72	6	2	VALLEY	73	6	2
DINNER	91	6	2	WINDOW	119	6	2

Note. KF = Kucera and Francis (1967) word frequency. Syll. = Syllables

Table A.2

Filler and Buffer Words from Experiment 1.

Word	KF	Letters	Syll.	Word	KF	Letters	Syll.
ACCOUNT	117	7	2	MARKET	155	6	2
AMOUNT	172	6	2	MARRIAGE	95	8	2
AMULETS	133	7	3	MASTER	72	6	2
ARGUMENT	63	8	3	MATERIAL	174	8	4
ARTICLE	68	7	3	MEMBER	137	6	2
AUDIENCE	115	8	3	MINIMUM	64	7	3
BENEFIT	63	7	3	MISTED	63	6	2
BREATH	53	6	1	MORESEL	93	7	2
BURLESQUE	97	9	2	MOTION	55	6	2
CAPITAL	85	7	3	NATION	139	6	2
CASHEWS	64	7	2	NATURE	191	6	2
CHANCE	131	6	1	NOBODY	74	6	3
CHECK	88	5	1	NOVEL	59	5	2
CONCERN	98	7	2	OPINION	96	7	3
CONFLICT	52	8	2	PEACE	198	5	1
CONSTRIC	93	9	2	PERFECT	58	7	2
CONTORT	69	7	2	PICTURE	162	7	2
CONTRACT	60	8	2	PLANT	125	5	1
CORNET	76	6	2	PLAYER	51	6	2
COUPLE	122	6	2	PORTION	62	7	3
CULTURE	58	7	2	PRICE	108	5	1
DANCE	90	5	1	PROPERTY	156	8	3
DANGER	70	6	2	REACTION	124	8	2
DEGREE	125	6	2	RESPECT	125	7	2
DESIGN	114	6	2	RIVER	165	5	2

Word	KF	Letters	Syll.	Word	KF	Letters	Syll.
DESIRE	79	6	2	SCALE	60	5	1
DISEASE	53	7	2	SCAMPER	75	7	2
DISTANCE	108	8	2	SCENE	106	5	1
DISTRICT	135	8	2	SEASON	105	6	2
DIVISION	107	8	3	SECRET	78	6	2
EFFORT	145	6	2	SERIES	130	6	2
ENGINE	50	6	2	SHAPE	85	5	1
ENTRENCH	63	8	2	SLEEP	65	5	1
EXPENSE	50	7	2	SOURCE	94	6	1
FLESH	52	5	1	SPACE	184	5	1
FLIGHT	46	6	1	SPEED	83	5	1
FORWARD	115	7	2	STAFF	113	5	1
FRAME	74	5	1	STAPLER	76	7	3
FRIEND	133	6	1	STATION	105	7	2
GARISH	117	6	2	STEEPLE	86	7	2
GROUND	186	6	1	STORY	153	5	2
GROWTH	155	6	1	SUPPLY	102	6	2
HALLOW	68	6	2	SURFACE	200	7	2
IMAGE	119	5	2	TASTE	59	5	1
INDUSTRY	171	8	3	THEME	55	5	1
INSTITUTE	50	9	3	THEORY	129	6	2
ISLAND	167	6	2	TRADE	143	5	1
KNIFE	76	5	1	TRUCK	57	5	1
LIBRARY	62	7	3	VICTORY	61	7	3
LINSEED	132	7	2	WEIGHT	91	6	1
MACHINE	103	7	2	WHEEL	56	5	1

Note. KF = Kucera and Francis (1967) word frequency. Syll. = Syllables

Table A.3

Critical Words from Experiment 2 and 4A – 5B.

Word	KF	Syll.	Letters	Imagery	Concrete
ANCHOR	15	2	6	6.5	6.3
ANGEL	18	2	5	5.9	4.0
APPLE	9	2	5	6.7	6.9
ARMOR	4	2	5	5.5	5.6
ARROW	14	2	5	6.4	6.8
AUTUMN	22	2	6	6.3	5.4
BANNER	8	2	6	5.8	5.8
BARGAIN	7	2	7	4.2	3.9
BARREL	24	2	6	5.6	6.8
BASIN	7	2	5	5.3	6.5
BEAVER	3	2	6	6.4	6.9
BEGGAR	2	2	6	6.1	5.8
BERRY	9	2	5	5.5	6.6
BOUNDARY	16	2	8	4.8	5.0
BUBBLE	12	2	6	6.1	6.2
BUSHEL	1	2	6	5.4	6.4
BUTCHER	8	2	7	6.0	6.7
BUTTON	10	2	6	6.3	6.8
CABIN	23	2	5	6.5	6.5
CABLE	7	2	5	5.0	6.4
CANDY	16	2	5	6.3	6.7
CANNON	7	2	6	6.5	6.4
CANOE	7	2	5	6.4	6.8
CANVAS	19	2	6	5.1	6.2
CARBON	30	2	6	4.5	5.9

Word	KF	Syll.	Letters	Imagery	Concrete
CARRIAGE	11	2	8	6.2	6.3
CASTLE	7	2	6	6.7	6.7
CHERRY	6	2	6	6.3	6.8
CHIMNEY	7	2	7	6.2	6.7
CIRCUIT	23	2	7	3.8	4.7
CLOTHING	20	2	8	6.2	6.4
CLUSTER	13	2	7	4.3	4.5
COLLAR	17	2	6	6.0	6.3
COMPASS	13	2	7	5.8	6.6
COMPOUND	11	2	8	2.7	3.3
COMRADE	4	2	7	4.2	5.2
CREATURE	15	2	8	4.4	5.0
CRYSTAL	23	2	7	5.8	6.1
DAYLIGHT	15	2	8	6.5	6.2
DEALER	25	2	6	4.2	5.7
DEVIL	25	2	5	5.8	3.9
DOORWAY	15	2	7	5.8	6.7
DRAGON	1	2	6	6.2	5.9
EAGLE	5	2	5	6.3	6.9
ELBOW	10	2	5	6.0	6.8
ELDER	15	2	5	5.2	5.1
FARMER	23	2	6	6.2	6.6
FEATHER	6	2	7	6.2	6.5
FLAVOUR	16	2	7	3.5	4.6
FOREHEAD	16	2	8	6.3	6.9
FOUNTAIN	18	2	8	6.4	6.6

Word	KF	Syll.	Letters	Imagery	Concrete
GARMENT	6	2	7	5.4	6.6
GENIUS	23	2	6	4.9	3.8
GODDESS	3	2	7	4.6	4.0
HAMMER	9	2	6	6.5	6.8
HELMET	1	2	6	5.9	6.8
HUNTER	18	2	6	6.1	6.3
INSECT	14	2	6	6.3	6.7
KEEPER	3	2	6	3.3	5.3
KINGDOM	26	2	7	4.8	4.8
KITTEN	5	2	6	6.4	7.0
LAYER	12	2	5	3.8	4.6
LEMON	18	2	5	6.5	6.8
LINEN	6	2	5	6.1	6.6
MADAM	2	2	5	6.5	5.7
MAIDEN	2	2	6	5.4	5.7
MEADOW	17	2	6	6.1	5.9
MERCHANT	20	2	8	5.1	6.2
MIRROR	27	2	6	6.4	6.7
MISTRESS	5	2	8	5.2	5.7
MONARCH	3	2	7	5.1	5.9
MONKEY	9	2	6	6.5	6.9
MONSTER	6	2	7	6.0	5.3
NEPHEW	9	2	6	5.0	6.4
ODOUR	14	2	5	5.1	5.0
OLIVE	5	2	5	6.1	6.8
ONION	15	2	5	6.1	6.7

Word	KF	Syll.	Letters	Imagery	Concrete
ORANGE	23	2	6	6.6	6.1
ORCHARD	3	2	7	6.4	6.4
ORGAN	12	2	5	5.9	6.6
OUTLINE	12	2	7	3.9	4.6
OYSTER	6	2	6	5.8	6.8
PAINTER	21	2	7	6.0	6.7
PARCEL	1	2	6	6.3	6.5
PARENT	15	2	6	6.3	6.2
PEPPER	13	2	6	6.3	6.8
PIGEON	3	2	6	6.4	6.9
PILLOW	8	2	6	6.5	6.6
PISTOL	27	2	6	6.1	3.8
PLANET	21	2	6	6.1	6.1
PRINCESS	10	2	8	6.0	6.0
PUPIL	20	2	5	5.8	6.2
PUZZLE	10	2	6	5.5	5.5
RABBIT	11	2	6	6.6	7.0
RAILWAY	12	2	7	6.5	6.5
RECEIPT	4	2	7	4.9	6.0
RIBBON	12	2	6	5.8	6.5
RIDER	16	2	5	5.3	6.2
ROBBER	2	2	6	5.8	6.0
ROBIN	2	2	5	6.6	6.7
SAILOR	5	2	6	6.3	6.6
SALAD	9	2	5	6.3	6.5
SANDWICH	10	2	8	6.7	6.7

Word	KF	Syll.	Letters	Imagery	Concrete
SCHOLAR	15	2	7	5.5	6.3
SERVANT	19	2	7	5.2	6.0
SHEPHERD	3	2	8	5.8	6.5
SHERIFF	20	2	7	6.3	6.8
SHIPPING	19	2	8	3.8	4.5
SINGER	10	2	6	4.2	3.3
SPARROW	0	2	7	5.9	6.8
SPIDER	2	2	6	6.8	7.0
SQUIRREL	1	2	8	6.3	6.9
STANZA	0	2	6	3.6	5.4
STATION	1.5	2	7	5.7	6.0
STEAMER	1	2	7	5.0	6.4
STOCKING	1	2	8	5.9	6.5
SULPHUR	3	2	7	4.1	6.5
TIGER	7	2	5	6.6	7.0
TRAITOR	2	2	7	4.2	4.8
TURKEY	9	2	6	6.6	6.8
TWILIGHT	4	2	8	5.3	4.8
VAPOR	12	2	5	5.0	4.6
VELVET	4	2	6	6.2	6.3
VICTIM	27	2	6	5.5	5.4
WIDOW	26	2	5	5.4	6.3
WILLOW	9	2	6	6.1	5.9
WITNESS	28	2	7	4.4	5.1
WORKER	30	2	6	5.4	6.3

Note. KF = Kucera and Francis (1967) word frequency. Syll. = Syllables. Word imagery and concreteness from the Toronto Word Pool (Friendly et al., 1982)

Table A.4

Buffer and Filler Words from Experiment 2 and 4A – 5B.

Word	KF	Syll.	Letters	Imagery	Concrete
AFFAIR	33	2	6	5.3	3.7
ARRAY	11	2	5	3.0	2.9
CANAL	3	2	5	5.4	6.4
COLONEL	37	2	7	4.5	6.4
CONQUEST	9	2	8	4.6	4.7
DIAMOND	8	2	7	6.6	6.6
FINGER	40	2	6	6.3	6.9
HARBOUR	37	2	7	6.2	6.4
HARNESS	10	2	7	5.5	6.2
INSTANT	38	2	7	3.0	3.0
JEWEL	1	1	5	6.0	6.4
MOISTURE	10	2	8	5.3	5.5
OVEN	7	2	4	6.0	6.9
PARTNER	32	2	7	5.4	5.7
PATENT	35	2	6	2.3	4.7
PROVINCE	15	2	8	4.5	5.3
SISTER	38	2	6	1.6	2.5
SLIPPER	3	2	7	4.8	5.3
SOLDIER	39	2	7	4.4	4.6
STOMACH	37	2	7	5.8	6.6
TUNNEL	10	2	6	6.3	6.4
WORSHIP	36	2	7	4.7	2.8

Note. KF = Kucera and Francis (1967) word frequency. Syll. = Syllables. Word imagery and concreteness from the Toronto Word Pool (Friendly et al., 1982)

Table A.5

Critical Words from Experiment 3.

Word	KF	Letter	Syll.	Word	KF	Letter	Syll.
ACCOUNT	117	7	2	LIBRARY	62	7	3
ADDITION	142	8	3	LINSEED	132	7	2
ADDRESS	77	7	2	LOCATION	63	8	3
AGENCY	56	6	3	MACHINE	103	7	2
AMOUNT	172	6	2	MARKET	155	6	2
AMULETS	133	7	3	MARRIAGE	95	8	2
APPEAL	62	6	2	MASTER	72	6	2
ARTICLE	68	7	3	MATERIAL	174	8	4
ARTISTS	55	7	2	MEMBER	137	6	2
AUDIENCE	115	8	3	MESSAGE	64	7	2
BALANCE	90	7	2	METAL	61	5	2
BENEFIT	63	7	3	METHOD	142	6	2
BOTTOM	88	6	2	MISSION	78	7	2
CAPITAL	85	7	3	MISTED	63	6	2
CAPTAIN	85	7	2	MORSEL	93	6	2
CAREER	67	6	2	NATION	139	6	2
CASHEWS	64	7	2	NATURE	191	6	2
CHAPTER	74	7	2	NOBODY	74	6	3
CHECK	88	5	1	OFFICER	101	7	3
COFFEE	78	6	2	OPINION	96	7	3
COMMAND	72	7	2	PATTERN	113	7	2
COMPLEX	91	7	2	PERCENT	53	7	2
CONCERN	98	7	2	PICTURE	162	7	3
CONGRESS	152	8	2	PLANT	125	5	1
CONSTANT	71	8	2	PLEASURE	62	8	2

Word	KF	Letter	Syll.	Word	KF	Letter	Syll.
CONSTRUCT	93	9	2	PLENTY	55	6	2
CONTORT	69	7	2	PROJECT	93	7	2
CONTRACT	60	8	2	PROPERTY	156	8	3
CORNET	76	6	2	RAILROAD	58	8	2
COUPLE	122	6	2	REACTION	124	8	3
COVER	88	5	2	RELIEF	66	6	2
CREDIT	64	6	2	RESPECT	125	7	2
DANCE	90	5	1	RIVER	165	5	2
DANGER	70	6	2	SCALE	60	5	1
DEAFEN	70	6	2	SCAMPER	75	7	2
DECISION	119	8	3	SEASON	105	6	2
DEGREE	125	6	2	SECRET	78	6	2
DESIGN	114	6	2	SERIES	130	6	2
DESIRE	79	6	3	SESSION	80	7	2
DETAIL	72	6	2	SHAPE	85	5	1
DINNER	91	6	2	SLEEP	65	5	1
DISTANCE	108	8	2	SPACE	184	5	1
DISTRICT	135	8	2	SPIRIT	182	6	2
DIVISION	107	8	3	STAPLER	76	7	3
EFFORT	145	6	2	STATION	105	7	2
ENTRENCH	63	8	2	STATUS	97	6	2
EVENING	133	7	3	STEEPLE	86	7	2
EXTENT	110	6	2	STORY	153	5	2
FACULTY	74	7	3	STYLE	98	5	2
FAILURE	89	7	2	SUCCESS	93	7	2
FOREST	66	6	2	SUPPLY	102	6	2

Word	KF	Letter	Syll.	Word	KF	Letter	Syll.
FORWARD	115	7	2	SUPREME	51	7	2
FRAME	74	5	1	SURFACE	200	7	2
GOVERNOR	83	8	3	TASTE	59	5	1
HALLOW	68	6	2	THEORY	129	6	2
HOTEL	126	5	2	UNCLE	57	5	2
IMAGE	119	5	2	VALLEY	73	6	2
INDUSTRY	171	8	3	VERSION	53	7	2
ISLAND	167	6	2	VILLAGE	72	7	2
KITCHEN	90	7	2	VOLUME	135	6	2
KNIFE	76	5	1	WARREN	51	6	2
LANGUAGE	109	8	3	WEAPONS	61	7	2
LEAGUE	69	6	1	WEIGHT	91	6	1
LETTERS	115	7	2	WINDOW	119	6	2

Note. KF = Kucera and Francis (1967) word frequency. Syll. = Syllables

Table A.6

Buffer and Filler Words from Experiment 3.

Word	KF	Letter	Syll.	Word	KF	Letter	Syll.
ANSWER	152	6	2	PARENTS	91	7	2
COUSIN	51	6	2	PRICE	108	5	1
DEMAND	102	6	2	PURPOSES	90	8	3
DIRECTOR	101	8	3	SCROLL		6	1
ENEMY	88	5	3	STAFF	113	5	1
ENERGY	100	6	3	STUDENT	131	7	2
FRIEND	133	6	1	SYSTEMS	129	7	2
GAMES	55	5	1	UNIVERSE	71	8	3
GROWTH	155	6	1	WRITER	73	6	2
LEADER	74	6	2				

Note. KF = Kucera and Francis (1967) word frequency. Syll. = Syllables

Appendix B

Hits and False Alarms for Experiments 1 – 5B

Table B.1

Means and Standard Deviations for Critical Recognition Hit and False Alarm Rates in Experiment 1 as a Function of Run Condition.

Run Condition	Hit		FA	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
1-trial	0.75	0.14	0.22	0.14
2-trial	0.73	0.14	0.26	0.14
4-trial	0.69	0.12	0.26	0.12
Control	0.75	0.15	0.26	0.16

Note. Critical recognition decisions in the 1-, 2-, and 4-trial run conditions are switch trials. Critical recognition decisions in the control run condition is a repetition trial. FA = false alarm

Table B.2

Means and Standard Deviations for Recognition Hit and False Alarm Rates in Experiment 2 as a Function of Test Block and Transition.

	Test Block			
	Test 1		Test 2	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Hits (deep)				
Repetition	0.81	0.17	0.88	0.15
Switch	0.81	0.16	0.82	0.17
Hits (shallow)				
Repetition	0.67	0.22	0.78	0.19
Switch	0.67	0.22	0.71	0.22
False Alarms				
Repetition	0.22	0.15	0.21	0.15
Switch	0.30	0.21	0.22	0.14

Note. Hit rates are reported as a function of level of processing at study.

Table B.3

Means and Standard Deviations for Recognition Hit and False Alarm Rates in Experiment 3 as a Function of Run Position and Test Block.

	Run Position							
	1		2		3		4	
	M	SD	M	SD	M	SD	M	SD
Test 1								
Hit	0.69	0.23	0.79	0.17	0.78	0.16	0.79	0.16
FA	0.26	0.21	0.32	0.22	0.29	0.20	0.27	0.19
Test 2								
Hit	0.73	0.21	0.82	0.16	0.83	0.17	0.78	0.16
FA	0.31	0.25	0.36	0.22	0.32	0.23	0.30	0.20

Note. Hit and false alarm rates are collapsed across the three run conditions used in Experiment 3. FA = false alarm.

Table B.4

Means and Standard Deviations for Recognition Hit and False Alarm Rates in Experiment 4A as a Function of RSI, Transition, and Test Block.

		Test Block			
		Test 1		Test 2	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Long (600 ms)					
Repetition	Hit	0.92	0.08	0.96	0.05
	FA	0.18	0.14	0.15	0.12
Switch	Hit	0.87	0.13	0.88	0.10
	FA	0.25	0.19	0.16	0.13
Short (200 ms)					
Repetition	Hit	0.93	0.09	0.94	0.07
	FA	0.19	0.12	0.13	0.12
Switch	Hit	0.90	0.08	0.90	0.08
	FA	0.20	0.11	0.20	0.18

Note. RSI = response-to-stimulus interval. Within-participant comparisons should be made diagonally from Test 1 to Test 2 (i.e., Test 1 Long RSI, Test 2 Short RSI).

Table B.5

Means and Standard Deviations for Recognition Hit and False Alarm Rates in Experiment 4B as a Function of RSI, Transition, and Test Block.

		Test Block			
		Test 1		Test 2	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Long (1200 ms)					
Repetition	Hit	0.88	0.09	0.90	0.08
	FA	0.18	0.16	0.15	0.09
Switch	Hit	0.85	0.13	0.90	0.09
	FA	0.21	0.15	0.18	0.10
Short (600 ms)					
Repetition	Hit	0.86	0.17	0.95	0.06
	FA	0.17	0.14	0.21	0.14
Switch	Hit	0.84	0.20	0.87	0.13
	FA	0.20	0.17	0.14	0.10

Note. RSI = response-to-stimulus interval. Within-participant comparisons should be made diagonally from Test 1 to Test 2 (i.e., Test 1 Long RSI, Test 2 Short RSI).

Table B.6

Means and Standard Deviations for Recognition Hit and False Alarm Rates in Experiment 5A as a Function of RSI, Transition, and Test Block.

		Test Block			
		Test 1		Test 2	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Long (2600 ms)					
Repetition	Hit	0.85	0.13	0.81	0.16
	FA	0.12	0.10	0.20	0.14
Switch	Hit	0.85	0.12	0.85	0.12
	FA	0.16	0.12	0.21	0.22
Short (200 ms)					
Repetition	Hit	0.88	0.10	0.91	0.09
	FA	0.15	0.14	0.11	0.10
Switch	Hit	0.85	0.12	0.92	0.08
	FA	0.19	0.14	0.18	0.13

Note. RSI = response-to-stimulus interval. Within-participant comparisons should be made diagonally from Test 1 to Test 2 (i.e., Test 1 Long RSI, Test 2 Short RSI).

Table B.7

Means and Standard Deviations for Recognition Hit and False Alarm Rates in Experiment 5B as a Function of RSI, Transition, and Test Block.

		Test Block			
		Test 1		Test 2	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Long (2600 ms)					
Repetition	Hit	0.87	0.12	0.84	0.12
	FA	0.18	0.12	0.17	0.12
Switch	Hit	0.84	0.12	0.82	0.14
	FA	0.18	0.18	0.18	0.14
Short (200 ms)					
Repetition	Hit	0.88	0.13	0.92	0.11
	FA	0.23	0.17	0.20	0.17
Switch	Hit	0.85	0.13	0.86	0.15
	FA	0.25	0.24	0.19	0.18

Note. RSI = response-to-stimulus interval. Within-participant comparisons should be made diagonally from Test 1 to Test 2 (i.e., Test 1 Long RSI, Test 2 Short RSI).

Appendix C

Signal Detection Analyses for Experiment 2

Measures of discrimination and bias were calculated to provide further evidence for the impact of task switching on recognition memory processes. The measure A' was selected over d' because it does not rely on equal variance of old and new distributions and is capable of calculating a value for A' when hit and/or false alarm rates are 0 and 1. This experiment used See, Warm, Dember and Howe's (1997) equations to calculate A' and B''_D values for individual participants. The current analysis focused on the A' data.¹⁸ As a reminder, A' is a measure of participants' ability to discriminate between old and new words—it is a measure of sensitivity.

The mean A' values are reported in Table C.1 for each block, level of processing (LOP), and transition. All values are considerably above chance (0.5), indicating that participants were relatively good at discriminating between old and new words. Consistent with the corrected accuracy data, the mean A' for switch trials were smaller

Table C.1

Means and Standard Deviations for A' in Experiment 2 as a Function of Test Block, Level of Processing at Study, and Task Transition.

	Test Block			
	Test 1		Test 2	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Deep				
Repetition	0.86	0.12	0.90	0.11
Switch	0.82	0.16	0.86	0.14
Shallow				
Repetition	0.78	0.19	0.86	0.13
Switch	0.74	0.24	0.83	0.09

¹⁸ For the sake of completeness, B''_D data indicated a slight bias towards responding 'old' which was not affected by task transition ($F = 0.05$, $p = 0.821$).

than that for repetition trials across all test block and LOP conditions. With regard to the LOP manipulation, the mean A' was larger for deeply encoded words (living or non-living?) than that for shallowly encoded words ('e' or no 'e?'). Finally, A' data showed similar overall improvement from the first to second test. Participants' ability to discriminate between old and new words improved with experience.

A 2 (test block) x 2 (level of processing) x 2 (transition) repeated measures ANOVA was conducted on A' scores replicated the findings of the corrected accuracy analyses. Critically, although main effects were observed for level of processing [$F(1,31) = 16.74$, $MSE = 0.01$, $p < 0.001$, $\eta_p^2 = 0.35$], and task transition [$F(1,31) = 6.06$, $MSE = 0.01$, $p = 0.020$, $\eta_p^2 = 0.16$], the interaction of these manipulations was not significant ($F = 0.02$, $p = 0.88$). Consistent with the corrected accuracy data, analyses of A' did not support the hypothesis that the level of processing effect would be reduced or eliminated when switching tasks. The significant main effect of test block [$F(1,31) = 9.48$, $MSE = 0.03$, $p = 0.004$, $\eta_p^2 = 0.23$] and non-significant interaction of block and transition ($F = 0.17$, $p = 0.681$) were also consistent with corrected accuracy analyses.

Appendix D

Supplementary Tables for Experiments 4A – 5B

Table D.1

Means and Standard Deviations for Correct RTs in Experiment 4A as a Function of Test Block, Transition, RSI, Task and Decision.

		Test 1				Test 2			
		Repetition		Switch		Repetition		Switch	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Long (600 ms)									
Recognition	New	381	91	492	154	336	59	410	84
	Old	329	95	455	176	270	44	363	95
Magnitude	More	740	123	799	182	638	77	701	87
	Less	712	96	774	118	669	59	721	94
Short (200 ms)									
Recognition	New	390	66	509	130	361	70	478	127
	Old	327	69	444	139	310	76	392	108
Magnitude	More	734	102	775	102	703	97	740	125
	Less	768	119	842	155	706	85	797	110

Note. RSI = response-to-stimulus interval. Recognition RTs do not include the 400-ms RSD. Within-participant comparisons should be made diagonally from Test 1 to Test 2 (i.e., Test 1 Long RSI, Test 2 Short RSI).

Table D.2

Means and Standard Deviations for Correct RTs in Experiment 4B as a Function of Test Block, Transition, RSI, Task and Decision.

		Test 1				Test 2			
		Repetition		Switch		Repetition		Switch	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Long (1300 ms)									
Recognition	New	405	145	491	183	384	98	458	127
	Old	353	135	482	226	342	119	423	123
Magnitude	More	773	228	850	257	669	127	739	152
	Less	840	230	856	228	730	91	726	120
Short (600 ms)									
Recognition	New	395	84	495	110	360	104	431	132
	Old	340	87	438	139	291	132	350	107
Magnitude	More	743	98	763	209	680	137	687	114
	Less	778	98	830	161	725	170	786	160

Note. RSI = response-to-stimulus interval. Recognition RTs do not include the 400-ms RSD. Within-participant comparisons should be made diagonally from Test 1 to Test 2 (i.e., Test 1 Long RSI, Test 2 Short RSI).

Table D.3

Means and Standard Deviations for Correct RTs in Experiment 5A as a Function of Test Block, Transition, RSI, Task and Decision.

		Test 1				Test 2			
		Repetition		Switch		Repetition		Switch	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Long (2600 ms)									
Recognition	New	484	146	619	177	454	103	521	152
	Old	447	153	599	210	439	127	522	137
Dot	More	807	142	878	184	814	222	873	228
	Less	824	195	838	160	842	169	886	256
Short (200 ms)									
Recognition	New	479	141	635	168	354	91	470	174
	Old	412	148	579	214	292	90	387	137
Dot	More	834	180	883	175	695	84	701	108
	Less	874	244	975	277	685	75	739	97

Note. RSI = response-to-stimulus interval. Recognition RTs do not include the 400-ms RSD. Within-participant comparisons should be made diagonally from Test 1 to Test 2 (i.e., Test 1 Long RSI, Test 2 Short RSI).

Table D.4

Means and Standard Deviations for Correct RTs in Experiment 5B as a Function of Test Block, Transition, RSI, Task and Decision.

		Test 1				Test 2			
		Repetition		Switch		Repetition		Switch	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Long (2600 ms)									
Recognition	New	473	152	591	196	488	118	553	130
	Old	440	133	570	201	436	101	573	151
Magnitude	More	760	149	896	224	702	132	749	107
	Less	854	178	904	206	813	134	811	120
Short (200 ms)									
Recognition	New	470	137	616	256	363	94	484	154
	Old	382	123	485	175	326	97	424	142
Magnitude	More	759	121	806	132	752	163	748	142
	Less	840	127	933	155	724	118	794	153

Note. RSI = response-to-stimulus interval. Recognition RTs do not include the 400-ms RSD. Within-participant comparisons should be made diagonally from Test 1 to Test 2 (i.e., Test 1 Long RSI, Test 2 Short RSI).

References

- Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà, & M. Moscovitch (Eds.), *Conscious and nonconscious information processing: Attention and Performance XV* (pp. 421-452). Cambridge, MA, US: The MIT Press.
- Allport, D. A., & Wylie, G. (2000). Task-switching, stimulus-response bindings, and negative priming. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 35-70). Cambridge, MA: MIT Press.
- Altmann, E. M. (2002). Functional decay of memory for tasks. *Psychological Research/Psychologische Forschung*, 66(4), 287-297. doi:10.1007/s00426-002-0102-9
- Altmann, E. M. (2004a). Advance preparation in task switching: What work is being done? *Psychological Science*, 15(9), 578-582. doi:10.1111/j.0956-7976.2004.00723.x
- Altmann, E. M. (2004b). The preparation effect in task switching: Carryover of SOA. *Memory & Cognition*, 32(1), 153-163. doi:10.3758/BF03195828
- Altmann, E. M. (2006). Task switching is not cue switching. *Psychonomic Bulletin & Review*, 13(6), 1016-1022. doi:10.3758/BF03213918
- Altmann, E. M. (2007a). Comparing switch costs: Alternating runs and explicit cuing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(3), 475-483. doi:10.1037/0278-7393.33.3.475
- Altmann, E. M., & Gray, W. D. (2002). Forgetting to remember: The functional relationship of decay and interference. *Psychological Science*, 13(1), 27-33. doi:10.1111/1467-9280.00405
- Altmann, E. M., & Gray, W. D. (2008). An integrated model of cognitive control in task switching. *Psychological Review*, 115(3), 602-639. doi:10.1037/0033-295X.115.3.602
- Altmann, E. M., & Trafton, J. G. (2002). Memory for goals: An activation-based model. *Cognitive Science: A Multidisciplinary Journal*, 26(1), 39-83. doi:10.1016/S0364-0213(01)00058-1

- Altmann, E. M., & Trafton, J. G. (2007). Timecourse of recovery from task interruption: Data and a model. *Psychonomic Bulletin & Review*, 14(6), 1079-1084. doi:10.3758/BF03193094
- Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American Psychologist*, 51(4), 355-365. doi:10.1037/0003-066X.51.4.355
- Arbuthnott, K., & Frank, J. (2000). Executive control in set switching: Residual switch cost and task-set inhibition. *Canadian Journal of Experimental Psychology/Revue Canadienne De Psychologie Expérimentale*, 54(1), 33-41. doi:10.1037/h0087328
- Arrington, C. M., & Logan, G. D. (2004). The cost a voluntary task switch. *Psychological Science*, 15(9), 610-615. doi:10.1111/j.0956-7976.2004.00728.x
- Arrington, C. M., & Logan, G. D. (2005). Voluntary task switching: Chasing the elusive homunculus. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(4), 683-702. doi:10.1037/0278-7393.31.4.683
- Arrington, C. M., Weaver, S. M., & Pauker, R. L. (2010). Stimulus-based priming of task choice during voluntary task switching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(4), 1060-1067. doi:10.1037/a0019646
- Atkinson, R. C., & Juola, J. F. (1973). Factors influencing speed and accuracy of word recognition. In S. Kornblum (Ed.), *Forth international symposium on attention and performance* (pp. 583-611). New York: Academic Press.
- Atkinson, R. C., & Juola, J. F. (1974). Search and decision processes in recognition memory. In *Contemporary developments in mathematical psychology: I. learning, memory and thinking*. (pp. 239-289). Oxford, England: W. H. Freeman.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory Vol 2*. (pp. 89-195). New York: Academic Press
- Baddeley, A., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: Evidence from task switching. *Journal of Experimental Psychology: General*, 130(4), 641-657. doi:10.1037/0096-3445.130.4.641

- Balota, D.A., Yap, M.J., Cortese, M.J., Hutchison, K.A., Kessler, B., Loftis, B., Neely, J.H., Nelson, D.L., Simpson, G.B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, 39, 445-459. doi:10.3758/BF03193014
- Banks, W. P. (1970). Signal detection theory and human memory. *Psychological Bulletin*, 74(2), 81-99. doi:10.1037/h0029531
- Boldini, A., Russo, R., & Avons, S. E. (2004). One process is not enough! A speed-accuracy tradeoff study of recognition memory. *Psychonomic Bulletin & Review*, 11(2), 353-361. doi:10.3758/BF03196582
- Brown, S., Lehmann, C., & Poboka, D. (2006). A critical test of the failure-to-engage theory of task switching. *Psychonomic Bulletin & Review*, 13(1), 152-159. doi:10.3758/BF03193827
- Craik, F. I. M. (2002). Levels of processing: Past, present . . . and future? *Memory*, 10(5-6), 305-318. doi:10.1080/09658210244000135
- Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, 125(2), 159-180. doi:10.1037/0096-3445.125.2.159
- Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior*, 11(6), 671-684. doi:10.1016/S0022-5371(72)80001-X
- Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104(3), 268-294. doi:10.1037/0096-3445.104.3.268
- Curran, T., DeBuse, C., & Leynes, P. A. (2007). Conflict and criterion setting in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(1), 2-17. doi:10.1037/0278-7393.33.1.2
- 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). Cambridge, MA: MIT Press.
- de Jong, R. (2001). Adult age difference in goal activation and goal maintenance. *European Journal of Cognitive Psychology*, 13(1-2), 71-89. doi:10.1080/09541440042000223
- Demanet, J., Verbruggen, F., Liefvooghe, B., & Vandierendonck, A. (2010). Voluntary task switching under load: Contributions of top-down and bottom-up factors in

- goal-directed behaviour. *Psychonomic Bulletin & Review*, 17(3), 387-393. doi: 10.3758/PBR.17.3.387
- Diana, R. A., Reder, L. M., Arndt, J., & Park, H. (2006). Models of recognition: A review of arguments in favor of a dual-process account. *Psychonomic Bulletin & Review*, 13(1), 1-21. doi:10.3758/BF03193807
- Dodson, C. S., Holland, P. W., & Shimamura, A. P. (1998). On the recollection of specific- and partial-source information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(5), 1121-1136. doi:10.1037/0278-7393.24.5.1121
- Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory & Cognition*, 24(4), 523-533. doi:10.3758/BF03200940
- Driesbach, B. & Haider, H. (2008). That's what task sets are for: Shielding against irrelevant information. *Psychological Research/Psychologische Forschung* 72(4), 355-361. doi: 10.1007/s00426-007-0131-5
- Dunn, J. C. (2004). Remember-know: A matter of confidence. *Psychological Review*, 111(2), 524-542. doi:10.1037/0033-295X.111.2.524
- Edwards, M. B., & Gronlund, S. D. (1998). Task interruption and its effects on memory. *Memory*, 6(6), 665-687. doi:10.1080/741943375
- Emerson, M. J., & Miyake, A. (2003). The role of inner speech in task switching: A dual-task investigation. *Journal of Memory and Language*, 48(1), 148-168. doi:10.1016/S0749-596X(02)00511-9
- Fernandes, M. A., & Moscovitch, M. (2000). Divided attention and memory: Evidence of substantial interference effects at retrieval and encoding. *Journal of Experimental Psychology: General*, 129(2), 155-176. doi:10.1037/0096-3445.129.2.155
- Field, A. (2009) *Discovering Statistics Using SPSS* (3rd Edition). London, England: Sage.
- Forstmann, B. U., Brass, M., Koch, I., & von Cramon, D. Y. (2006). Voluntary selection of task sets revealed by functional magnetic resonance imaging. *Journal of Cognitive Neuroscience*, 18(3), 388-398. doi:10.1162/jocn.2006.18.3.388
- Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). The Toronto word pool: Norms for imagery, concreteness, orthographic variables, and

- grammatical usage for 1,080 words. *Behavior Research Methods*, 14 (4), 375-399. doi: 10.3758/BF03203275
- Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition*, 16(4), 309-313. doi:10.3758/BF03197041
- Glisky, E. L., Polster, M. R., & Routhieaux, B. C. (1995) Double dissociation between item and source memory. *Neuropsychology*, 9(2), 229-235. doi:10.1037/0894-4105.9.2.229
- Gillie, T., & Broadbent, D. E. (1989). What makes interruptions disruptive? A study of length, similarity, and complexity. *Psychological Research*, 50(4), 243-250. doi:10.1007/BF00309260
- Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, 91(1), 1-67. doi:10.1037/0033-295X.91.1.1
- Ginsburg, N. (1978). Perceived numerosity, item arrangement, and expectancy. *The American Journal of Psychology*, 91(2), 267-273. doi:10.2307/1421536
- Gopher, D., Armony, L., & Greenshpan, Y. (2000). Switching tasks and attention policies. *Journal of Experimental Psychology: General*, 129(3), 308-339. doi:10.1037/0096-3445.129.3.308
- Green, D. M., & Swets, J. A. (1974). *Signal detection theory and psychophysics*. Oxford, England: Robert E. Krieger.
- He, L., Zhang, J., Zhou, T., & Chen, L. (2009). Connectedness affects dot numerosity judgment: Implications for configural processing. *Psychonomic Bulletin & Review*, 16(3), 509-517. doi:10.3758/PBR.16.3.509
- Heathcote, A., Ditton, E., & Mitchell, K. (2006). Word frequency and word likeness mirror effects in episodic recognition memory. *Memory & Cognition*, 34(4), 826-838. doi:10.3758/BF03193430
- Heathcote, A., Raymond, F., & Dunn, J. (2006). Recollection and familiarity in recognition memory: Evidence from ROC curves. *Journal of Memory and Language*, 55(4), 495-514. doi:10.1016/j.jml.2006.07.001
- Hicks, J. L., & Marsh, R. L. (2000). Toward specifying the attentional demands of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(6), 1483-1498. doi:10.1037/0278-7393.26.6.1483
- Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, 95(4), 528-551. doi:10.1037/0033-295X.95.4.528

- Hintzman, D. L., & Curran, T. (1994). Retrieval dynamics of recognition and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory and Language*, 33(1), 1-18. doi:10.1006/jmla.1994.1001
- Hockley, W. E. (1992). Item versus associative information: Further comparisons of forgetting rates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(6), 1321-1330. doi:10.1037/0278-7393.18.6.1321
- Hockley, W. E., & Consoli, A. (1999). Familiarity and recollection in item and associative recognition. *Memory & Cognition*, 27(4), 657-664. doi:10.3758/BF03211559
- Hogan, M. J., Kelly, C. A. M., & Craik, F. I. M. (2006). The effects of attention switching on encoding and retrieval of words in younger and older adults. *Experimental Aging Research*, 32(2), 153-183. doi:10.1080/03610730600553935
- Hunt, A. R., & Klein, R. M. (2002). Eliminating the cost of task set reconfiguration. *Memory & Cognition*, 30(4), 529-539. doi:10.3758/BF03194954
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, 30(5), 513-541. doi:10.1016/0749-596X(91)90025-F
- Jacoby, L. L. (1998). Invariance in automatic influences of memory: Toward a user's guide for the process-dissociation procedure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(1), 3-26. doi: 10.1037/0278-7393.24.3.1.3
- Jacoby, L. L. (1999). Ironic effects of repetition: Measuring age-related differences in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(1), 3-22. doi:10.1037/0278-7393.25.1.3
- Jacoby, L. L., Debner, J. A., & Hay, J. F. (2001). Proactive interference, accessibility bias, and process dissociations: Valid subject reports of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(3), 686-700. doi:10.1037/0278-7393.27.3.686
- James, W. (1890). *The Principles of Psychology, Volume One*. New York, NY: Dover Publications.
- Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology*, 14 (Whole No. 89).
- Jones, T. C. (2005). Study repetition and the rejection of conjunction lures. *Memory*, 13(5), 499-515. doi:10.1080/09658210444000197

- Jones, T. C., & Atchley, P. (2006). Conjunction errors, recollection-based rejections, and forgetting in a continuous recognition task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(1), 70-70. doi:10.1037/0278-7393.32.1.70
- Jones, T. C., & Bartlett, J. C. (2009) When false recognition is out of control: The case of facial conjunctions. *Memory & Cognition*, 37(2). 143-157. doi: 10.3758/MC.37.2.143
- Jones, T. C., Bartlett, J. C., & Wade, K. A. (2006). Nonverbal conjunction errors in recognition memory: Support for familiarity but not for feature bundling. *Journal of Memory and Language*, 55(1). doi: 10.1016/j.jml.2006.01.002
- Jones, T. C., Burson, R. A., & Aronsen, D. (2011, April). *Building a bridge from interruptions to task switching*. Paper presented at the 38th meeting of the Australasian Experimental Psychology Conference, Auckland.
- Jones, T. C., & Jacoby, L. L. (2001). Feature and conjunction errors in recognition memory: Evidence for dual-process theory. *Journal of Memory and Language*, 45(1), 82-102. doi:10.1006/jmla.2000.2761
- Joordens, S., & Hockley, W. E. (2000). Recollection and familiarity through the looking glass: When old does not mirror new. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(6), 1534-1555. doi:10.1037/0278-7393.26.6.1534
- Kavcic, V., Krar, F. J., & Doty, R. W. (1999). Temporal cost of switching between kinds of visual stimuli in a memory task. *Cognitive Brain Research*, 9(2), 199-203. doi:10.1016/S0926-6410(99)00059-2
- Keele, S. W., & Rafal, R. (2000). Deficits of task set in patients with left prefrontal cortex lesions. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 627-651). Cambridge, MA: MIT Press.
- Kieffaber, P. D., & Hetrick, W. P. (2005). Event-related potential correlates of task switching and switch costs. *Psychophysiology*, 42(1), 56-71. doi:10.1111/j.1469-8986.2005.00262.x
- Klein, K. A., Shiffrin, R. M., & Criss, A. H. (2007). Putting context in context. In J. S. Nairne (Ed.), *The foundations of remembering: Essays in honor of Henry L. Roediger, III*. (pp. 171-189). New York, NY, US: Psychology Press.

- Koch, I. (2001). Automatic and Intentional Activation of Task Sets. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(6), 1474-1486. doi: 10.1037//0278-7393.27.6.1474
- Koch, I. (2003). The role of external cues for endogenous advance reconfiguration in task switching. *Psychonomic Bulletin & Review*, 10(2), 488-492. doi:10.3758/BF03196511
- Koch, I. (2005). Sequential task predictability in task switching. *Psychonomic Bulletin & Review*, 12(1), 107-112. doi:10.3758/BF03196354
- Koch, I., & Allport, A. (2006). Cue-based preparation and stimulus-based priming of tasks in task switching. *Memory & Cognition*, 34(2), 433-444. doi:10.3758/BF03193420
- Knott, L. M., & Dewhurst, S. A. (2007a). The effects of divided attention at encoding and retrieval on false recognition: A comparison of DRM and categorized lists. *Memory & Cognition*, 35(8), 1954-1965. doi:10.3758/BF03192928
- Knott, L. M., & Dewhurst, S. A. (2007b). Divided attention at retrieval disrupts knowing but not remembering. *Memory*, 15, 664-674. doi:10.1080/09658210701467137
- Kramer, A. F., Hahn, S., & Gopher, D. (1999). Task coordination and aging: Explorations of executive control processes in the task switching paradigm. *Acta Psychologica*, 101(2-3), 339-378. doi:10.1016/S0001-6918(99)00011-6
- Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging*, 15(1), 126-147. doi:10.1037/0882-7974.15.1.126
- Kucera, H. & Francis, W.N. (1967). *Computational Analysis of Present-day American English*. Providence, RI: Brown University Press.
- Liefooghe, B., Demanet, J., & Vandierendonck, A. (2009). Is advance reconfiguration in voluntary task switching affected by the design employed? *The Quarterly Journal of Experimental Psychology*, 62(5), 850-857. doi:10.1080/17470210802570994
- Lindsen, J. P., & de Jong, R. (2010). Distinguishing between the partial-mapping preparation hypothesis and the failure-to-engage hypothesis of residual switch costs. *Journal of Experimental Psychology: Human Perception and Performance*, 36(5), 1207-1226. doi:10.1037/a0020362
- Lockhart, R. S., & Craik, F. I. (1990). Levels of processing: A retrospective commentary on a framework for memory research. *Canadian Journal of*

- Psychology/Revue Canadienne De Psychologie*, 44(1), 87-112.
doi:10.1037/h0084237
- Logan, G. D. (2003). Executive control of thought and action: In search of the wild homunculus. *Current Directions in Psychological Science*, 12(2), 45-48. doi: 10.1111/1467-8721.01223
- 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.
doi: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. doi:10.3758/BF03196709
- Logan, G. D., & Schneider, D. W. (2010). Distinguishing reconfiguration and compound-cue retrieval in task switching. *Psychologica Belgica*, 50(3-4), 413-433.
- Logan, G. D., Schneider, D. W., & Bundesen, C. (2007). Still clever after all these years: Searching for the homunculus in explicitly cued task switching. *Journal of Experimental Psychology: Human Perception and Performance*, 33(4), 978-994. doi:10.1037/0096-1523.33.4.978
- Malmberg, K. J. (2008). Recognition memory: A review of the critical findings and an integrated theory for relating them. *Cognitive Psychology*, 57(4), 335-384.
doi:10.1016/j.cogpsych.2008.02.004
- 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. (2007). Inhibition of task sets. In D. S. Gorfein, & C. M. MacLeod (Eds.), *Inhibition in cognition*. (pp. 27-44). Washington, DC, US: American Psychological Association. doi:10.1037/11587-002
- Mayr, U., & Bell, T. (2006). On how to be unpredictable: Evidence from the voluntary task-switching paradigm. *Psychological Science*, 17(9), 774-780.
doi:10.1111/j.1467-9280.2006.01781.x
- 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. doi: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
- McClelland, J. L., & Chappell, M. (1998). Familiarity breeds differentiation: A subjective-likelihood approach to the effects of experience in recognition memory. *Psychological Review*, 105(4), 724-760. doi:10.1037/0033-295X.105.4.734-760
- Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(6), 1423-1442. doi:10.1037/0278-7393.22.6.1423
- Meiran, N. (2000a). Reconfiguration of Stimulus Task Sets and Response Task Sets during Task Switching. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 377-399). Cambridge, MA: MIT Press.
- Meiran, N. (2000b). Modeling cognitive control in task-switching. *Psychological Research/Psychologische Forschung*, 63(3-4), 234-249. doi:10.1007/s004269900004
- Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. *Cognitive Psychology*, 41(3), 211-253. doi:10.1006/cogp.2000.0736
- Minear, M., & Shah, P. (2008). Training and transfer effects in task switching. *Memory & Cognition*, 36(8), 1470-1483. doi:10.3758/MC.36.8.1470
- Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J. (2004). Inner speech as a retrieval aid for task goals: The effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica*, 115(2-3), 123-142. doi:10.1016/j.actpsy.2003.12.004
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134-134. doi:10.1016/S1364-6613(03)
- 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., Sumner, P., & Waters, H. (2003). Task-set reconfiguration with predictable and unpredictable task switches. *Memory & Cognition*, 31(3), 327-342. doi:10.3758/BF03194391
- Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Behavior*, 16(5), 519-533. doi:10.1016/S0022-5371(77)80016-9
- Murdock, B. B. (1997). Context and mediators in a theory of distributed associative memory (TODAM2). *Psychological Review*, 104(4), 839-862. doi:10.1037/0033-295X.104.4.839
- Murdock, B. (2006). Decision-making models of remember-know judgments: Comment on Rotello, Macmillan, and Reeder (2004). *Psychological Review*, 113(3), 648-655. doi:10.1037/0033-295X.113.3.648
- Naveh-Benjamin, M., Craik, F. I. M., Gavrilescu, D., & Anderson, N. D. (2000). Asymmetry between encoding and retrieval processes: Evidence from divided attention and a calibration analysis. *Memory & Cognition*, 28(6), 965-976. doi:10.3758/BF03209344
- Nieuwenhuis, S., & Monsell, S. (2002). Residual costs in task switching: Testing the failure-to-engage hypothesis. *Psychonomic Bulletin & Review*, 9(1), 86-92. doi:10.3758/BF03196259
- Odegard, T. N., Koen, J. D., & Gama, J. M. (2008). Process demands of rejection mechanisms of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(5), 1296-1304. doi:10.1037/a0013034
- Park, H., Reder, L. M., & Dickison, D. (2005). The effects of word frequency and similarity on recognition judgments: The role of recollection. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(3), 568-578. doi:10.1037/0278-7393.31.3.568
- Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, 116(2), 220-244. doi:10.1037/0033-2909.116.2.220
- Paulitzki, J. R., Risko, E. F., Oakman, J. M., & Stolz, J. A. (2008). Doing the unpleasant: How the emotional nature of a threat-relevant task affects task-switching. *Personality and Individual Differences*, 45(5), 350-355. doi:10.1016/j.paid.2008.05.003

- Philipp, A. M., & Koch, I. (2006). Task inhibition and task repetition in task switching. *European Journal of Cognitive Psychology.Special Issue: Integrative Views on Dual-Task Costs*, 18(4), 624-639. doi:10.1080/09541440500423269
- Poljac, E., Koch, I., & Bekkering, H. (2009). Dissociating restart cost and mixing cost in task switching. *Psychological Research/Psychologische Forschung*, 73(3), 407-416. doi:10.1007/s00426-008-0151-9
- Pollack, I. (1959). Identification of elementary auditory displays and the method of recognition memory. *Journal of the Acoustical Society of America*, 31(8), 1126-1128.
- Pösse, B., Waszak, F., & Hommel, B. (2006). Do stimulus-response bindings survive a task switch? *European Journal of Cognitive Psychology.Special Issue: Integrative Views on Dual-Task Costs*, 18(4), 640-651. doi:10.1080/09541440500423285
- Postman, L., & Kruesi, E. (1977). The influence of orienting tasks on the encoding and recall of words. *Journal of Verbal Learning & Verbal Behavior*, 16(3), 353-369. doi:10.1016/S0022-5371(77)80056-X
- Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition*, 21(1), 89-102. doi:10.3758/BF03211168
- Ratcliff, R., & Starns, J. J. (2009). Modeling confidence and response time in recognition memory. *Psychological Review*, 116(1), 59-83. doi:10.1037/a0014086
- Reder, L. M., Nhonyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember-know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(2), 294-320. doi:10.1037/0278-7393.26.2.294
- Roediger, H. L., Gallo, D. A., & Geraci, L. (2002). Processing approaches to cognition: The impetus from the levels-of-processing framework. *Memory*, 10(5-6), 319-332. doi:10.1080/09658210224000144
- Rotello, C. M., & Heit, E. (1999). Two-process models of recognition memory: Evidence for recall-to-reject? *Journal of Memory and Language*, 40(3), 432-453. doi:10.1006/jmla.1998.2623

- Rotello, C. M., Macmillan, N. A., & Van Tassel, G. (2000). Recall-to-reject in recognition: Evidence from ROC curves. *Journal of Memory and Language*, 43(1), 67-88. doi:10.1006/jmla.1999.2701
- 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
- Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, 11(6), 251-257. doi:10.1016/j.tics.2007.04.004
- Ruiz, J. C., Soler, M. J., & Dasí, C. (2004). Study time effects in recognition memory. *Perceptual and Motor Skills*, 98(2), 638-642. doi:10.2466/PMS.98.2.638-642
- Rubin, O., & Meiran, N. (2005). On the origins of the task mixing cost in the cuing task-switching paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(6), 1477-1491. doi:10.1037/0278-7393.31.6.1477
- Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance*, 27(4), 763-763. doi:10.1037/0096-1523.27.4.763
- 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. doi:10.1037/0096-1523.27.6.1404
- Salthouse, T. A., Fristoe, N., McGuthry, K. E., & Hambrick, D. Z. (1998). Relation of task switching to speed, age, and fluid intelligence. *Psychology and Aging*, 13(3), 445-461. doi:10.1037/0882-7974.13.3.445
- See, J. E., Warm, J. S., Dember, W. N., & Howe, S. R. (1997). Vigilance and signal detection theory: An empirical evaluation of five measures of response bias. *Human Factors*, 39(1), 14-29. doi:10.1518/001872097778940704
- Schacter, D. L., & Tulving, E. (1994). What are the memory systems of 1994? In D. L. Schacter & E. Tulving (Eds), *Memory Systems 1994* (pp. 1-38). Massachusetts: The MIT Press.
- Schneider, D. W. (2007). Task-set inhibition in chunked task sequences. *Psychonomic Bulletin & Review*, 14(5), 970-976. doi:10.3758/BF03194130
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools Inc.

- 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. doi:10.1037/0096-3445.134.3.343
- Schneider, D. W., & Logan, G. D. (2006). Hierarchical control of cognitive processes: Switching tasks in sequences. *Journal of Experimental Psychology: General*, 135(4), 623-640. doi:10.1037/0096-3445.135.4.623
- Schneider, D. W., & Logan, G. D. (2007). Task switching versus cue switching: Using transition cuing to disentangle sequential effects in task-switching performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(2), 370-378. doi:10.1037/0278-7393.33.2.370
- Schneider, D. W., & Logan, G. D. (2011). Task-switching performance with 1:1 and 2:1 cue-task mappings: Not so different after all. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(2), 405-415. doi:10.1037/a0021967
- Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning & Verbal Behavior*, 6, 156-163. doi:10.1016/S0022-5371(67)80067-7
- Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin & Review*, 4(2), 145-166. doi:10.3758/BF03209391
- Skinner, E. I., & Fernandes, M. A. (2008). Interfering with remembering and knowing: Effects of divided attention at retrieval. *Acta Psychologica*, 127(2), 211-221. doi:10.1016/j.actpsy.2007.05.001
- Slotnick, S. D., & Dodson, C. S. (2005). Support for a continuous (single-process) model of recognition memory and source memory. *Memory & Cognition*, 33(1), 151-170. doi:10.3758/BF03195305
- Slotnick, S. D., Klein, S. A., Dodson, C. S., & Shimamura, A. P. (2000). An analysis of signal detection and threshold models of source memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(6), 1499-1517. doi:10.1037/0278-7393.26.6.1499
- Sohn, M., & Anderson, J. R. (2001). Task preparation and task repetition: Two-component model of task switching. *Journal of Experimental Psychology: General*, 130(4), 764-778. doi:10.1037/0096-3445.130.4.764

- Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, 117(1), 34-50. doi:10.1037/0096-3445.117.1.34
- Spector, A., & Biederman, I. (1976). Mental set and mental shift revisited. *American Journal of Psychology*, 89(4), 669-669.
- Sternberg, S. (1969). The discovery of processing stages: Extensions of donders' method. *Acta Psychologica*, 30, 276-315. doi:10.1016/0001-6918(69)90055-9
- Swets, J. A. (1973). The relative operating characteristic in psychology. *Science*, 182(4116), 990-1000. doi:10.1126/science.182.4116.990
- Trafton, J. G., Altmann, E. M., Brock, D. P., & Mintz, F. E. (2003). Preparing to resume an interrupted task: Effects of prospective goal encoding and retrospective rehearsal. *International Journal of Human-Computer Studies.Special Issue: Notification User Interfaces*, 58(5), 583-603. doi:10.1016/S1071-5819(03)00023-5
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie Canadienne*, 26(1), 1-12. doi:10.1037/h0080017
- Verbruggen, F., Liefoghe, B., Vandierendonck, A., & Demanet, J. (2007). Short cue presentations encourage advance task preparation: A recipe to diminish the residual switch cost. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(2), 342-356. doi:10.1037/0278-7393.33.2.342
- 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
- Wais, P. E., Wixted, J. T., Hopkins, R. O., & Squire, L. R. (2006). The hippocampus supports both the recollection and familiarity components of recognition memory. *Neuron*, 49(3), 459-466. doi: 10.1016/j.neuron.2005.12.020
- 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
- Waszak, F., Hommel, B., & Allport, A. (2004). Semantic generalization of stimulus-task bindings. *Psychonomic Bulletin & Review*, 11(6), 1027-1033. doi:10.3758/BF03196732
- Watkins, M. J. (2002). Limits and province of levels of processing: Considerations of a construct. *Memory*, 10(5-6), 339-343. doi:10.1080/09658210244000162

- Weldon, M. S., & Roediger, H. L. (1987). Altering retrieval demands reverses the picture superiority effect. *Memory & Cognition*, 15(4), 269-280.
doi:10.3758/BF03197030
- Werkle-Bergner, M., Mecklinger, A., Kray, J., Meyer, P., & Düzel, E. (2005). The control of memory retrieval: Insights from event-related potentials. *Cognitive Brain Research*, 24(3), 599-599. doi:10.1016/j.cogbrainres.2005.03.011
- Wixted, J. T. (2007a). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, 114(1), 152-176. doi:10.1037/0033-295X.114.1.152
- Wixted, J. T. (2007b). Signal-detection theory and the neuroscience of recognition memory. In J. S. Nairne (Ed.), *The foundations of remembering: Essays in honor of Henry L. Roediger, III*. (pp. 67-82). New York, NY, US: Psychology Press.
- Wixted, J. T., & Stretch, V. (2004). In defense of the signal detection interpretation of remember/know judgments. *Psychonomic Bulletin & Review*, 11(4), 616-641.
doi:10.3758/BF03196616
- Wylie, G., & Allport, A. (2000). Task switching and the measurement of "switch costs.". *Psychological Research/Psychologische Forschung*, 63(3-4), 212-233.
doi:10.1007/s004269900003
- Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(6), 1341-1354. doi:10.1037/0278-7393.20.6.1341
- Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition*, 25(6), 747-763. doi:10.3758/BF03211318
- Yonelinas, A. P. (1999). The contribution of recollection and familiarity to recognition and source-memory judgments: A formal dual-process model and an analysis of receiver operating characteristics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(6), 1415-1434. doi:10.1037/0278-7393.25.6.1415
- Yonelinas, A. P. (2001a). Components of episodic memory: The contribution of recollection and familiarity. *The Philosophical Transactions of the Royal Society, Series B*, 356 (1413), 1363-1374.

- Yonelinas, A. P. (2001b). Consciousness, control, and confidence: The 3 Cs of recognition memory. *Journal of Experimental Psychology: General*, 130(3), 361-379. doi:10.1037/0096-3445.130.3.361
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, 46(3), 441-517. doi:10.1006/jmla.2002.2864
- Yonelinas, A. P., Aly, M., Wang, W., & Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus*, 20(11), 1178-1194. doi:10.1002/hipo.20864
- Yonelinas, A. P., Dobbins, I., Szymanski, M. D., Dhaliwal, H. S., & King, L. (1996). Signal-detection, threshold, and dual-process models of recognition memory: ROCs and conscious recollection. *Consciousness and Cognition*, 5, 418-441. doi: 10.1006/ccog.1996.0026
- Yonelinas, A. P., & Jacoby, L. L. (1995). The relation between remembering and knowing as bases for recognition: Effects of size congruency. *Journal of Memory and Language*, 34(5), 622-643. doi:10.1006/jmla.1995.1028