An associative approach to task-switching.

Submitted by Charlotte Louise Forrest to the University of Exeter
as a thesis for the degree of
Doctor of Philosophy in Psychology
In July 2012

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature:…………………………………………………….

Abstract

This thesis explores the behaviour of participants taking an associative approach to a task-cueing paradigm. Task-cueing is usually intended to explore controlled processing of task-sets. But small stimulus sets plausibly afford associative learning via simple and conditional discriminations. In six experiments participants were presented with typical task-cueing trials: a cue (coloured shape) followed by a digit (or in Experiment 5 a symbol) requiring one of two responses. In the standard Tasks condition (Monsell Experiment and Experiments 1-3), the participant was instructed to perform either an odd/even or a high/low task dependent on the cue. The second condition was intended to induce associative learning of cue + stimulus-response mappings.

In general, the Tasks condition showed a large switch cost that reduced with preparation time, a small, constant congruency effect and a small perturbation when new stimuli were introduced. By contrast the CSR condition showed a small, reliable switch cost that did not reduce with preparation time, a large congruency effect that changed over time and a large perturbation when new stimuli were introduced. These differences may indicate automatic associative processing in the CSR condition and rule-based classification in the Tasks condition. Furthermore, an associative model based on the APECS learning algorithm (McLaren, 1993) provided an account of the CSR data. Experiment 3 showed that participants were able to deliberately change their approach to the experiment from using CSR instructions to using Tasks instructions, and to some extent vice versa.

Experiments 4 & 5 explored the cause of the small switch cost in the CSR condition. Consideration of the aspects of the paradigm that produced the switch cost in the APECS model produced predictions, which were tested against behavioural data. Experiment 4 found that the resulting manipulation made participants more likely to induce task-sets. Experiment 5 used random symbols instead of numbers, removing the underlying task-sets. The results of this experiment broadly agreed with the predictions made using APECS. Chapter 6 considers an initial attempt to create a real-time version of APECS. It also finds that an associative model of a different class (AMAN, Harris & Livesey, 2010) can provide an account of some, but not all, of the phenomena found in the CSR condition.

This thesis concludes that performance in the Tasks condition is suggestive of the use of cognitive control processes, whilst associatively based responding is *available* as a basis for performance in the CSR condition.

Acknowledgements

Firstly, I would like to acknowledge the contributions of my supervisors Ian McLaren and Stephen Monsell. I would like to thank Ian for his undying enthusiasm, dedication to supporting my research, and understanding. I would like to thank Stephen not only for providing the initial impetus for this thesis, but also for his helpful inputs and suggestions along the way. I have learnt a great many things from both my supervisors and it has been an interesting challenge combining the thoughts of two experts approaching the same problem from different angles.

One of the best things about doing my PhD at Exeter has been the friendship of my fellow PhD students. I particularly want to acknowledge those in the cognition group: Tobias Stevens, Felice van ‘t Wout, Heike Elchlepp, Hannah Filmer, Fayme Yeates, Amy McAndrew, Christina Meier and Rossy McLaren, for their friendship, support and creating a creative research atmosphere within the department. I am particularly grateful to Rossy, whose expertise in running experiments and managing the labs has been vital to the experiments presented in this thesis. I would like to acknowledge in particular the contributions of Tobias and Felice, who have on many occasions provided thoughtful and critical comments on my work.

I am also thankful for the contributions of two academics on their visits to Exeter. Firstly Ben Newell, who introduced me to state-trace analysis and very kindly ran an introductory session. And secondly Evan Livesey who inspired me to think more deeply about the associative structure of the CSR group and very generously took the time to discuss AMAN with me.

I also want to thank the University of Exeter, which funded my PhD, allowed me to get extensive teaching experience and offered a great research environment, with access to many training opportunities and seminar series.

Finally I would like to acknowledge the contribution of Adam Drury, who throughout my thesis has not only provided emotional support when needed but has also acted as an excellent sounding board to help me get my ideas straight.

Contents Page

List of Figures

List of Tables

Declaration

The research reported in this thesis was carried out at the University of Exeter between October 2008 and July 2012 and was supervised by Professor Ian McLaren and Professor Stephen Monsell. Except for the "Monsell Experiment" which was an unpublished pilot study conducted by Stephen Monsell and colleagues.

This dissertation has not been submitted, in whole or in part, for any other degree, diploma or qualification at any university.

Charlotte Forrest
Exeter, July 2012.

Chapter 1: General introduction

1.1 Dual Process theories

> *"…they're finding out what being human really means… that you're not as much in control as you think" (Pratchett, 2001, pg 305)*

> *"It was his subconscious which told him this – that infuriating part of a person's brain which never responds to interrogation merely gives little meaningful nudges and then sits humming quietly to itself saying nothing" (Adams, 1988, p148)*

> *"Women observe subconsciously a thousand little details, without knowing they are doing so. Their subconscious mind adds these little things together and they call the result intuition." (Christie, 1926, p148)*

It is a widely held belief that, although some aspects of our behaviour are under our control, there are aspects that are not. In recent years there has been a growth in popular science books claiming that the human mind is governed by two systems: one subject to our conscious control and the other operating without conscious intention (e.g. Thaler & Sunstein, 2009; Kahneman, 2011; Gigerenzer, 2007). This is not a new way of viewing the human mind; indeed Freud began popularizing the idea of a subconscious mind as early as 1896 - around the time of the beginning of modern psychology (Bowler & Morus, 2005). The more recent conceptions of dual process theories, however, present the unconscious mind less as a repository of our darkest wants and desires and more as an associative learning machine picking up the regularities in our environment. Over the years many different dualistic theories have been developed that describe the two systems in subtly different ways. Table 1.1 below gives an indication of some of the typical descriptors used.

System 1/ Type 1 processes	System 2/ Type 2 processes
Automatic	Controlled
Associative	Intentional
Effortless	Effortful
Fast	Slow
Unconscious	Conscious
Skilled	Rule-based
Parallel	Serial
Only final product subject to conscious awareness	Uses working memory
Based on past experience	Hypothetical thinking/ planning (predicts future)
Tacit	Declarative
Implicit	Explicit
Incidental	Deductive
Perceptual	Hypothesis Testing
Procedural	Uses executive attention
Exemplar based	Linguistic
Link formation	Propositional
Statistical learning	Reasoning
Relatively undemanding of cognitive capacity	Demanding of cognitive capacity
Interactive intelligence	Analytical intelligence
Heuristic processing	Analytic processing/ cognition
Experiential system	Belief based
Quick and inflexible modules	Rational choice strategy
Highly contextualized	Decontextualized
Acquisition by biology, exposure and personal experience	Acquisition by cultural and formal tuition
Intuitive cognition	
Recognition primed decisions	
Can be programmed to mobilize attention	
Executes skilled response	
Biased to believe and confirm	
Distinguishes the surprising from the normal	

Table 1.1 Words and phrases used to describe the properties of the hypothetical two systems (Thaler &Sunstein, 2009; McLaren, Green & Mackintosh, 1994; Evans, 2003; Ashby, Paul & Maddox, 2011, Mitchell, De Houwer & Lovibond, 2009, Stanovich & West, 2000, Kahneman, 2011).

Attempts have been made to dissociate the two systems, using a range of paradigms to demonstrate their independent existence and characteristics (see Mitchell, De Houwer & Lovibond 2009; Shanks, 2010 for reviews). Some of these properties are still highly debatable, e.g. unconscious/conscious (Shanks & St John, 1994), as it hard to prove that something was learnt unconsciously and so there is some difficulty in using this as a criterion. In light of this, I will focus on using the first two properties in Table 1.1 as the criterion for distinguishing the two, considering the distinction between an automatic/associative system and a controlled/intentional system (McLaren, Green & Mackintosh, 1994; Jacoby, 1991). Using this criterion, we can assume that if a process is under an individual's control then they should be able to stop and start the process at will. This criterion has been successfully used to assay the relative contributions of controlled and automatic processes for recall (Jacoby, 1991; Jennings & Jacoby, 1993). However, although this criterion is easier to operationalise and use as a basis for experimental manipulation than consciousness, it still presents issues as we cannot necessarily assume that participants *will* follow instructions even if they *can*.

In addition to the question over how the distinction is made, there are other assumptions made in a two-system account. Typically, processes at both levels are thought to operate simultaneously, with a degree of independence (with the notable exception that the final product of System 1 (Table 1.1, henceforth referred to as the associative system) is often thought to be subject to conscious awareness). To a first approximation, this means that the automatic system is always operating in the background, regardless of what the controlled system is doing. Additionally it is also assumed that the automatic/associative system is evolutionarily older than the controlled system, and so the behaviour of other animals, such as pigeons and rats is often thought to be governed solely by an automatic/associative system (McLaren, Green & Mackintosh, 1994 but see Beckers, Miller, De Houwer & Urushihara, 2006).

Over the years many attempts have been made to study the controlled processes, which supposedly are a defining feature of humanity. One of the archetypal classes of paradigms considered has been the task-switching paradigms of which the task-cueing paradigm (Shaffer, 1965; Biederman, 1972; Sudevan & Taylor, 1987; Meiran, 1996; see Section 1.2.2 for a definition) considered in this thesis is a common example. This paradigm is widely used to assess people's ability to exert control on their actions. It has been used in research on areas such as aging (Kramer et al, 1999; Kray & Lindenberger, 2000; Mayr, 2001) and schizophrenia (Meiran, 2000) and more recently has also been included in many “brain training” packages as a standard way to increase ones ability to both multitask and better focus one’s attention (http://www.lumosity.com/knowledge-center/brain-reference/task-switching). As this paradigm is used so widely as a tool, it is important to consider whether it actually tests what it is claiming to test – namely people’s ability to switch in a controlled manner between tasks.

In the work reported in this thesis I took a task-cueing paradigm, and examined what happens if participants approach this paradigm in a way that should rely more on associative/automatic processes. Beginning with the assumption that the associative system requires no knowledge of tasks, as it simply learns the statistical regularities in the environment around it, the participants were given no knowledge of the tasks but simply encouraged to learn the correct responses based on the input given and the feedback received after responding.

Researchers studying control sometimes deny the need for controlled processes to explain performance in a supposedly control based task (e.g. Logan & Bundesen, 2003), whilst those who study associative learning deny the need for associative learning (e.g. Mitchell, De Houwer, Lovibond, 2009). This approach often generates a great deal of controversy as researchers deny the existence of the very concept that interested other researchers in the research they were doing in the first place. In the case of Mitchell et al (2009) this approach also denies the intuitive and generally held belief that our behaviour is governed by two systems. Although this approach may provide a good platform for debate within the respective fields, it may not actually provide the best explanation of the behaviour. This thesis aims not to necessarily logically falsify either of these positions but to claim that conceiving of behaviour in terms of automatic and controlled processes provides a good explanation of the data gathered (for a definition of a good explanation see Lipton, 2004).

This thesis draws on ideas from several areas of psychology with large existing literatures. This introduction focuses on areas that are of the most direct relevance to the paradigm in question. I begin with an overview of the task switching literature and the standard results in task-cueing paradigms, then I consider what we might expect the results to be when participants take a more associative approach, and finally what might be a suitable associative model to use for this paradigm. Although there has also been considerable research into the neural correlates of task switching and relevant associative learning, these are not among the issues considered in this thesis.

1.2 Task-switching

Task-switching has attracted considerable interest in recent years (see Kiesel, Steinhauser, Wendt et al 2010; Vandierendonck, Liefooghe, & Verbruggen, 2010; Monsell, 2003 for reviews). Reconfiguring one's mind to perform a different task, especially when the environment continues to afford the previous task, seems a paradigmatic case of a controlled (endogenous, top-down, voluntary) cognitive process. Task switching itself has been studied using many different paradigms: list procedure (Jersild, 1927; Allport 1994; Spector & Biederman, 1976), alternating runs (Rogers & Monsell, 1995), intermittent instructions (Gopher, 1996) voluntary task switching (Arrington & Logan, 2004) and task-cueing (Meiran, 1996, Sudevan & Taylor, 1987). As task switching is usually considered to be the study of changing between task-sets it is important to understand what is meant by this. A 'task-set' is usually considered to consist of a collection of control settings and task parameters which configure the cognitive system to perform processes such as stimulus identification, response selection and response execution (Vandierendonck et al, 2010). Although all of these paradigms are relevant to the study of set-shifting, and similar phenomena are seen across the various paradigms, the focus here will be on the task-cueing paradigm, as this is the paradigm for which the most explicit argument has been made that performance can be explained without reference to selection of "tasks", but rather in terms of memory retrieval (possibly based on associative learning) (Logan & Bundesen, 2003, see Section 1.2.4).

1.2.1. Task-cueing paradigms

In a task-cueing paradigm, the participant is usually presented on each trial with a stimulus that they must classify or otherwise respond to according to a task rule. For example, in the experiments I will report, the stimulus is a digit between 1 and 9, and the two task-rules are to classify it as odd/even or as higher/lower than 5. As this paradigm is commonly construed, the cognitive demands of changing between task rules can be examined by comparing performance, or brain activity, on trials on which the task changes to trials on which the same task is performed as before, all other things being equal. This and some other examples of task pairs are listed in Table 1.2, to give an idea of the range of typical realizations of the task-cueing paradigm.

Stimuli	Tasks	Cues	References
8 digits	Classify odd/even or high/low	Words (“parity”, “magnitude”) or arbitrary letters, shapes or colors.	Monsell, Sumner and Waters (2003) Logan and Bundesen (2003) Schuch and Koch (2003)
16 colored shapes	Identify color or shape	Words ("color", "shape") and pictorial cues.	Monsell and Mizon (2006) Expts 4 and 5.
A smiley face in one of four quadrants.	Classify as above/below, or left/right of, center.	Arrow cues indicating relevant dimension.	Meiran (1996)
4 rectangles	Classify by height or width.	Cues are the letters “h” or “w”	Altmann (2004)
8 letters colored red or blue.	Classify vowel/ consonant or red/blue	Stimulus position in a 2 by 2 grid.	De Jong (2000)
4 or A displayed large or small in red or blue	Identify size, color or form	A double arrow, yellow patch or “$”	Koch (2001)
6 colored words	Name the word or the color.	location –top/bottom of screen.	Allport and Wiley (2000)
4 colored rectangles	Indicate orientation color	The digits 3, 4, 5 or 6.	Slagter et al (2006)
8 colored face pictures.	Classify or color gender	The words "gender" or "color" with and without S-R mappings	Shi, Zhou, Muller & Schubert, (2010)
Letter-digit pair	Classify digit as odd/even or letter as vowel/consonant.	One of 4 “hot” or “cold” colored squares.	Jamadar, Hughes, Fulham, Michie & Karayanidis (2010)

Table 1.2 shows examples of the types of tasks and cues used in task-cueing paradigms.

1.2.2. Task switching phenomena

Among the phenomena seen in such experiments, I focus on three: the switch cost, effects of preparation on the switch cost and the response congruence effect.

1.2.2.1 The switch cost.

This is the often substantial increase in reaction time (and error rate) on trials when the task changes from the previous trial, compared to trials where the task repeats (see Kiesel et al 2010; Monsell, 2003, for reviews). These switch costs are usually around 200ms relative to a baseline of around 500ms (Monsell, 2003).

1.2.2.2 The reduction in switch cost with preparation.

The task-cueing paradigm allows one to manipulate the time available for preparation by varying the cue-stimulus interval (CSI) while keeping the time since (and hence any carry-over from) the previous task execution constant (Meiran, 1996). Provided participants are adequately motivated, an opportunity for preparation improves performance overall —as one might expect from a century of experiments on fore-period effects and phasic alertness (Niemi & Näätanen, 1981; Nobre, Correa & Coull, 2007), and also because a long enough CSI removes the dual task costs of processing cue and stimulus simultaneously. Critically, however, an opportunity for preparation also reduces the switch cost. This *reduction in switch cost* (RISC) with preparation is frequently interpreted as resulting from top-down control operations: the time to prepare allows subjects to accomplish at least some part of the task-set reconfiguration that would otherwise have to be done after stimulus onset. Doing it in advance may save processing time consumed by such processes (Rogers &Monsell, 1995); it may reduce interference due to competition from the other task (because of advance activation of the relevant task-set and/or suppression of the irrelevant task-set); it may reduce interference due to concurrence of control processes and stimulus processing after the stimulus; it may do all three. However, there appears to be a limit to how much can be achieved in readying the system for a change of task through advance preparation. Preparation almost never entirely eliminates the switch cost; as the CSI increases, the RT switch cost usually decreases to an asymptote, typically reached at a CSI between 500 and 1000 ms (Meiran, 1996; Meiran, Chorev & Sapir, 2000; Monsell & Mizon, 2006).

One account of this asymptotic "residual" cost is that it is due to some irreducible persisting activation of the previous task-set - "task-set inertia" (Allport, Styles & Hsieh, 1994; Meiran 1996; Yeung & Monsell, 2003; Yeung, Nystrom, Aronson & Cohen, 2006). Another (not mutually exclusive) possibility is that the stimulus associatively reactivates task-sets previously associated with it, including competing task-sets (Koch & Allport, 2006; Rubin & Meiran, 2005; Waszak, Hommel & Allport, 2003; but see Monsell & Mizon, submitted). A third possibility is that task-set preparation simply "fails to engage" on a proportion of trials (De Jong, 2000; Nieuwenhuis & Monsell, 2002). Another measure of the limits of preparation is that merely preparing to do Task A, without actually performing it, is not sufficient to cause a switch cost when Task B follows on the next trial; Schuch & Koch (2003) showed that aborting the response with a “no-go” signal abolished the switch cost on the next trial.

1.2.2.3 The response congruence effect.

In many task-switching experiments the same small set of responses is used for both tasks, and all or many of the stimuli afford both tasks: they are “bivalent”. For example, if the two tasks are odd/even and high/low classifications of a digit, the left key might be used to respond both "odd" and "low" and the right "even" and "high". In this case, some digits (e.g. 1) would be *congruent* — receiving the same response irrespective of task — while for others (e.g. 7) the appropriate response depends on the task in play — these are *incongruent* stimuli. Congruent stimuli are usually responded to faster than incongruent stimuli, (e.g. Monsell & Mizon, 2006; Rogers & Monsell, 1995; Kiesel, Wendt & Peters, 2007), which has been taken as suggesting that when an irrelevant task-set has recently been active and/or must be kept available for the next switch, it is not completely disabled — its S-R mappings still mediate some degree of response activation, which either combine (congruent) or compete (incongruent) with the response activation generated by the current task's S-R rules. The congruence effect is usually amplified on task-switch trials, suggesting that persistence (or reactivation) of the previous trial's S-R rules is at least part of the problem causing the switch cost. But whereas the switch cost is transient, with performance recovering to asymptote within 3 or 4 trials of an unpredictable switch, the effects of congruence are much more persistent (e.g. Monsell, Sumner & Waters, 2003). There is also some evidence to suggest that the congruence effect can be reduced with time to prepare (Kiesel et al, 2010, Monsell & Mizon, 2006), however, the evidence for this effect has been somewhat mixed, with some experiments finding the effect and others not. The evidence suggests that whether or not this effect is found is independent of whether a RISC effect is found (Monsell & Mizon, 2006).

1.2.2.4 Other task switching phenomena

It is worth noting that there are numerous additional phenomena seen in task-cueing experiments which will not be considered in any detail in this thesis, though their interpretation has generally assumed a "control" or task based account of task-cueing rather than the "associative" account to which we will contrast such an account. For example, the n-2 "backward inhibition" (Mayr & Keele, 2000) in which participants show a disadvantage in the sequences of tasks ABA relative to CBA, has generally been interpreted in terms of task-set A having to be actively inhibited when the switch is made to task B, and the inhibition taking some time to "wear off". The finding of asymmetrical switch costs where there is a greater switch cost when switching to the easier task than the harder task, which has been interpreted in terms of more top-down control needing to be asserted to permit the weaker task-set to dominate (Allport et al, 1994; Yeung & Monsell, 2003). The complex interaction between response and task switches, whereby a response repeat has an advantage on a task repeat, but a disadvantage on a task switch (Brown, Reynolds & Braver, 2006). Also the Gratton effect, whereby if the previous trial is incongruent then the congruency effect is smaller on the next trial (Gratton, Coles, & Donchin, 1992). This is usually interpreted as an increase in cognitive control processes after an incongruent trial (Botvinick, Braver, Barch, Carter, & Cohen, 2001) though there has been some argument about whether the effect is merely an artefact of stimulus repetition (Mayr, Awh & Laurey, 2003).

1.2.3 Computational models of task-switching

In addition to numerous descriptive theoretical accounts of task-switching performance several computational models have been developed to account for aspects of performance in these paradigms (e.g. Altmann & Gray, 2008; Botvinick & Plaut 2002; Brown, Reynolds & Braver, 2007; Gilbert & Shallice, 2002; O'Reilly, 2006; Roelofs, 2003, Yeung & Monsell, 2003). The computational models presented within this thesis were *not* designed as models of cognitive control, but as a realisation of a control-free associationist account of performance in the task-cueing paradigm. It is nevertheless worth considering briefly Gilbert & Shallice's (2002) model and its more recent extension by Brown, Reynolds & Braver (2006), which are developments of Cohen, Dunbar & McClelland's (1990) parallel distributed processing (PDP) model of the Stroop effect. Although these models and the associative model to be presented below are both inspired by a connectionist architecture, they are completely different in operation, sharing only their parallel distributed processing heritage.

1.2.3.1 Gilbert & Shallice (2002)'s model

Gilbert & Shallice (2002) applied their model to a Stroop-based task switching experiment, in which one task involved the participant naming the ink colour of a colour word, e.g. RED would require the response "black", and the other task involved reading the colour word, e.g. RED would require the response "red". In order to model this they had separate input units coding both the ink colour and the word aspects of a stimulus. As reading words is generally considered to be the dominant task, the links for the word-reading task were set to be stronger than those for the colour-naming task. These units directly mapped onto their own output units. In addition to these units the model also contained task demand units. These units gave additional input to the output units of the currently relevant task's output units. More details of this structure can be seen in Figure 1.1.

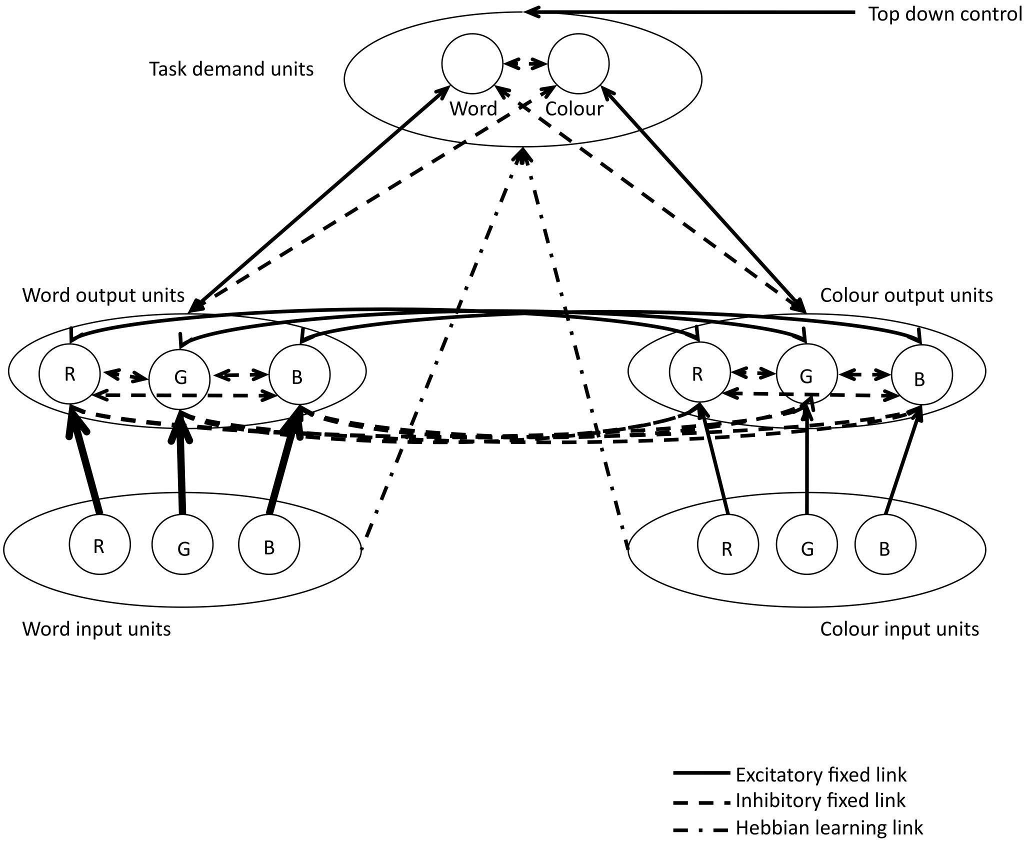


Figure 1.1 shows the structure of Gilbert and Shallice's (2002) model. Note that the arrows between the ovals that group units of similar type indicate that every unit in that group has a link, whilst links from specific units only indicate what that specific unit is linked to. Note that, although there is Hebbian learning between the input and task demand units this learning only persists from trial N-1 to trial N to avoid unbounded increments to the links.

In order to simulate a trial the appropriate top down input (dependent on the task in play) was added to the task demand units on every cycle through the network within the trial. This is in some ways a basic model of task-set reconfiguration. This process continues for the duration of the preparation interval. At the end of the preparation interval, the appropriate word and/or colour inputs are then set to their maximum value. Activation is the allowed to propagate until the response threshold is reached. In order to propagate activation, any unit whose activation needs updating has its new activation calculated by adjusting its current state by a factor related to the sum of the weighted input to that unit – that is the product of the activation of a contributing unit multiplied by the strength of the weight between that unit and the unit of interest. This is then used to calculate the change in the activation of a given unit. This is done using the following equations:

$$Net\,input > 0 : \Delta_{act} = step \times net \times (max - act)$$

$$Net\,input < 0 : \Delta_{act} = step \times net \times (act - min)$$

Where *act*= current activation, *step*= step size, net=net input, *max*= maximum activation value and *min* = minimum activation value (note that the activation of a given unit is constrained within this maximum and minimum activation). In addition to this the model also adds some Gaussian random noise into the system. The model is deemed to reach a response when the difference between the most active output unit and its nearest competitor (excluding the unit which would create an identical response) has reached a threshold.

It is worth noting that only links between the input units and the task demand units are capable of learning and this learning is only sustained from the previous trial, as at the end of a trial these weights are determined by the multiplication of the activations of the two units in question, with all other links set to zero. The activation of the task demand units is also carried over from the previous trial, but reduced by a squashing parameter from its terminal value at the end of the previous trial. These two aspects of the model are the only connections that vary on a trial by trial basis. This means that this model is a model of performance rather than learning as there is no long-term learning within the model. Both of these mechanisms represent different aspects of interference accounts of the switch cost– the learnt connections to the task demand units represent interference from stimulus-task associations, whilst the sustained but squashed previous activation of the task demand units represents the inertia of the task-set from the previous trial.

Among other effects this model was able to account for the greater effects of congruence (i.e. the differences between neutral, congruent and incongruent words) in the colour naming task (Dunbar and MacLeod, 1984) than the word reading task due to the word reading task interfering more with the colour naming task, of course this was already true of Cohen, Dunbar & McClelland's (1990) precursor model. The model also produces the position in run effects found in an alternating runs paradigm (Rogers and Monsell, 1995) whereby the switch cost is confined to the first trial in the run (albeit by assuming that activation only persists from trial N-1). It also demonstrated asymmetrical switch costs such that switching to the easier task produces larger switch costs (Allport et al, 1994); and was able to show a reduction of the switch cost with time to prepare (Rogers and Monsell, 1995). Hence a model which just combines a very basic interpretation of task-set inertia and task-set reconfiguration is able to account for a wide range of commonly found task switching effects. Note that the task-sets (implemented through the structure of the model and then selected by the task demand units) are important to this model for its account of performance in a task-switching paradigm.

Although this model is able to the account for the basic findings there are many more complex findings within the literature it is unable to account for. For example, given that it only has any remaining representation of the previous trial it could not account for any effect that is caused by a longer lasting persistence, e.g. the n-2 "backward inhibition" (Mayr & Keele, 2000).

1.2.3.2 Brown Reynolds & Braver (2007)'s model

The Gilbert and Shallice model was more recently updated by Brown, Reynolds & Braver (2007). They significantly increased the complexity of the model and made attempts to incorporate neuronal information into their model. This model was designed to take account of higher-order sequential effects, e.g. the Gratton effect (where the congruency effect is smaller after an incongruent trial) and its interaction with task switching (that following an incongruent trial there is an advantage on a repeat trial but a disadvantage on a switch trial, on the grounds that prior incongruency consolidates the current task-set (Goschke, 2000)). It was also designed to accommodate response alteration costs and their interaction with switch costs (there is an advantage for response repeats when the task repeats, but a disadvantage for response repeats when the task switches).

Their model contained a task-set layer which had many properties similar to the task demand layer of Gilbert & Shallice (2002). In Brown et al's model the input from the task cue[1], instead of directly biasing the output layer, biases the hidden/"plan" layer of the model to interpret the correct aspect of the target input and this in turn activate the correct response from the output layer. And so unlike Gilbert and Shallice's model where the task input directly biases the output layer, in Brown et al's model it biases the

---

[1] Brown et al unlike Gilbert and Shallice simply conceptualise the input to the network as coming from the task cue rather than coming from top down control.

way in which the target stimulus is interpreted. For a diagram of the structure of Brown et al's model see Figure 1.2.

In addition to this component the model also had two other "supervisory control" components – one sensitive to monitoring incongruence and the other sensitive to monitoring task-set switches. The incongruence monitor (INC loop, Figure 1.2) detects conflict between incompatible stimulus-response mappings at the hidden/plan layer. This conflict detector is proposed to be a separate monitor as the regions for conflict detection are distinct from those areas that deal with cognitive control (Botvinik et al, 2001). This is coded by having one unit for each of the possible incongruent stimuli in each possible task, which simply detects this pattern of activation in the hidden/plan layer. When this loop detects conflict at the hidden/plan layer it acts to strengthen the current task-set. This activation was persistent across trials, and so facilitated situations in which the task-set remained constant. This allows the network to account for the Gratton effect (see above (1.2.3.4)) as previously presenting an incongruent trial acts to strengthen the currently active task-set and so performance on a repeat trial following an incongruent trial will be enhanced.

The other component of the model, the change detector (CH loop, Figure 1.2), monitors for task or response changes across trials. It monitors the level of lateral inhibition in the task-set level, and, if this is high (i.e. when the task has recently changed), then it activates a general slowing mechanism (tonic, Figure 1.2). The change detector also monitors if the response has changed from the previous trial. In order to do this it has a response buffer that maintains the previous response. As with detecting task-set conflict, its response to detecting response conflict is to the impose a general slowing mechanism (tonic, Figure 1.2). This means it contributes to the switch cost, by producing a general slowing of the response on switch trials. This mechanism posits that the switch cost is partially generated by increased caution on switch trials as compared to repeat trials, rather than being entirely due to the controlled process being needed to reconfigure the task-set.

This image has been removed by the author of this thesis/dissertation for copyright reasons

Figure 1.2 Structure of Brown et al model, adapted from Brown et al (2007, pg55 Figure 8). Where the INC loop monitors for congruency and the CH loop monitors for change (task or response).

However, the switch cost in Brown et al's model is also caused by associative task-set interference. This is implemented in the model in a different way from the Gilbert & Shallice model which had a link from the stimulus layer to the task-set layer, in that in Brown et al's model this is simulated as a link between the hidden/plan layer and the task-set layer. In order to simulate the effect of response repetition being an advantage in the same task but a disadvantage if the task changed, the link between the hidden/plan layer and the output layer was strengthened for the relevant task units and the relevant response on each trial. This was at the expense of the link to the opposing task for that response due to normalization of the total link strength.

As with Gilbert & Shallice's model the connections between the cue input and task-set and between the target input-> plan layer and the plan -> output layer were one-to-one mappings, rather than every unit being connected to every other unit (as in the connectionist architecture presented below). Again as with Gilbert & Shallice's model the task-set units are able to sustain their activation. They also had a "feedback" connection to the hidden/plan layer which allowed the task-set units to activate the relevant units for that task over the relevant units for the other tasks. This connection was bidirectional so that there could be target/stimulus based activation of the task-set layer- similar to that in Gilbert & Shallice, but now via the hidden/plan representations. Again, as with the Gilbert & Shallice model, there was lateral inhibition between the two task units. The plan units were connected up to the output units to create the appropriate response. In line with Gilbert & Shallice's approach, the network was defined as having made a response when the activity of one of the units reached a pre-specified threshold – this enabled measures of both accuracy and latency.

The processing in the model was done in a real-time manner. During the CSI the relevant cue input was activated for the duration of the cue. This activation was then propagated forwards to the task-set layer where the relevant unit would then become active (an inhibitory current meant that the baseline activation remained low). After the CSI, the target/stimulus input was then activated and propagated forward to the hidden/plan layer where the activation combined with that from the task-set layer. In addition to the input from the hidden/plan layer, the output layer also received a non-response-specific input from the target layer – where each target unit was connected to each output unit.

Slightly more of the links are subject to learning in this model than the Gilbert & Shallice model as both the links between the hidden/plan and output units, and the links between the hidden/plan and task-set units, were modified by a Hebbian learning algorithm (Hebb, 1949). The learning in the link between the hidden/plan and task-set units was included to simulate TSI. The learning in the link between the hidden/plan and output units represents a form of response priming. As well as learning in a Hebbian manner, these links also experienced rapid exponential decay back to baseline (although less rapid decay than in the Gilbert & Shallice model).

They found that their model was able to account for the complex pattern of data generated by a task cueing paradigm, using an odd/even and vowel/consonant task on a letter digit pair. In their experiment, they found that there was an interaction between response switching and task switching, such that the switch costs were larger for response repetition trials, although the opposite effect was present in their error data. They also showed that previous task switch or response alteration trials resulted in longer RTs on the subsequent trial, including when the trial in question was a task switch or a response alteration. This was their justification for their change monitor detector.

They showed that a preceding incongruent trial speeds up responses more on a task repeat trial than a task switch trial. In addition, this effect is amplified if the preceding trial is also a repeat trial. This was their justification for the conflict detector, because they explained this effect as a previous incongruent trial establishing the task more – akin to an explanation of the Gratton effect (Gratton et al, 1994, but see Mayr, Awh & Laurey, 2003 for an alternative explanation based on stimulus repetitions).

Again, as with Gilbert & Shallice's model, Brown et al's model is a model of performance and not learning. They assume that the majority of the links are pre-existing, such as the link between the cue and task-set representations. This additionally means that every run of the model would be nigh on identical (if the sequence presented was similar enough). Indeed when they analyse the model they simply show the model the total number of trials experienced by all the participants in the experiment and then chop this up into participant sized chunks as it makes no difference doing this or running the "participants" separately. This means that the model has no way of encoding any individual differences in task switching ability (e.g. Miyake, Friedman, Emerson, Witzki, Howerter & Wager, 2000) or changes in task switching ability with training (e.g. Karbach & Kray, 2009). In fairness to the models considered, these issues have never been the focus of the debate in the task switching literature.

The Brown et al model assumes multiple different drivers of the switch cost – the activation of the task-set which is driven by response conflict and the amount of time the cue is presented for, as well as the strength of the link to the relevant hidden/plan units from the task-set units (task-set inertia), and the amount of general slowing in the system that is driven by switching the task and changing the response. The model suggests that the congruency effect comes from competition at the hidden/plan and output level which is driven mostly by the stimulus exciting two different responses and also by incorrectly activating the wrong task-set. However, it is exactly this link to the task-set level which allows the model to handle incongruent stimuli as the task-set units will also act to reduce the competition from the irrelevant attribute of the stimulus by strengthening the processing of the relevant attribute. This model and Gilbert & Shallice's model both attribute their explanation of the phenomena in task switching in large part to the representations of task-set (task demands). For both models, performance is mediated via knowing about the task based structure of the paradigm. In Section 1.4 these models are compared to an association formation model.

1.2.4 Associative approach

The purpose of this thesis is not to decide among these various accounts of how task-set processes, e.g. task-set reconfiguration, task-set inertia (as instantiated in the above models), might generate the phenomena in task-switching experiments, but instead to contrast them as a class with an associative learning account of performance in this situation, when the models in question have no inbuilt knowledge of the tasks as such. Most task-cueing experiments use a small pool of stimuli that repeatedly recur throughout the experiment, and just one or two cues per task (see Table 1.2). In such circumstances, it is a logical option for the participant simply to learn a combination of component and biconditional discriminations (as they are sometimes described in associative learning experiments, e.g. Livesey et al, 2011). For example, if the task cue is a background colour and the stimulus is a digit, the participant might learn a set of contingencies of the form: "pink background+7 → right key", and then on each trial retrieve from memory the response associated with the current cue + stimulus pair. Note that for congruent stimuli, the stimulus alone is completely predictive of the response and so only simple S-R associations need be learned: e.g. "3 → left" to solve the discrimination. Hence although the experimenter may construe the experiment as one that requires switches between task-sets, the participant might (after some practice) be operating with just a single task-set: retrieve the learned response associated with the cue + stimulus combination presented. We will refer to this strategy as CSR learning (where C stands simultaneously for "cue", "contextual" and "conditional"; S for "stimulus" and R for "response").

1.2.4.1 Logan and colleagues' compound retrieval model

An example of such a theory, albeit not under this name, has been proposed by Logan & Bundesen (2003) and developed in a number of papers (Arrington & Logan, 2004, Logan et al, 2004, Schneider & Logan, 2005, 2007). Their starting point was the finding by Logan and Bundesen (2003) that (in at least some circumstances) when two cues per task were used, a large part of the switch cost could be accounted for by cue changes not task changes and that masking the cue had a larger impact on repeat trials than switch trials[2]. They suggested that the participant encodes the cue, then the stimulus, and the combination simply retrieves the appropriate response from memory — which implies, of course, that the CSR association has been learned in some form. What appears to be task-switch costs, they argued, arise because cue encoding is primed

---

[2] Although further investigation into this finding suggested that this initial experiment perhaps over estimated the contribution of the effect of changing the cue to the switch cost, see for example Monsell & Mizon (2006) and Schneider & Logan (2011).

when the same cue was encoded on the previous trial. More formally they assumed that there was both a long-term and a short-term representation of the cue. On a repeat trial both the long-term and short-term memory representations match the current cue (if there is only one cue per task and the cue is repeated from the previous trial) and so it is encoded more quickly, whereas on a switch trial only the long term representation matches the current cue. So the mean cue encoding time on a repeat trial (μ_{rep}) is:

$$\mu_{rep} = \frac{1}{v_{STM} + v_{LTM}}$$

where ν_{STM} is the rate at which the cue is compared with the short-term memory and v_{LTM} is the rate at which the cue is compared with long-term memory. In contrast on switch trials the mean cue encoding time (μ_{sw}) is:

$$\mu_{sw} = \frac{1}{v_{LTM}}$$

Hence $\mu_{sw} > \mu_{rep}$ and so there is a repetition benefit for cue encoding on a repeat trial which causes the switch cost.

The driver of the RISC effect is the time taken for cue encoding. This exponentially distributed duration is assumed to frequently extend into the post-stimulus interval when the CSI is short, and Logan et al assume that retrieval of the response associated with the cue + stimulus compound waits until cue encoding is complete. Hence the shorter the CSI, the more often and longer on average CSR retrieval has to wait for cue encoding, and the greater the contribution of cue priming to performance. More formally, the probability that cue encoding is not finished at a given CSI on a repeat trial is $e^{\frac{-CSI}{\mu_{rep}}}$ and similarly for a switch trial. Hence the RT on a repeat trial (RT_{rep}) is given by:

$$RT_{rep} = RT_{base} + \mu_{rep} \times e^{\frac{-CSI}{\mu_{rep}}}$$

Similarly on a switch trial:

$$RT_{sw} = RT_{base} + \mu_{sw} \times e^{\frac{-CSI}{\mu_{sw}}}$$

This initial instantiation of the model is therefore able to account for two of the key findings in the task switching literature – the switch cost and the RISC effect- without the need to make reference to the tasks. They were able to demonstrate that cue switch costs were a large part of the task switch cost (Logan & Bundesen, 2003) but see (Mayr & Kliegl, 2003; Monsell & Mizon, 2006). The model was also able to account for the increase in the size of the switch cost with the reduction in the probability of a task switch by suggesting that when a certain transition type is more frequent then this transition type is more primed and so cue encoding for this type of transition is faster (Schneider & Logan, 2007).

The elegant simplicity of the model is its strength, as with very few assumptions it can explain important effects. It is, however, also its downfall as it is easy to come up with aspects of a task-cueing paradigm that it has no mechanism to account for. The most obvious example is the congruency effect as the model makes no reference to even encoding the stimulus, let alone considering how it is mapped to a response. This issue was addressed in Schneider & Logan (2005)'s adaption of the model by including a part of the model which retrieves the response based on the cue + stimulus compound.

They assumed the same model of cue encoding as above but additionally specified that the representation of the cue in short-term memory is only affected by the previous trial. They additionally assumed that two cues used for the same task would prime each other into short term memory via their representation in long term memory where they are both linked to the same task. In this model also the cue and stimulus retrieved the task SR mapping so the odd/even cue and 7 would retrieve the representation of "odd" which would then retrieve the appropriate response. This still requires retrieval of the relevant task-response rules, so this is different from the purely associative model outlined below where there is no mediation by task-response rules and the odd/even cue and 7 would simply retrieve the appropriate response.

Their adaptation of the model uses a parameter which determines the strength of evidence (η) for a given response category. This is thought to vary depending on the cue used on the previous trial – with a higher value being given if the cue repeats than if the same cue from a given category is used and again than if a cue from the other task was used. This is used to determine the rate at which information can be retrieved from short-term memory. It is this variation which drives a cue-stimulus congruency effect found when one of the response values is used as the task cue: for example, when the cue is congruent (ODD-3) then there is more evidence for a given category and so the category will be retrieved faster than if they are incongruent (EVEN-3). Additionally it is also used to determine the probability (p) of retrieving a given response category.

$$p(i|x,y)=\frac{\eta(i|x,y)}{\sum_{j\in R}\eta(j|x,y)}$$

Where *x* is the cue, *y* is the stimulus, *i* is the response category and *j* is any of the possible response categories (*R*). The value of η is determined by

$$\eta(i,x,y)=\eta(i|x)\times\sum_{y\in S_i}\eta(i|y)$$

Where S_i is the stimulus set *S* associated with the response category *i*. This equation represents compound cue retrieval.

Schneider & Logan (2005) considered the case of their model where two response categories mapped onto each response key. As mentioned above, this is the structure used in the experiments presented below. The probability of selecting a response key (A_n) is then the sum of the retrieval probabilities of the response categories assigned to that response:

$$p(A_n|x,y)=\sum_{i\in A_n}p(i|x,y)$$

In a way this is similar to the way in which the responses are modelled in the associative model used in this thesis, as the same response is coded via the same output unit and so both models assume that participants keep a concept of which response they are going to make in a physically relevant form. Schneider & Logan's model then uses a random walk procedure (Nosofsky & Palmeri, 1997) in order to select which response the model actually makes. In this case response counters are incremented in discrete time steps to implement the random walk until the difference between the counters reaches a given threshold. The drift rate of the random walk is defined by comparing the evidence for that response category compared to the summed evidence for all response categories – in other words the relative probability of that category. So on a given time step

$$RT_{step}(x,y)=\frac{1}{\sum_{j\in R}\eta(j|x,y)}$$

The number of steps for the random walk to terminate (N_{step}) is calculated by the following equations derived by Busemeyer (1982) and adapted by Nosofoky & Palmeri (1997)

$$N_{step}(x,y)=\frac{\theta_1(2\mathrm{K})-\theta_2(K)}{p(A_1|x,y)-p(A_2|x,y)}$$

Where K is the criterion and θ *is*

$$\theta_1=\frac{\left[\frac{p(A_1|x,y)}{p(A_2|x,y)}\right]^{2\mathrm{K}}+1}{\left[\frac{p(A_1|x,y)}{p(A_2|x,y)}\right]^{2\mathrm{K}}-1}$$

$$\theta_2=\frac{\left[\frac{p(A_1|x,y)}{p(A_2|x,y)}\right]^{K}+1}{\left[\frac{p(A_1|x,y)}{p(A_2|x,y)}\right]^{K}-1}$$

So the RT_{CCR} for the compound cue retrieval time is

$$RT_{CCR}(x,y)=RT_{step}(x,y)\times N_{step}(x,y)$$

This is simply amended into the final equation given by Logan & Bundesen (2003) so that it is now:

$$RT_{rep}=RT_{base}+RT_{CCR}+\mu\times \mathrm{e}^{\frac{-CSI}{\mu}}$$

Schneider & Logan (2005) were able to show that this model accounts for cue encoding benefits, switch costs and cue-target congruency effects (in which a congruent cue-target relationship has an advantage). More recently Schneider & Logan (2009) tested this compound cue model and found that a slightly adapted model was able to account for the influence of including a prime. They noted that the cues used in Schneider &

Logan (2005) could also act as primes, as the cue ODD would prime the response ODD, and that this might be different from the cues usually used in task-cueing experiments, which aimed to prime the task (e.g. parity) and not a specific response to the task. To pursue this issue they ran an experiment which included both a cue (e.g. parity) and a prime (ODD). They validated the assumption in the model above that the way in which the prime would enact a congruency effect is through the ease of retrieving the correct response, as this model provided the best fit.

Although the compound retrieval model is able to provide a possible account of performance in a task-cueing paradigm without the need to appeal to control or task based processes, the assumptions made in the model have come into question. Monsell & Mizon (2006) questioned the amount of time the model assumes cue encoding takes. In the model for a CSI of 100ms cue encoding usually has failed to succeed, whereas it has usually succeed at a CSI of 800ms. This seems like a long time for cue encoding given that the average RT to the stimulus is usually around 700ms, which would suggest that participants would be failing to encode the stimulus (if the stimulus needs to be encoded in a similar manner to the cue).

1.2.4.2 Evidence against the compound retrieval model

This *compound-retrieval* model has excited controversy because it rejects what is for most users of the task-cueing paradigm its raison d'être: the hope that we can use it to measure the process and/or limitations of endogenous cognitive control. However not all studies are in agreement with the compound retrieval model outlined above. Some studies using two cues per task have found substantial task-switch costs when cue switch is controlled (Mayr & Kliegl, 2003), and also reductions in those costs with preparation (Monsell & Mizon, 2006). However, as mentioned above, the elaborated theory (Schneider & Logan, 2005) argues that even when the cue changes, there is priming of cue encoding on task-repeat trials, either because two cues signaling the same task become associatively related (and hence primed by spread of activation) or because they retrieve a mediating task-representation (and hence its retrieval is primed because the same mediating representation was retrieved on the previous trial). This idea will be returned to later in the discussion of how we might expect CSR participants to behave (see Section 1.3.4).

One piece of evidence directly problematic for the theory in its chosen domain of application comes from the same laboratory. Arrington, Logan and Schneider (2007) used a so-called "double registration" procedure in which subjects were required to make two responses per trials. First they responded as soon as they had identified a cue's category (either with a response which corresponds to the cued task, as in their Experiments 2, 3b, and 4b, or with a response which simply identified the cue: Experiments 1,3a, and 4a). Only then was the stimulus presented for the classification response (e.g. odd/even). If "task-switch" and RISC effects are driven entirely by priming of cue encoding, such effects should appear only in the first RT (which presumably includes cue encoding). When the response to the cue was simply to identify the cue (Experiments 1, 3a and 4a) then there were still cue switch costs present in the stimulus RT. However, there were substantial task-switch costs in the second RT when one response per cued task was used (Experiments 2,3b and 4b), not in the first RT. That is, requiring the subject to interpret the task signaled by the cue (which is what "cue encoding" does in the Logan/Bundesen/Schneider theory) before the stimulus is presented did not eliminate the substantial task switch cost, which would seem to suggest it cannot result from priming of a cue encoding process that extends into the latent interval.

Regardless of such difficulties, the Logan/Bundesen/Schneider theory provides a concrete example of a theory according to which the phenomena of cued task-switching can in principle be accounted for by cue-stimulus response (CSR) learning and retrieval, rather than processes of task-set control. Curiously, however, although the compound retrieval theory requires that CSR learning of some kind must have taken place, little is said about that learning, nor have the authors' experiments assessed this acquisition process. Yet it is not trivial to acquire (for 8 digits and 4 cues, for example) 32 CSR associations or, if we assume that congruent stimuli are learned as simple S-R pairs, 16 CSR and 4 S-R associations. The amount of practice given before data are collected in such experiments has typically been quite small.

If the compound-retrieval model is a valid account of how participants perform in general in a task-cueing paradigm then we would expect little or no difference between performance of participants given standard instructions in a task-cueing experiment (henceforth referred to as a *Tasks* condition), and performance when participants are induced to learn cue-stimulus response associations and use a CSR strategy for retrieving the response (a *CSR* condition). The experiments presented in this thesis will demonstrate that performance in the Tasks and CSR conditions is in fact very different. I now review some other evidence which bears on how participants in the CSR condition might be expected to perform: first some evidence from within the task-switching literature, and then evidence from the associative learning literature.

1.2.5 The Dreisbach et al. experiments.

The most relevant experiments are reported in a series of papers by Dreisbach and colleagues. (Dreisbach, Goschke & Haider, 2006; Dreisbach, Goschke & Haider, 2007; Dreisbach & Haider, 2008; Dreisbach & Haider, 2009). They used an instructional manipulation related to the one used in this thesis to contrast an S-R classification performed according to a task-rule with the same classification accomplished by means of individual S-R rules. In general participants were taught a set of individual S-R rules for binary classification of a set of stimuli. The stimulus set was designed so that an additional task-rule-based classification strategy is available, but it was sufficiently non-obvious that participants generally did not notice it until it was drawn to their attention.

For example, participants were taught a binary classification of a set of 4 red and 4 green words specified by eight apparently arbitrary S-R mappings (half the words of each colour mapped to a left key, half to a right key). In fact the mappings were such that there were also hard-to-spot task-based rules available to specify the response: for a red word, the response was specified by whether the word began with a consonant or vowel, and for a green word, by whether it referred to an animal or not: the classic structure of a task-cueing experiment. Participants given the task-rules at the outset showed switch costs (as estimated in relation to the tasks structure) when the colour switched, whereas uninformed participants did not, and late-informed participants showed only modest switch costs after the rule was given. Interestingly, performance was notably slower, especially to start with, in the early-informed group who were using the task rule. That is, under these conditions, where the stimuli are univalent (i.e. each word associated with one response) so that only S-R, not CSR, associations have to be learned, using the task-rules seems to add difficulty, at least initially. They showed this was the case for both 1 syllable and 4 syllable words. Why then did the participants tipped off to the task-rule elect to use it if it leads to worse performance?

One possible reason is to do with the ability to generalize. One might intuit that the use of task rules would allow the participant to generalize to new situations to which the task rules also apply. As an extension of the experiments described above Dreisbach, Goshke & Haider (2006) added in a transfer condition for the early and SR conditions. They found that there was reliable increase in the RTs for the task informed group but not for the SR group – opposite to the intuition that task-rules allow generalistion. Dreisbach et al argue that when participants know the tasks the improvement in their performance over blocks is due to participants having learned the stimulus-specific mappings – hence when the stimuli are changed they get no benefit from the stimulus specific learning and so there is a marked decline in performance. This seems entirely reasonable as an argument. What is slightly odd is that Dreisbach et al then claim that the reason there is no cost of transfer in the SR condition is because they have implicitly extracted some rules. However, what they mean by this is not that they have extracted the task rules but that they have learnt 4 categorisation rules based on the colour and response, e.g. red words that started with a consonant require a left response and green animal words require a left response. In support of this they showed a disadvantage for introducing rule incongruent stimuli. This is more in line with how an associative model might solve this situation as it extracts the statistical regularities in the environment. So based on this data one would find it difficult to argue that using task-sets actually enables easier generalization (but see Chapters 2 & 3).

Dreisbach and Haider (2008, 2009) raise one other possibility. They report further experiments that suggest that one advantage of using a task-set rule is that it "shields" against irrelevant sources of variation that perturbs the performance of those using individual S-R rules. In their first experiment they showed that if participants approached the experiment using just one task rule then they were unaffected by irrelevant changes. That is to say in their SR condition and their 2 tasks condition they found an advantage if the colour repeated and the response repeated but a disadvantage if the colour repeated and the response switched. The opposite was true if the colour switched – that is there was an advantage for a switched response over a repeated response (although this was not found in previously in their SR condition, Dreisbach, Goshke & Haider, 2006).

In Dreisbach & Haider (2009) they showed that this also applied to other types of distractor. In two experiments, they had either a one task group (decide if the clothes word presented was an item that covered the legs) or an SR group. The difference between the two experiments was whether these groups were assigned a priori (by telling half the participants the task) or post hoc (by asking the participants what their strategy was). In addition to the word being presented on each trial they also included a distractor image which was either another item of clothes or an animal facing either to the left or right. They showed that performance when participants were using tasks was only interfered with when a relevant distractor was present. In other words only the images of clothes interfered – helping when they would require the same response and hindering when they would require the opposite response. When participants were relying on SR rules this pattern was again observed but additionally the animal distractors interfered – when they were facing the same direction as the required response there was an advantage and a disadvantage when they faced the wrong direction. Again they took this as evidence that task-sets shield participants against irrelevant information.

So these experiments suggest that there would be a difference between participants instructed to use tasks and those instructed to use SR rules – specifically it suggests that participants using SR rules show no switch cost. These results are, taken at face value, in disagreement with the compound retrieval model, as in the Dreisbach et al experiments cue switch and task switch were confounded, and so a switch cost should have been observed in the SR condition if cue priming did drive a switch cost. However, Dreisbach et al's stimuli only ever appeared in one task (unlike most task-switching experiments), and as such participants need not retrieve the cue-stimulus compound but just the stimulus to know how to respond. As the cue is not relevant to responding its priming effects may not be sufficient to cause a switch cost.

1.2.6 Task-switching in monkeys.

Another basis for predictions in the CSR condition is research by Stoet and Snyder (2003a, 2003b, 2007), who gave Rhesus monkeys *(Macaca mulatto)* thousands of trials in a simple task-switching paradigm (actually not dissimilar methodologically to some of the conditional discrimination work in monkeys described below). Without language to mediate the development of a task-set strategy (cf. Goschke, 2000), and after so much training, it seems likely that the monkeys were using a CSR strategy developed through trial and error learning rather than task-sets. Nevertheless, Stoet and Snyder (2007) found a small switch cost in the monkeys; human participants (who received very much less training) showed a much larger cost. Stoet and Snyder (2007) also found a large congruence effect in the monkeys they tested, suggesting that approaching the experiment without a task representation should yield a large congruence effect in humans too. Stoet and Snyder (2003b) did find a RISC effect in the two monkeys they studied. However, they used only one cue per task, so that a change in the cue to be processed was confounded with a change in task. In these circumstances, having more time to process the cue before the stimulus appeared may benefit the monkey more on a switch than it would on a repeat trial because the cost of a task switch is actually the cost of a cue switch (as Logan & Bundesen, 2003, suggested). Hence it may be the case that if cue change and task change were un-confounded, then little or no RISC effect would be observed when using a CSR strategy.

1.2.7 Summary of task-cueing evidence

The task-cueing literature has reported many phenomena, but three appear to be key: the switch cost, the response congruence effect and the reduction in switch cost with preparation effect. The last is usually taken as an index of cognitive control. However, Logan and colleagues have suggested that this effect need not index task-set control but could instead simply result from the participants learning and retrieving cue+stimulus-->response mappings. Against this view of the task-cueing paradigm as being reduced to a cue-priming model, the animal research suggests that there may be a difference between performance given standard task-based instructions and more associatively based performance. In order to further examine what we might expect if participants are simply learning the associations present I now turn to the relevant associative learning literature.

1.3. Associative learning literature

		Cues for “Task 1”		Cues for “Task 2”		
		W	X	Y	Z	
Stimuli	A	L	L	L	L	Component Discriminations
	B	R	R	R	R	
	C	L	L	R	R	Biconditional Discrimination
	D	R	R	L	L	

Table 1.3 shows the contingencies present with a task-cueing experiment with two cues per task and four stimuli (2 congruent (A &B) and 2 incongruent (Y & Z)). A&B are known as component discriminations, whilst C &D represent a biconditional discrimination.

Table 1.3 shows the associative structure of the most commonly used task-cueing paradigm used in this thesis. As noted briefly above, this consists of a mixture of what have been referred to as *component* and *biconditional* discrimination (e.g. Livesey et al, 2011). A brief overview of the general findings of the literature on these types of discriminations will now be given. Another property of these discriminations is that W & X and Y & Z are functionally equivalent to one another, and so there will also be a brief discussion of *acquired equivalence*. The cue and stimulus occur at different times, with the cue occurring first, so it could be said that the cue sets the occasion for the stimulus, hence a brief overview of *occasion setting* will also be given. The literature will also help with the consideration of what sort of associative model would be appropriate for modelling the CSR group. This section will begin with a discussion of what is meant by an associative model, then consider component discriminations, biconditional discriminations, cue equivalence and finally occasion setting.

1.3.1 Associative models

What is an associative model? Are the models mentioned above, Gilbert & Shallice (2002), Brown et al (2006) and Schneider & Logan (2005) associative? All three models talk about links between various aspects in the task-cueing paradigm, e.g.

stimuli and responses (all three), cue input and task-set (Brown et al, 2006), and the cue and its representation in long term memory (Schneider & Logan, 2005). So in some ways these models are association based. However, what I really mean by an associative model is not just a link-based model but an *association formation model* (De Houwer, 2009; Dickinson, 2009), that is, a model that learns what the links should be based on input it receives from the environment. Apart from a few specific links in the models described above, e.g. the input-task demand link in the Gilbert & Shallice model (and even here is debatable), they are not models that learn over time, rather they are models of performance. That is not to say there is no merit in creating models of performance, given that most adult humans have had many years worth of experience, any aspect of which they could call upon to assist them in performing within a given experiment. Also, in the context of task-switching, it is the performance which is often of interest rather than the learning per se. Often little attention is paid to the ways in which participants are trained to respond, rather the focus is on what is happening when they have reached asymptotic performance, as it is in some ways the robustness of the switch cost over time that is of interest. And with these goals in mind such models may be appropriate.

However, although performance may be the focus of the task-switching literature, this is not the focus of the conditional discrimination literature. Here the focus has been more on how the discriminations are learnt, and any model inevitably will have some mechanism for learning, e.g. Livesey et al (2011). Hence these are the types of models that will be considered to model the CSR condition.

1.3.2 Component discriminations

When considering the associative structure of the task- cueing paradigm (Table 1.3) in question it has been noted that, for the congruent trials, if the participant can learn to “ignore” the cue on these trials then they can be solved easily by learning a simple stimulus-response association. This type of learning can easily be accommodated by common models of associative learning such as those of Rescorla and Wagner (1972), Mackintosh, (1975). As such, learning this discrimination is not the limiting factor in model selection.

For the most part, experiments using the structure of the congruent trials (component discrimination) have been used in investigations of learned predictiveness (Le Pelley & McLaren, 2003; Le Pelley, Suret, & Beasley, 2009; Livesey, Thorwat, De

Fina & Harris, 2011). These have shown that, when participants are trained on AX+, AY+, BX-, BY-, first, this is a relatively easy discrimination to solve and second, that the associability of A and B is higher than that of X and Y. In other words it is easier to learn new associations to A and B than it is to X and Y, a result that favours a Mackintosh (1975) type attention mechanism. Although this is an interesting issue, none of the experiments in this thesis look at making new associations to A and B, and so the modelling in the thesis will not directly address this issue.

1.3.3 Biconditional discriminations

The incongruent stimuli in a task-cueing experiment make up part of what is known as a biconditional discrimination. This type of discrimination has been of considerable interest due to its theoretical construction, as no single stimulus predicts the outcome. That is, for XA+, XB-, YA-, YB+, each of the elements A, B, X and Y are followed by reinforcement half the time and non-reinforcement the other half of the time. This poses issues for single layer models of learning, e.g. Rescorla and Wagner (1972), Mackintosh (1975). The evidence on acquisition of biconditional discriminations and comparison with component discriminations will now be considered in more depth.

There is a long history of showing biconditional discriminations in animals other than humans. For example, North, Maller and Hughes (1957) inspired by Lashley (1938), showed that rats were capable of learning a conditional discrimination. Both experiments trained rats first to discriminate an upright and inverted triangle on a black background by jumping towards one of the triangles. Then, subsequently the rats were trained to respond in the opposite manner to upright and inverted triangles on striped backgrounds. They were finally tested on all the possible combinations of positive and negative stimuli, and this showed that the rats had indeed learnt the full biconditional discrimination.

There is also evidence that monkeys can solve biconditional discriminations (rhesus macaques: Warren, 1964, Riopelle and Copelan, 1954; squirrel monkeys: Burdyn and Thomas, 1984). Burdyn and Thomas (1984)'s experiment is of particular note as their design was similar to a task-cueing paradigm, although the tasks here were not orthogonal like those used by Stoet & Snyder (see 1.2.5.2). In their experiment they initially trained squirrel monkeys to make sameness and difference discriminations, by training the monkeys (in separate blocks) to choose the side on which the two objects were the same (sameness) or choose the side where the two objects were different (difference). They then trained the monkeys that a triangle was present during the sameness discrimination whilst a heptagon was present when the requirement was to select the two different objects. They then randomly intermixed the trial types. Finally, they introduced an interval between the presentation of the triangle/heptagon and the discriminanda, which was extended until the monkey could no longer reach criterion (13 correct sameness and 13 correct difference judgments in the last 30 trials). They showed that monkeys were indeed capable of learning a conditional discrimination of this type, and some were able to cope with extended delays of up to a maximum of 16seconds. As by the end of the experiment there were 492 different objects being used this could be taken as evidence that the monkeys had acquired and were using the concepts of sameness and difference. However, as this experiment was run to see whether monkeys could learn this type of discrimination they did not consider transitional effects such as whether the monkeys were slower/less accurate when the type of discrimination required switched as compared to when it repeated. Clearly if one wants to truly understand whether monkeys are using the concepts of "same" and "different" one has to examine whether there are parallels to the ways in which humans use conceptual tasks, more in the way of Stoet & Snyder's experiments.

So animals can solve biconditional discriminations, but how does their learning of them compare to solving component discriminations? Saavedra (1975)'s Experiment 1 compared biconditional and component discriminations in an eyelid conditioning paradigm in rabbits. This was done between subjects with half the rabbits receiving a biconditional discrimination (e.g. A_1L_1+, A_2L_2+, A_1L_2-, A_2L_1-) to learn and the other half a component discrimination (e.g. A_1L_1+, A_1L_2+, A_2L_2-, A_2L_1-). The stimuli used were a combination of two different auditory tones and two different lights, whilst the unconditional stimulus (US) was an electric shock to the eyelid. They showed that, although the rabbits could learn the biconditional discrimination, it was much harder for the rabbits to learn than the component discrimination. This result is of particular relevance to the matter at hand as it suggests that participants learning using associative mechanisms would be expected to find the incongruent trials much harder than the congruent trials. Hence any model should be able to capture this difference.

This evidence suggests that various animals, whose learning is often considered to be associative (McLaren, Green & Mackintosh, 1994), can learn a biconditional discrimination, but given that the subjects in this thesis are humans it is important to consider some comparable human learning experiments. Harris and Livesey (2008) showed that biconditional discriminations are harder for people to learn than negative patterning (A+ B+ AB-) and positive patterning (A- B- AB+). In two experiments they used variants of the allergy paradigm (in Experiment 1, the relationship between wine varieties and hangovers; in Experiment 2, the relationship between demographic variables and whether a child had an above average literacy level). They used a within subjects design where all participants learnt both a negative patterning and a biconditional discrimination simultaneously with different stimuli used for each discrimination. In their second experiment they examined the possibility that the result in the first experiment was due to the negative patterning rule being easier to induce than the biconditional discrimination. In other words, the rule that the compound performs in the opposite manner to the elements is easier to induce and use than the rule if AB+ and CD+ then AC- and BD-. They showed, by means of increasingly more specific tests, that more participants induced the negative patterning rule than the biconditional discrimination rule. This suggests that part of the reason negative/positive patterning is shown to be easier is that it is the result of rule-based behaviour rather than associative learning. This potential for a rule-based explanation of participants' behaviour raises theoretical possibilities additional to the accounts typically applied in the animal literature, and something to consider when human participants are used. However, when only participants who did not induce a rule were included it was still true to say that negative patterning was easier to learn than biconditional discriminations. Harris and Livesey took their result to support Harris (2006)'s model and McLaren & Mackintosh (2000; 2002)'s model.

Although of interest in its own right, negative patterning is not part of the task-cueing paradigm. Of more relevance is Livesey, Thorwart De Fina & Harris (2011)'s first two experiments which constrasted the predictiveness of the different elements in a biconditional discrimination and a component discrimination, in a way similar to the learned predictiveness experiments mentioned above (see Section 3.2). They trained participants on 2 biconditional (JN-O1, JO-O2, KN-O2, KO-O1, LP-O1, LQ-O2, MP-O2, MQ-O1) and 2 component discriminations (AW-O1, AX-O1, BW-O2, BX-O2, CY-O1, CZ-O1, DY-O2, DZ-O2) concurrently. In line with the research outlined above they showed that the biconditional discrimination was much harder for people to learn than the component discrimination. They also showed that the predictive elements (of the outcome) in the component discrimination gained a higher associability than the elements in the biconditional discrimination and the non-predictive elements. That is, these predictive elements were able to enter into new associations more easily.

To test this in Experiment 1, after training the two biconditional discriminations and two component discriminations, they then had a second phase of training in which novel combinations were created and trained to new outcomes. In the case of the component discrimination these were composed of one previously predictive element and one non-predictive element which had not been paired before, i.e. AY-O3, BZ-O4, CW-O4, DX-O3. The biconditional discrimination elements were swapped in an equivalent way i.e., JP-O3, KQ-O4, LN-O4, MO-O3. The amount of learning was assessed by doing a summation test where elements that had been trained to the same outcome in stage two were combined e.g. AD and a negation test where elements that had been trained to opposite outcomes in stage two were combined, e.g. AZ. In these tests participants gave a rating on how likely they thought a given outcome would be. The results showed the expected pattern of results: predictive elements had maintained a higher associability than biconditional elements which in turn had a higher associability than un-predictive elements.

However, the design of this experiment did not directly pit elements in the component discrimination against those in the biconditional discrimination (though the contrast can be inferred). . The deficit was rectified in Experiment 2, which differed in stage two where compounds AJ-O3, BK-O4, CL-O4, DM-O3 and WN-O3, XO-O4, YP-O4, ZQ-O3 were trained instead, thus combining one part of an original biconditional discrimination with one part of the original component discrimination. The results showed that the elements of the biconditional discrimination and the non-predictive elements of the component discrimination had the same low associability. There was no difference in the amount of learning to the elements of the compounds containing non-predictive elements and biconditional elements but less was learnt about the previous biconditional component as compared to the previous predictive component (as little as about the non-predictive elements in Experiment 1).

These experiments suggest that as well as the biconditional discrimination being harder to learn than component discriminations, the status of the elements of the discrimination may vary. Although the associability of the cues and stimuli is not directly assessed in the experiments presented in this thesis, Chapter 5 does consider differences in the status of the elements based on the structure of the associations formed.

1.3.4 Cue equivalence

In task-cueing paradigms there are often cues or stimuli which are responded to in a manner equivalent to other cues or stimuli which are present. When considering the associative structure outlined earlier, this acquired equivalence would apply to the CSR group as two of the cues, W& X and Y& Z predict the same outcome for a given stimulus, and we would expect participants to learn this equivalence over time. The notion of acquired cue equivalence was also suggested in the task-switching literature by Schneider & Logan (2005), when they suggested that two cues per task were not really two cues per tasks as they would prime each other from long term memory into short term memory via their task representation. Consistent with this proposition, there is evidence in the associative learning literature to suggest that humans and animals do acquire equivalence between cues treated in the same manner (Honey, Close & Lin, 2010; Honey & Ward-Robinson, 2002; Coutureau, Kilcross, Good, Marshall, Ward-Robinson & Honey, 2002, Honey & Watt, 1998, 1999; Honey & Ward-Robinson, 2001; Ward-Robinson & Honey, 2000, Honey & Hall, 1989).

Honey and Ward-Robinson (2002) extended the typical biconditional discrimination by in effect doubling the number of cues/contexts, i.e. including the trials AX+, AY-, BX+, BY-, CX-, CY+, DX-, DY+, (N.B. the same effect can be obtained in a task cueing paradigm if there is more than one stimulus which behaves in the same way, e.g. 1 & 3 would both be odd and low and so require the same responses as each other in each task, or by doubling up the number of cues per task.) To test for cue equivalence, they tested responding to AB vs AD. A and B are equivalent cues whereas A and D are non-equivalent. They argued, based on a connectionist account (see Figure 1.3), that combining equivalent cues should lead to more variable responding than combining non-equivalent cues. This, they argued, was due to the fact that presentation of both A and B would lead to the hidden units used to carry the mappings for A and B being highly active (similarly to if they had seen AX etc) and in turn this would lead to both the relevant outcome units being turned on. This would lead the rat to simultaneously respond to food and no-food leading to highly variable responding. However, in the case of AD the hidden units involved would each be equally slightly active in a way which would not be enough for both the relevant output units to reach behavioural threshold leading to the rat to be less likely to respond. This is indeed what they found.

This image has been removed by the author of this thesis/dissertation for copyright reasons

Figure 1.3 adapted from Honey et al (2010). The figure shows the way in which the model proposed by Honey & Ward-Robinson structurally solves biconditional discriminations of the nature AX+, AY-, BX+, BY-, CX-, CY+, DX- DY+.

As already mentioned, Honey & Ward-Robinson (2002) provided a connectionist explanation of their findings (see 1.4.1 for a more detailed explanation of connectionist modelling). They argued that the reason cue equivalence occurs is because the mappings from A and B to the outcomes are mediated through the same hidden unit (see Figure 1.3). In order to achieve this in their model they allowed reciprocal connections between the hidden layer and the response layer (see Honey et al, 2010 for more details on how the model can be implemented). To provide further evidence in support of their model, in their second experiment they looked at what happened if the rat was primed, either with no food or food. They found that for compounds of stimuli that were treated equivalently in the biconditional discrimination, AB, compared to compounds of non- equivalents stimuli, AD, priming the rat with food produced more nose pokes into the magazine (a higher expectancy of food), compared to the situation where rats received priming with no-food. This they argued was because presenting the food has an impact on which hidden units are active, i.e. it primes those associated with food. This means that because A and B are activating just two hidden units (the one carrying their mapping to food, and the one carrying their mapping to no-food) with the addition of a prime to the food link this food unit will then govern behaviour leading to the rat to expect food. Whereas for AD, as the constituents are then activating all 4 hidden units, this distribution of input still fails to allow the rat to expect food.

Hodder, George, Killcross & Honey 's (2003) Experiment 2 used an allergy paradigm in humans to examine cue equivalence in a way that was similar to Honey & Ward-Robinson (2002) Experiment 1. Their first experiment compared congruent and incongruent transfer. Participants were originally trained, using a biconditional discrimination where two elements signaled the same thing e.g. AX+ AY- BX+ BY- CX- CY+ DX- DY+, where A and B interact with X and Y in the same way as do C and D. V and W then replaced X and Y. In the congruent case, when V and W were combined with A and B, they still gave the same outcome (i.e. AV produced the same outcome as BV), as did C and D (though these outcomes were the opposite of those for A and B). However, in the other, incongruent case, A and D now gave the same outcome, as did B and C. They found that participants were faster to learn the congruent case than the incongruent case, suggesting again that they were able to learn cue equivalence and generalize this to the transfer test with V and W. They took this as support for their connectionist model of learning.

Coutureau, Kilcross, Good, Marshall, Ward-Robinson & Honey, 2002 added a neural manipulation to their computational modelling. They showed that lesioning the entorhinal cortex but not the hippocampus affected rats' ability to acquire equivalence, but not their ability to learn the conditional discrimination. Additionally Iordanova, Killcross & Honey (2007) found that lesions to the medial prefrontal cortex showed the same effect. More specifically, Marquis, Killcross & Haddon (2007) showed that having no prelimbic prefrontal cortex (but not no infralimbic prefrontal cortex) impairs performance on these discriminations.

So it seems that there is some evidence that both humans and animals acquire equivalence and that this tendency to acquire equivalence can be accounted for by a connectionist model which creates configural representations.

1.3.5 Occasion setting

In addition to the nature of the conditional discriminations being learnt and acquired equivalence there is another literature which is reminiscent of the task-cueing paradigm – occasion setting (e.g. Bouton & Swartzentruber, 1986; Rescorla, 1985; Holland, 1986; Morell & Holland, 1986; Bonardi 1996; Bonardi & Jennings, 2009), where a contextual stimulus ("occasion setter") governs the UCS-CS relationship. Stimuli acting as occasion setters were thought to differ in 4 keys ways from a standard CS (Bouton & Swartzentruber, 1986). Firstly, they were not thought to enter into their own relationship with the US as there was limited responding to an occasion setter on its own (Rescorla, 1985). Secondly, they were also thought to be generated only in a situation where the occasion setter preceded the CS, and so the temporal training of an occasion setter required different parameters to those for an effective CS (Ross & Holland, 1981; Holland, 1986). Thirdly, it has been claimed that they have different summation properties from normal CSs, as if trained with AX- X+ where A occurs before X and so is thought to become an occasion setter, A does not become a conditioned inhibitor (e.g. Holland 1993, but see Rescorla, 1985). Fourthly, they show little effect of their previous training as an occasion setter when acquiring a new link to a US (Rescorla, 1985).

It is the temporal aspect of occasion setting which makes it particularly relevant to the task-cueing paradigm. As noted above, in a task-cueing paradigm the cue is often presented before the stimulus, and then compared to times when it is presented with (or just before) the stimulus in order to measure participant’s ability to prepare the task (see Section 2.2). This would mean that when the cue was presented before the stimulus it may become an occasion setter, whereas it would just be another stimulus when presented near-simultaneously with the stimulus (Holland, 1986).

Also of interest is the idea of occasion setting being similar to the way in which the tasks are implemented in the task-set control models described earlier. Occasion setters are proposed to mediate the link between the CS and the US – in a way similar but not identical to the way the task demand units are proposed to help activate an intermediate node between the stimulus input and the response in Brown et al’s (2007) model, – see Bonardi & Jennings (2009) for a brief description of a similar concept to that present in Brown et al’s model. It is this intuitive nature of the explanation of occasion setters (Bonardi & Jennings, 2009) which makes this a logical way to attempt to implement a model of task-cueing.

Early occasion setting research focused on feature negative (AX-, X+) and feature positive discriminations (AX+, X-) (e.g. Holland, 1986; Holland, 1989; Bonardi, 1996; Young, Johnson & Wasserman, 2000). Although reminiscent of the task-cueing structure, there is never a case when a stimulus is presented without a cue in this paradigm. More recent research has examined a structure which is more similar to the task-cueing paradigm and those structures used by Honey & Ward-Robinson (2002). Bonardi & Jennings (2009) trained rats in a biconditional discrimination with the same structure as the incongruent trials in the task-cueing paradigm – WC+,WD-, XC+, XD-, YC-,YD+, ZC-, ZD+, where + indicates food and – indicates no food. The key difference here from Honey & Ward-Robinson (2002) was that the cues (they called them occasion setters) were presented before the stimulus, and were not present at the time of stimulus presentation, whereas Honey & Ward-Robinson used contexts that were also present during stimulus presentation. They then paired one of the cues, W with a shock whilst simply exposing another one, Y. They then tested the generalization of the fear conditioning to both the equivalent cues and the stimuli when paired with food or not, e.g. C+, D+, C-and D-. They did not find significant evidence of cue equivalence, in that performance to X was not significantly different from performance to Z (although numerically there was more conditioned inhibition, see Honey & Watt, 1998, for evidence for this effect). They did, however, find more generalization of the conditioned inhibition to C+ and D- which were the outcomes when W (which had been devalued by the shock) had been present in training. The authors took this as support for occasion setters having a different status from normal CSs, arguing that a purely configural model, where XC has a link to the food, could not account for these findings. However, they also admit that the connectionist account given by Honey & Ward-Robinson could explain their data, and so there is no need to give occasion setters a different status.

Although the concept of an occasion setter is of interest, the early evidence for occasion setters having a different status from other CSs has been mixed, and mostly based on work with single layer associative networks. It seems from the descriptions of occasion setters that we have now that perhaps the concept is more similar to a task-cue when using task instruction, as the task in some ways sets the occasion for how to treat the new stimulus. Although it is common in the task-cueing literature to have cues which terminate before the presence of the stimulus, such as using auditory cues (e.g. Stevens, 2011), in the experiments presented in this thesis the cue remains visually present with the stimulus. As this type of temporal relationship may not necessarily encourage occasion setting (Holland, 1986 but see Holland, 1989 in which stimulus salience is also shown to play a role in occasion setting), I will not give the cues a different status in the models presented later in this thesis, but do acknowledge that this is in some ways an assumption that could be made.

1.3.6 What does the associative learning literature predict?

The existing evidence comparing biconditional and component discriminations (Saavedra, 1975; Livesey et al 2011) suggests that biconditional discriminations are much more complex to learn than component discriminations. This supports earlier evidence from the task switching literature, which suggested that performance on incongruent stimuli would be worse than performance on congruent stimuli, without participants necessarily having task knowledge (Stoet & Snyder, 2007). There is also reason to believe that, when cues or stimuli are treated equivalently, then their mappings are mediated in the same way. That is, they become functionally equivalent to one another. This is in agreement with arguments put forward by Schneider & Logan (2006), who suggested that even with two disparate cues per task, that these cues become equivalent to one another. The difference here is that the equivalence by task may only exist for the incongruent trials, as for the congruent trials all the cues are equivalent (in that they all require the same response to the stimulus) this differs from cues being equivalent, but mediated via tasks. In order to model both these effects it has been suggested by Honey & Ward-Robinson (2002) that a specialized connectionist network is required. It may be the case that a model such as Harris (2006) might also be able to account for the findings described above, however, the initial focus here will be on an approach similar to that taken by Honey & Ward-Robinson (2002). The section below will explain in more detail what is meant by a specialized connectionist model and discuss APECS as a suitable model for making predictions.

1.4. An introduction to connectionist modelling.

Earlier I mentioned that Honey & Ward-Robinson (2002) suggested that a connectionist model was one way to account for the ability to solve biconditional discriminations and to acquire cue equivalence. But what is meant by a connectionist model? The seminal books on connectionist modelling in the 1980s that set the agenda for much of the field thereafter were the two volumes of *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* published by James McClelland, David Rumelhart and colleagues in 1986 and 1987. These books paved the way for many interesting variants of connectionist model to be developed, e.g. simple recurrent networks or SRN (Elman, 1989), augmented SRN (Cleermans & McClelland, 1991), APECS (McLaren, 1997).

Connectionist models are defined as sharing the following properties: a set of processing units with states of activation, an output function for each unit, a pattern of connectivity among units, a propagation rule, an activation rule and a learning rule (Rumelhart, Hinton and McClelland, 1986). In addition to these properties, the specific use of a back propagation algorithm and the inclusion of a hidden layer have become common features in most connectionist modelling.

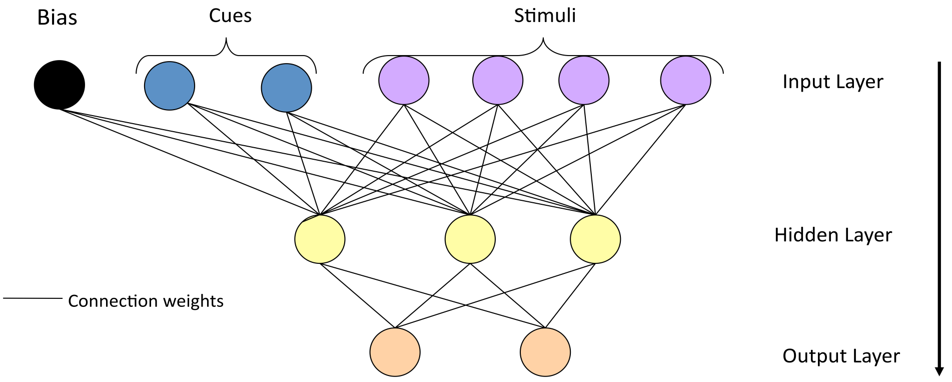


Figure 1.4 shows a typical structure of a connectionist network.

Figure 1.4 is designed to illustrate a standard structure for a connectionist network. In these feed-forward networks, every unit in the input layer is connected to every unit in the hidden layer, which in turn is connected to every unit in the output layer. The number of units in each layer varies between different incarnations of the model. Typically the activations of the input layer are set by the programmer. The pattern of activation across these units (and the desired pattern across the output units) is used to represent the input-output relationship to be learned. However, one input unit it always kept turned on and this is known as the bias for the hidden layer. The activations of these units are then propagated forwards to the hidden layer by multiplying the activations of each of the input units with the strength of their connection to the hidden layer unit in question.

$$input_j = \sum_{1}^{i} A_i W_{ij}$$

Where input is the input to a unit, A is the activation of the unit and W is the weight connection that unit to the unit in question.

The input to each unit is then calculated using a non-linear activation function:

$$a_j = \frac{1}{\mathrm{e}^{-input_j} + 1}$$

This process is then repeated for the connections between the hidden layer and the output layer. The network then uses back propagation in order to try to reduce the amount of error in the system. To do this, the error between the target activation and the calculated activation of the output units needs to be found.

$$\Delta o_j = (At_j - Ao_j)(1 - Ao_j) Ao_j$$

Where Δo is the output layer error, At is the target output activation and Ao is the output activation.

This can then be used to calculate the change in the connection strength between the hidden units and the output units.

$$\delta W(h-o)_{ij} = S(\Delta o)_j (Ah)_i$$

Where W(h-o) is the weight between the hidden and output layer, S is the learning rate parameter (a typical value is 0.01), and Ah is the hidden layer activation.

The problem then is how to assign the error to the hidden units. This is known as the credit assignment problem and a common solution is to use the following formula:

$$\Delta h_i = (\Delta o_j W(h-o)_{ij})(1 - Ah_i)(Ah_i)$$

Where Δh is the error at the hidden layer. This is the essence of the backpropagation algorithm.

The weights between the input and the hidden layers can then be adjusted in a similar way to the weights between the hidden and the output layers.

$$\delta W(i-h)_{ij} = S(\Delta h)_j (Ai)_i$$

Where W(i-h) is the weight between the input and hidden layer, S is the learning rate parameter (e.g. 0.01), and Ai is the input layer activation.

Over time a network of this type will come to solve a wide variety of problems presented to it (Rumelhart & McClelland, 1986). Its lack of specificity means that it can be used in many different fields in psychology and thus has the power to explain many different phenomena. However, there are some problems that the structure outlined above cannot solve, e.g. cue equivalence (Honey & Ward-Robinson, 2002) and the sequential learning problem (Barnes & Underwood, 1959) (more detail below). In order to solve these problems we need to slightly modify the network in a similar way to that suggested by Honey & Ward-Robinson (2002). The modification used in this thesis is McLaren's (1993) APECS model.

1.4.1 APECS

Although the model presented by Honey and Ward-Robinson (2002) has merit in learning biconditional discriminations and acquiring cue equivalence, it was not formally implemented in their paper and was not the first model of this form to be developed. McLaren (1993) developed a model of a similar class. This model was originally developed to solve the problem of catastrophic interference (McCloskey & Cohen, 1989). This problem was identified by McCloskey & Cohen (1989) when they tried to model a sequential learning experiment originally run by Barnes & Underwood (1959). In Barnes & Underwood’s experiment participants were first asked to learn one list of consonant-vowel-consonant (CVC) – adjective pairs (list 1), e.g. dax-regal. Then they were asked to learn a second list (list 2) which used the same CVCs but now paired with new adjectives, e.g. dax-sleek. They found that learning list 2 interfered with the learning of list 1 so that performance on list 1 dropped to about 50% after learning list 2. However, participants did show some retention of list 1.

McCloskey & Cohen (1989) found that when they trained a standard feed-forward connectionist network on this problem that, after learning list 2, the model had no memory of list 1. In other words the learning of list 2 had "catastrophically interfered" with list 1. In their paper they tried various ways of reducing catastrophic interference. They tried distributed coding for the inputs and outputs instead of localist coding, they tried a slower learning rate parameter, overtraining list 1, using more hidden units and freezing the weights for list 1. None of these prevented catastrophic interference from occurring. This effect was so powerful that it even occurred with just one run through of list 2, even though list 1 had taken hundreds of training trials to learn. The only way in which the model was able to learn both list 1 and list 2 was for the learning of the two lists to be intermixed.

McLaren’s (1993) Adaptively Parameterised Error Correct Systems (APECS) model was developed as a way round this problem. It does this by ensuring that list 1 learning takes a form that survives list 2 learning. To do this the model selects one (or a few) of the hidden units to carry the mapping by giving it (or them) a higher learning rate. This makes the learning more discretely coded at the hidden layer level. The model also gives these selected units a higher negative bias after learning involving them. This helps to preserve their learning. These changes mean that after training on dax-regal the model develops a structure as in the left hand side of Figure 1.5.

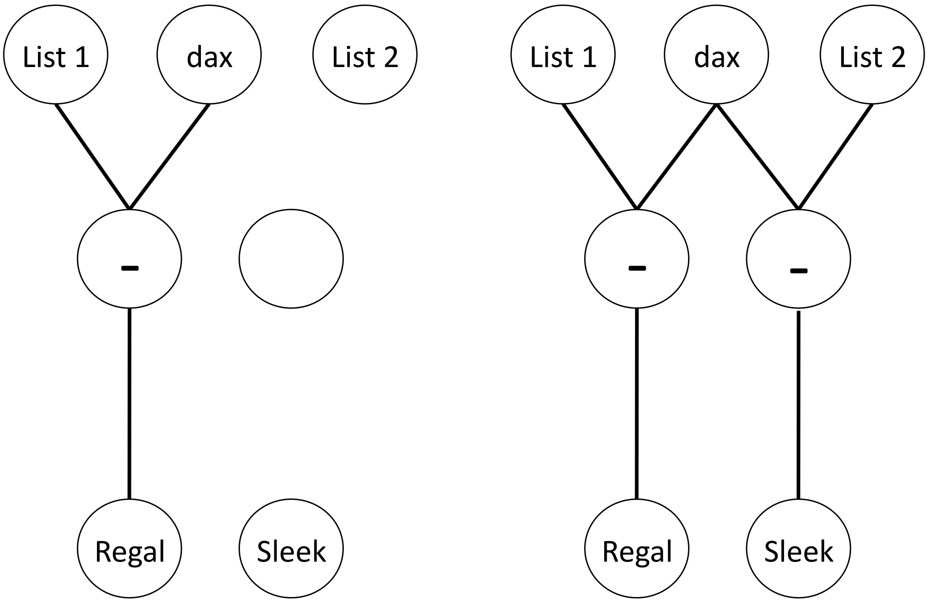


Figure 1.5 shows APECS's solution for the learning of a paired associate (dax-regal) from list 1 on the left and then the subsequent structure when it is then asked to learn dax-sleek from list 2.

These changes mean that when the model learnt about list 2, the learning about list 1 was protected. When dax-sleek is now presented in list 2, on its own it does not have a strong enough link with the hidden unit to overcome the bias and so does not activate the hidden unit, and hence the link to regal. As this link is not active and hence creating a large prediction error by wrongfully predicting regal when in fact the correct response is now sleek, the error-correcting learning system will not erase this learning. This means that learning dax-sleek will be undertaken by a different hidden unit to the learning of dax-regal, see Figure 1.5.

In addition to the experiments described above APECS has also been shown to be able to model:

- Backwards blocking (AB+ |A+ |B?) and un-overshadowing (CD+ |C- |D?). In backward blocking presentation of A+ after presentation of AB+ reduces creates reduced responding to B, compared to the case where AB+ had been followed by A+/- training. Le Pelley & McLaren (2001) were able show that APECS could account for these findings. Their interpretation of why APECS is able to account for these findings will be returned to in Chapter 5.

- Dickinson & Burke's (1996) experiment that examined the role of within compound associations in backwards blocking. Le Pelley & McLaren (2001) showed that APECS could account for the findings without the need to posit within compound associations.
- Exclusive OR (otherwise known as negative patterning) McLaren (2011)
- Surface dyslexia – APECS can be damaged in such a way that it will regularize the pronunciation of irregular words, e.g. pint. McLaren (2011)
- The context specificity of extinction found by Bouton & King, (1983), McLaren (1994, 2011).
- Second order backwards blocking (BC+| AB+| A+| B? C?) – here APECS suggested that performance to C would mimic performance to B. This was found to be the case in data briefly reported in McLaren, Forrest & McLaren (2012).

Given that APECS has been demonstrated to learn a wide range of varied psychological phenomena and the existing evidence suggests that it will be able to solve a biconditional discrimination and acquire cue equivalence – it seemed sensible to pick it as a candidate to use to try and model performance in a task-cueing paradigm when participants are told to learn the cue-stimulus response mappings. The version of APECS considered here is the same as the one used in both McLaren (2011) and McLaren, Forrest & McLaren (2012) and the set of parameters has been kept the same.

1.4.2 Real-time associative models

Although connectionist networks like the one described above can be used in many different areas of psychology, there are certain aspects of experiments which they are unable to simulate unless further modified. For example, typically the networks work at a trial level with a one-shot computation. That is, for each trial, the stimulus input is presented to the network, the output calculated and then the network moves on to the next experimental trial. Thus, there is no room to include within-trial temporal effects, for example, temporal discrimination effects (Kehoe and Boesenberg, 2002) or effects of CSI or ITI in task switching (e.g. Monsell 2003). Although APECS is typically run with more than one cycle through the network, this does not make it a real-time model as there is no analogy between a cycle and time nor are the activations within the network determined by their previous state, i.e. the activation of a hidden layer unit is determined purely by the input it receives from the input layer and not its activation on the previous cycle.

The problem of encoding time properly in a connectionist network usually addressed by allowing the activation of given unit to not only be determined by the input its receiving but also by its previous activation, e.g. the task demand and output units in Gilbert & Shallice's (2002) model or Brown et al (2006)'s model or AMAN (Harris & Livesey, 2010; Thorwat, Livesey & Harris, 2012). The precise nature of the recursive formula used varies between the models, but the principle of how time is encoded remains the same. There is an attempt to create a more realistic version of an APECS model in Chapter 6, and so further issues regarding real-time modelling (such as an appropriate decision mechanism) will be discussed in further detail in Chapter 6.

1.5. Concluding remarks

The above review suggests that much may be gained by examining and modelling the performance of participants under CSR instructions in a task-cueing paradigm with bivalent stimuli. To start with, the comparison of performance under CSR and Tasks instructions provides a test of Logan & Budensen's (2003) claim that learning cue-stimulus response mappings is all that is necessary to explain the patterns of performance seen in task-cueing paradigms. If (as turns out to be the case) performance under Tasks and CSR instructions is very different, a number of useful consequences and developments follow. From the perspective of cognitive control researchers, this outcome defends the task-cueing paradigm (under standard instructions) as a way of exploring task-set control. From the perspective of associative learning theorists, participants' ability to perform using a CSR strategy requires explanation, and prompts an exploration of the applicability of associative learning account to performance in the CSR condition, and its relation to animal and human behaviour in certain formally similar learning paradigms. For cognitive science more generally, it raises larger issues of the relationship between higher-order cognitive control and (processes mediated by) lower-level associative learning. For example, while participants perform in the task-cueing paradigm using a task-set strategy, do they also implicitly learn the CSR associations? What are the advantages of a hierarchical control strategy (select task, select response) over a one-step CSR strategy? Is performance of the CSR strategy mediated by "dumb" associative learning, or could it be mediated by a propositional strategy different from, but just as "controlled" and high-level as that used to perform the Tasks strategy.

Chapter 2: Initial investigation into a CSR approach

This chapter presents two initial investigations into what happens when participants perform a task-cueing experiment in circumstances in which either they are given standard task-cueing instructions supplemented by encouragement to learn the CSR rules, or they are merely instructed to learn the CSR rules presented in a list. The first experiment is an unpublished experiment performed by Monsell and colleagues at Exeter.[3] It is reported and reanalysed here as, although it used a rather weak instructional manipulation, the results were striking enough to motivate my own more extreme manipulations of instruction and training reported in this thesis, and core features of its design are used in my own experiments. Together with two of my experiments, it forms a set of three reported in a submitted paper (Forrest, McLaren and Monsell, under revision).[4]

2.1 Introduction

The experimental design used is based on the standard task-cueing paradigm. In this paradigm participants are instructed and trained in two or more tasks afforded by a set of stimuli. Then, on each trial, a stimulus from that set is presented, preceded by a task cue indicating which task the participant is to perform. The participant is required to make an appropriate response as rapidly and accurately as possible. They may then receive accuracy feedback, there is a pause, the next cue is presented, and so on. The experiments I report in this chapter required participants to classify digits as odd/even or high/low based on arbitrary shape or colour cues.

In the task-cueing paradigm used in these experiment, under standard task-cueing instructions participants were told, for example, that if the background is a diamond then they should classify the digit they then saw as odd/even, where odd required a response with the left key and even with the right key. However, if the background was a square they should classify the digit as higher/lower than 5, with high mapped to the right key and low to the left key. The right hand side of Figure 2.1 illustrates the task structure of the experiment. However, participants did not logically

---

[3] This experiment in turn was based on a pilot experiment conducted by Monsell and Rahamtulla.

[4] The paper was initially submitted to Journal of Experimental Psychology: General, and is currently under revision for Journal of Experimental Psychology: Learning Memory and Cognition.

need to use this task structure to know how to respond, as the shape and the number combination by themselves were completely predictive of the required response, e.g. a 2 on a square background will always require a left response; hence it is logically possible for participants in such an experiment not to use the tasks at all provided they have had adequate experience with all the CSR combinations they will encounter. The two experiments reported in this chapter are an initial investigation comparing a group of participants who are instructed to use the tasks (the Tasks group) with another group given instructions intended to encourage a CSR strategy.

2.1.1 Task-cueing effects

In order to examine whether, and in what ways performance differs between the two conditions we will focus on the three core phenomena described in Chapter 1: the switch cost, the reduction in the switch cost with time to prepare (RISC effect) and the congruency effect (Monsell, 2003, Kiesel et al, 2011).

It is typically found that when participants change from performing one task to performing another task there is a switch cost; participants are generally slower and less accurate on a task-switch trial than a task-repeat trial. Participants are also able to reduce this switch cost when they are given more time to prepare the task-set, i.e. when there is a longer time between the cue (coloured circle) appearing and the stimulus (number) appearing the switch cost declines. Theoretical accounts of the switch cost and its reduction with preparation were reviewed in Chapter 1.

In the task-cueing experiment described used here, the responses for two tasks were mapped onto the same keys, i.e. the left key represented odd and high, whilst the right key represented even and low. This means that for some numbers the response is always the same regardless of the task in play, e.g. 1 always requires a left response; these are called *congruent* stimuli. For other numbers the response depended on the task in play, e.g. 4 requires a left response if the task is high/low but a right response if the task is odd/even; these are *incongruent* stimuli. Typically, it is found that participants are faster and more accurate for congruent than incongruent stimuli. As with the switch cost and reduction in switch cost (RISC) effect theoretical accounts of the congruency effect were reviewed in Chapter 1.

2.1.2 The manipulation and predictions

The main manipulation was an instructional manipulation by which some participants were encouraged to use the aforementioned task-sets whilst others were either not informed of the task-sets or encouraged to use the cue-stimulus response mappings instead. Why might this manipulation be of interest? Because this manipulation takes seriously the general proposition that CSR learning, or learning how to respond without reference to task-sets, is a plausible strategy for performance in the typical task-cueing paradigm. (In our experiments, we did our best to make it a viable strategy by keeping the number of stimuli and cues small.) These experiments test the notion that if participants are in fact using a CSR strategy in the Tasks condition and so are not actually relying on the task-sets (as implied by the compound-retrieval model (Logan & Bundesen, 2003)), we would not expect any qualitative difference between conditions, though participants might learn CSR associations faster if we explicitly push them to do so. If we obtained evidence suggestive of CSR learning in the CSR condition, and a qualitatively different pattern of results in the Tasks condition, this would have suggested that performance in task-cueing cannot be simply reduced to CSR learning. It would also suggest measures diagnostic of these two strategies, and allow us to probe their relative efficiency.

Chapter 1 made some suggests about the pattern of behaviour we might expect to find if participants were learning how to respond using associative mechanisms. This suggested that there might be a large congruency effect (due to the relative difficulty of learning a biconditional discrimination compared to a simple SR mapping), a small switch cost if any (perhaps due to cue processing times) and potentially little or no effects of having time to prepare on performance (unless due to extended cue processing time or possible occasion setting which was unlikely as the cue and stimulus are still concurrently presented).

2.1.3 Transfer to a new stimulus set.

In addition to examining the switch cost, RISC effect and congruence effects, these two experiments also included a manipulation especially relevant to assessing stimulus-specific learning. Towards the end of each experiment, during which participants had experienced only four digits, the stimulus set was changed to a new set of four. This manipulation should have a more detrimental effect on CSR-based performance than on performance mediated by task-sets. If one is applying categorical task rules, they will continue to apply in the transfer test, and the participant must merely get used to applying them to a novel set of stimuli; this might be expected to slow performance transiently, but not differentially on task-switch and task-repeat trials. However, in the

CSR condition new stimuli require the learning of new S-R and CSR associations. This leads one to expect a more dramatic impairment on the transfer test in the CSR condition. One would also predict quicker post-transfer recovery for the congruent stimuli, for the same reason that one predicts faster learning of the congruent stimuli during training. However, there is no reason to expect performance on the transfer test to regress all the way back to the level seen at the beginning of the experiment. Training optimizes multiple aspects of performance such as efficiency of stimulus encoding and response readiness. Thus, we would expect the decrement in performance resulting from the need for new learning in the CSR condition to exceed and differ in pattern from the transitory adjustment to new stimuli expected in the Task condition.

The two experiments initiated the comparison between groups receiving standard task-cueing instructions (the *Tasks* group) to groups who received the same trials, but were encouraged to learn and use CSR associations (*CSR group)*. In the Monsell Experiment, the CSR participants were given the same standard task-set instructions as the Tasks group, and were merely exhorted, in addition, to learn S-R and CSR mappings. However, this meant that participants in the Monsell Experiment did have knowledge of the tasks. In Experiment 1, the instructions given to the CSR group made no reference to high/low and odd/even classification tasks; participants were simply told how to respond on the basis of individual CSR mappings, made available as a list for learning at the beginning of the experiment.

2.2 The Monsell Experiment

On each trial the participant saw a task cue (a square or diamond filled with pink or blue) in the center of the screen. After a 100 or 1200 ms CSI, one of four digits appeared superimposed on the coloured shape. The task, an odd/even or a low/high classification, was specified by the colour of the cue for some participants, the shape for others. Either the shape or colour changed on every trial, in such a way that the task changed on half the trials. All the participants received standard task-set instructions; half were in addition encouraged to learn the CSR mappings.

2.2.1 Method

2.2.1.1 Participants

The participants were 48 psychology undergraduates (age range 18-35) at the University of Exeter. Participants were paid for their participation.

2.2.1.2 Apparatus

The experiment was programmed in SuperLab (v1.5) and presented on a Macintosh Performa computer with a 13″ screen. Participants were tested individually.

2.2.1.3 Design

The experiment had a between-subjects factor of instruction (*CSR* versus *Tasks*), and within-subjects factors of trial type (*switch* versus *repeat*), response congruency (*congruent* versus *incongruent*), and CSI (*long*: 1200 ms versus *short:* 100ms). After practice and instruction, a session consisted of 14 blocks of 61 or 63 trials. For the first 10 of these blocks participants saw only four of the eight digits (1,4,7,8 for half the participants, 2,3,6,9, for the rest). For blocks 11-14, the stimulus set changed to the other four digits, as a transfer test. The trial sequences were fully randomized with the following constraints. Switch and repeat trials were equally frequent. The lengths of runs of trials on the same task were controlled to approximate within a block the expected distribution of run lengths for a binomial sequence with p(switch)=0.5. Each digit appeared equally often on switch, first repeat and second repeat trials in a run, and then randomly on other trials with the constraint that each digit appeared roughly equally often during the experiment.

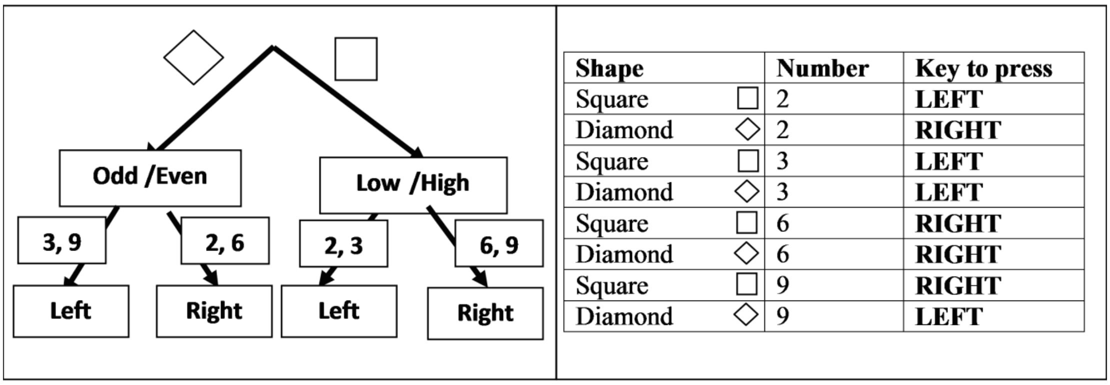
2.2.1.4 Stimuli

The task cue was a blue or pink diamond or square, displayed centrally. On the next trial one of the dimensions of the cue changed to the other value, so that there were no immediate exact repetitions of the cue; for half the subjects in each group the colour signaled the task, for the other half the shape. The side of the square was 170 pixels (~5.2° of visual angle); the diamond was the same shape rotated by 45°. The blue shapes had RGB values of (153, 204, 255) whilst the pink shapes had RGB values of (255, 153, 153). The digit was displayed in the center of the shape in Geneva font, and was 56 pixels (~1.7°) high.

2.2.1.5 Procedure

Participants sat with their eyes about 50 cm from the computer screen, left and right index fingers on the “z” and "/" keys of the computer keyboard. All participants were given onscreen instructions that introduced the odd-even and high-low classification tasks, and explained how they were cued.

Participants were first given four single-task practice blocks of 20 trials, one with a long CSI and then one with a short CSI for each task. There followed two blocks of 30 trials to practice switching between the tasks, first with a long CSI and then with a short CSI. During the practice blocks all 8 digits appeared. After practice the CSR group of participants received additional instructions designed to encourage them to learn and use “pairings” such as "Square + 2 → RIGHT". Half the participants were assigned to the CSR condition and the other half to the Tasks condition. As previously mentioned Figure 2.1 illustrates the mappings used under both conditions in a form that emphasizes either the hierarchical Tasks structure or the "flat" list of CSR mappings, for a participant for whom the cue's shape was the task cue.



Shape	Number	Key to press
Square □	2	LEFT
Diamond ◇	2	RIGHT
Square □	3	LEFT
Diamond ◇	3	LEFT
Square □	6	RIGHT
Diamond ◇	6	RIGHT
Square □	9	RIGHT
Diamond ◇	9	LEFT

Figure 2.1. The left hand panel indicates the hierarchical structure of the task-set strategy. In this example, square versus diamond cue specifies odd/even versus high/low classification of the digit; the next or the low/high task; the next level specifies the response assignments for the selected classification. The right hand panel gives the CSR mappings offered to participants in that condition.

Participants then began the main part of the experiment. CSI alternated by block, with the starting CSI balanced across participants. In order to keep the response-stimulus interval (RSI) constant at 1700ms, the response-cue interval was 1600ms in the short CSI blocks, and 500ms in the long-CSI blocks. If the response was correct, the next trial began immediately. If the participant pressed the wrong key, “WRONG!” was displayed (in red) for an extra 1500ms, before the experiment continued; if he/she failed to respond within 4000ms, “Time Out” appeared (in black). At the end of the block the participants’ mean RT and number of errors for that block were displayed. Participants rested between blocks for as long as they wished, and were advised to stretch their legs after block 8. Participants were informed after block 10 that the numbers they were going to see would be different.

2.2.2 Results

In the two experiments in this chapter, trials following an error were excluded from the analysis, and error trials were excluded from the RT analysis. Where departures from sphericity may be an issue throughout this thesis the Huyn-Felt corrected F and p values are given, but with the uncorrected degreed of freedom.

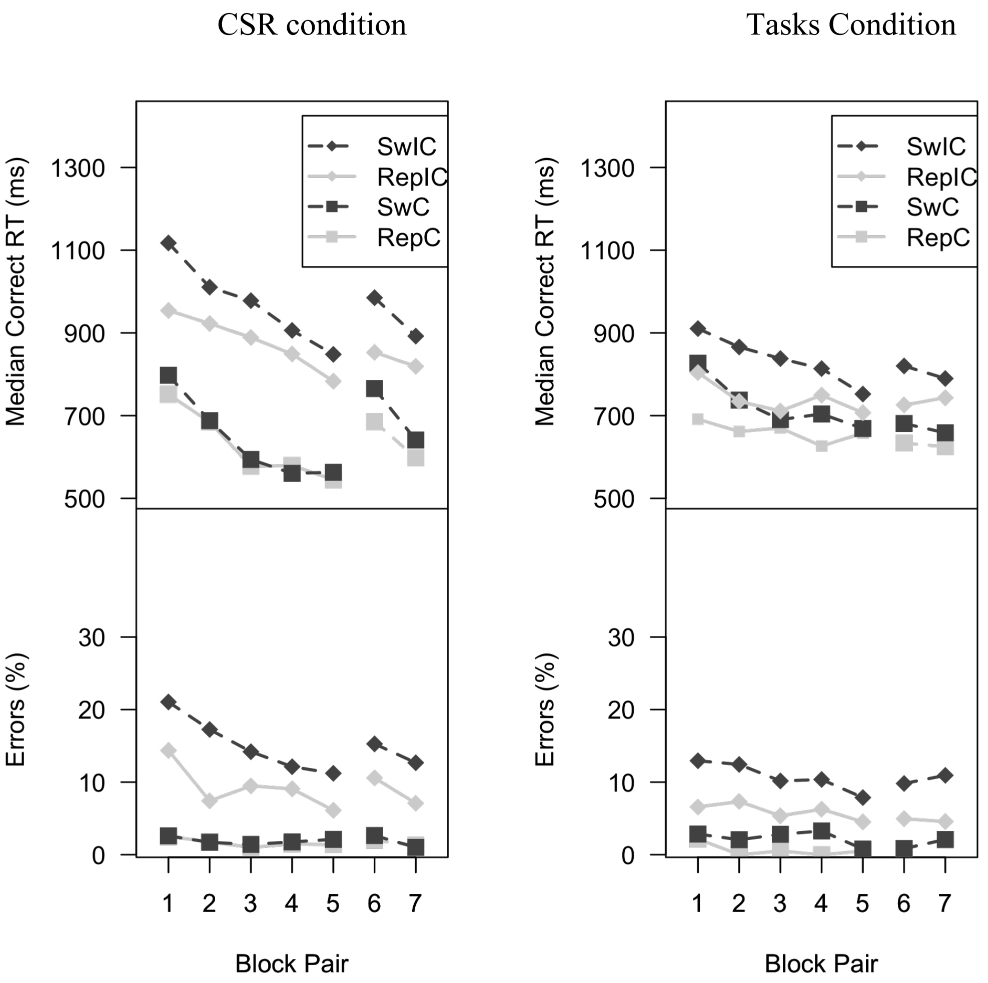


Figure 2.2 shows the means of median correct RTs and error % over successive block pairs in the Monsell Experiment, for each combination of congruent/incongruent stimuli and switch/repeat trials. Darker dotted lines represent switch trials, lighter solid lines are repeat trials. Square symbols represent congruent trials, diamonds represent incongruent trials. The breaks in the lines mark transfer to a new stimulus set for block pairs 6 and 7. The left hand panel shows results for the CSR condition; the right hand panel shows the results for those participants in the Tasks condition.

As illustrated in Figure 2.2, participants in both the Tasks and CSR conditions showed an improvement in performance from block pair 1 to 5. (Each block pair comprised a short-CSI and a long-CSI block, with their order balanced across participants). Although the CSR group, who received instructions to learn, started off slower and less accurate than the Tasks group, by the end of block pair 5 there was little difference in overall performance between the groups, with the CSR group having a median correct RT of 685ms and 5.21% errors, compared to 676ms and 4.83% errors for the Tasks group. Figure 2.2 also illustrates how performance was perturbed by the introduction of novel stimuli at block pair 6 (the transfer test). The perturbation appears more marked and persistent for the CSR group. Each of the effects and interactions (switch cost, RISC effect and congruency effect) will now be considered in turn, first for the main experiment (block pairs 1-5), then for the transfer blocks.

2.2.2.1 Block pairs 1-5

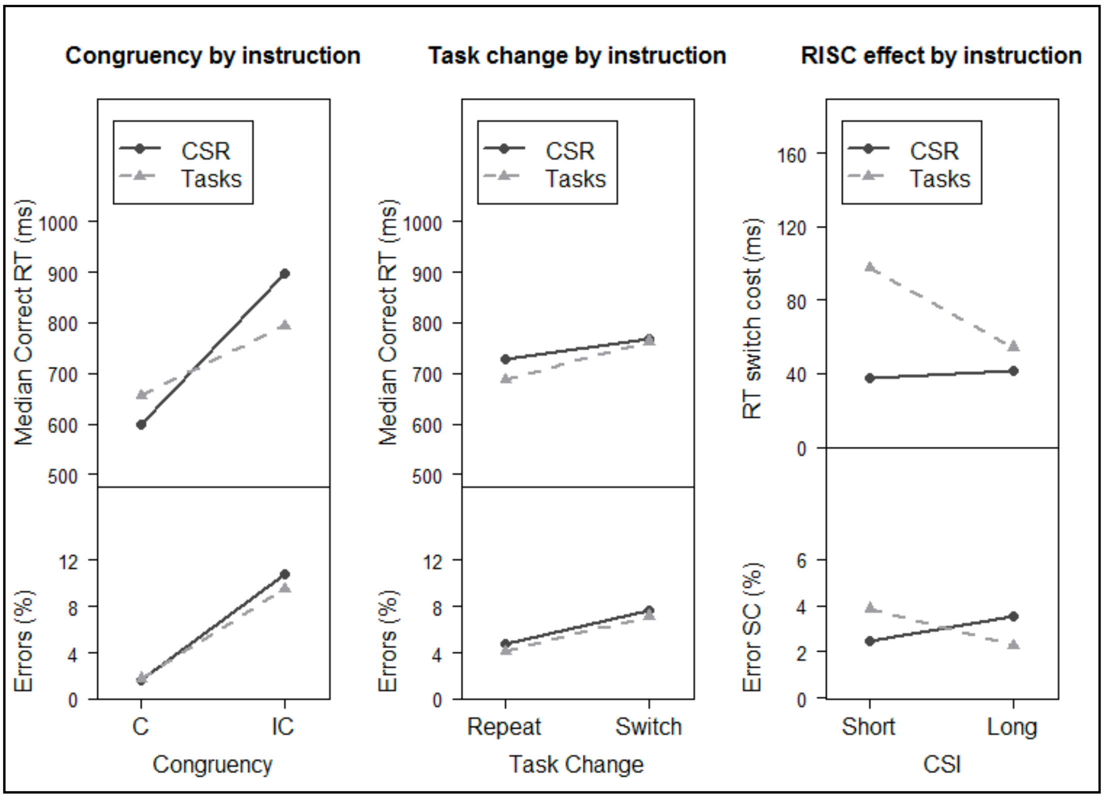


Figure 2.3 shows contrasts indicating three principal findings of the Monsell Experiment. Reaction time contrasts are shown in the upper panels, error rate contrasts in the lower panels. The three sets of panels show interaction of instruction group with the effects of stimulus congruence (left), a task switch (middle), and the reduction in switch cost with increasing CSI (right).

To test the critical interactions of instruction condition with effects of a task switch, congruency and preparation time, a mixed ANOVA with the factors: task switch (repeat, switch), congruency (congruent, incongruent), CSI (long or short), block pair (1,2,3,4 or 5) and instruction condition (Tasks versus CSR) was run on both the median correct reaction time (RT) and percent errors, and followed up by separate ANOVAs for the two groups where appropriate. Salient contrasts are shown in Figure 2.3.

Task switches and instruction. The Tasks group had a larger RT switch cost (76ms) than the CSR group (40ms), F(1,46)=6.76, $p<0.05$, but the switch costs in both the Tasks condition, F(1,23) = 69.8, $p<0.001$, and the CSR condition, F(1,23) = 34.9, $p<0.001$ were reliable. For errors, the overall main effect of a task switch was reliable, F(1,46)=39.7, $p<0.001$ (repeat: M=4.42%, switch: M= 7.45%), but there was no interaction with instruction: the switch cost was 3.06% for the Tasks group and 3.01% for the CSR group. Hence, as expected, the switch cost over these blocks was greater in the Tasks than in the CSR condition (see Figure 2.3).

Preparation and instruction. As Figure 2.3 shows, preparation reduced the switch cost in the Tasks group RT from 97ms in the short-CSI blocks to 65ms in the long CSI blocks. There was no reduction in the CSR group; their switch cost was 37ms in the short-CSI blocks and 42ms in the long-CSI blocks. The three-way interaction was marginally reliable F(1,46)= 3.30, $p=0.075$. Separate analyses revealed a significant interaction of CSI and task switch in the Tasks condition, F(1,23)=5.39, $p<0.05$, but not in the CSR condition, $F<1$. For error rates, a similar pattern was evident. For the Tasks group there was a larger switch cost in the short-CSI (3.85%) than the long CSI blocks (2.26%). For the CSR group the switch cost was 2.45% for the short-CSI blocks, and 3.56% for the long CSI. This interaction was reliable, F(1,46)=4.96, $p<0.05$. There was a nearly reliable RISC effect for the Tasks group, F(1,23)=3.22, $p=0.081$, but none for the CSR group, F(1,23)=1.80, $p=0.19$, and the difference between the two groups is reliable. Hence overall there is evidence here to suggest a RISC effect in the Tasks group but not in the CSR group.

However, both groups benefited more generally from the opportunity to prepare, as we may see by examining task-repeat trials alone. The Tasks group had faster repeat RTs after a long CSI (627ms) than after a short CSI (747ms), F(1,23) = 30.1, $p<0.001$. A similar but less marked effect was seen in the CSR condition, where repeat trials were faster in the long CSI (714ms) than in the short CSI (744ms), F(1,23) = 832, $p<0.01$.

The difference in this effect between the two groups was significant, F(1,46) = 8.06, p<0.01. This effect was not significant in the errors for either group.

Congruency and instruction. RT and error rates showed a much larger effect of congruence (incongruent minus congruent) in the CSR group (299ms, 9.28%) than in the Tasks group (138ms, 7.74%); the interaction was highly reliable for RTs, F(1,46)=11.1, p<0.01, but not error rates. In separate analyses of the two groups, the congruence effect was reliable for both the Tasks group, F(1, 23) =32.2, p<0.001, for RT, F(1,23)=46.0, p<0.001 for errors, and for the CSR group: F(1,23) = 61.6, p<0.001, for RT; F(1,23)=60.9, p<0.001 for errors.

Acquisition effects. As may be seen in Figure 2.2, overall performance evidently improved from block pair 1 to 5, and the effect of block pair was reliable for RTs F(4,184)= 21.8, p<0.001. Although the RT improvement looks greater for the CSR group than the Tasks group, the difference was not reliable, F<1. For the error data the CSR group improved more from block pair 1 to 5 than the Tasks group, F(4,184) =4.10, p<0.05. Separate analyses revealed significant effects of block pair in both the Tasks, F(4,92)=2.73, p<0.05, and CSR groups, F(4,92)=10.0, p<0.001.

One might expect that congruent stimuli should be learned more rapidly than incongruent under CSR instructions. This is supported by the interaction between block pair, congruency and instructions in the error rates, F(4,184)= 3.47, p<0.05, driven largely by an interaction between block pair and congruency in the CSR group, F(4,92)= 7.84, p<0.001, but not in the Tasks group, F<1, and is also supported by the non-reliable interaction in the RTs (see Figure 2.2). For the CSR group, performance on the congruent stimuli appears asymptotic in the error rates relatively early on, and after only three block pairs in the RTs, while the much poorer performance on the incongruent stimuli continues to improve steadily across all five block pairs. For the Tasks group, there was more modest improvement with practice, and it was not obviously different for congruent and incongruent stimuli.

Other significant interactions. In task-switching experiments, the effect of incongruence is usually amplified by a task switch, and vice versa, and this would be expected if the irrelevant task-set is more active on switch than repeat trials. One might also expect this for the CSR condition as the cue is more relevant for incongruent than congruent trials. The switch cost was larger for incongruent (Tasks: 95ms, CSR: 75ms) than for congruent stimuli, (Tasks: 57ms, CSR: 4ms), F(1,46)= 18.5, p<0.001, and this effect was numerically much more pronounced in the CSR condition than in the Tasks condition, but the three way interaction was not significant, F(1,46)=1.66, p=0.20. Separate analyses revealed a significant interaction in the CSR group, F(1,23)= 59.7, p<0.001, but not in the Tasks group. In the error data the pattern was similar. Again there were larger switch costs in the incongruent trials (Tasks: 4.96%, CSR: 5.68%) than the congruent trials (Tasks: 1.15%. CSR: 0.34%). This did not interact with instructions. Separate analyses revealed a significant effect in both the CSR group, F(1,23)= 18.3, p<0.001, and the Tasks group, F(1,23) = 19.2, p<0.001.

Inasmuch as task-set preparation might be expected to reduce the impact of the irrelevant S-R mapping, one might expect a smaller congruence effect at the longer CSI. In the task-switching literature such an interaction has sometimes been observed, but sometimes not (Monsell & Mizon, 2006; Kiesel et al, 2010). In the present experiment, the congruence effect was indeed larger with a short-CSI (Tasks: 170ms; CSR: 315ms), than with a long-CSI (Tasks: 105ms; CSR: 283ms), F(1,46)= 10.7, p<0.01, but this change did not vary much with instruction condition. This effect was not significant in the errors.

2.2.2.2 The transfer test

In order to examine the effect of transfer from one set of stimuli to another, performance in block pair 5 was compared with performance in block pair 6, using a mixed ANOVA with the factors: task switch (repeat or switch), congruency, CSI, block pair (5 or 6) and instruction condition. Only factors that interacted with block pair will be discussed. This analysis does not capture all the effects of transfer, but is the most readily interpretable. For example, a comparison between performance on block pairs 1 and 6 would be ambiguous because of the large difference in baseline performance on block pair 1 between the CSR and Tasks groups. As performance in the two conditions in block pair 5 is similar, it provides a better baseline for comparing the impact of transfer to new stimuli across conditions. The caveats around our ability to interpret this analysis are inherent in the manner in which the transfer test was conducted.

As expected, the CSR group was more affected by the transfer to new stimuli; their mean RT (error rate) increased by 137ms (2.41%) as compared to the Tasks group's 55ms (1.15%). The interaction was significant in RTs, F(1,46)= 4.14, p<0.05, but not error rate, F(1,30)=1.36, n.s. although there was a main effect of transfer in the errors, F(1,46)=10.7, p<0.01. Under both instruction conditions, transfer to new stimuli increased the switch cost, by 48ms (1.17%) in the Tasks group, and increased the RT switch cost in the CSR group by 64ms but with a very slight decrease in the error switch cost (-0.23%). In the RTs there was a significant interaction between block pair and task switch, F(1,46)=16.6, $p<0.001$, but not in the error rate F(1,46)=0.711, n.s. This interaction was separately significant in RTs in the Tasks, F(1,23)= 9.70, $p<0.001$ and CSR groups, F(1,23)= 7.99, $p<0.05$, but did not reach significance for either group in the errors, $F<1.5$.

A further analysis showed that performance improved from block pair 6 (Tasks: 731ms, 5.98%, CSR: 822ms, 7.62%) to 7 (Tasks: 688ms, 5.89% CSR: 738ms, 5.52%) significantly in both the RTs, F(1,46)=18.3, $p<0.001$ and the errors, RT, F(1,46)=6.81, $p<0.05$. The improvement was larger in the CSR group than the Tasks group in the error data, F(1,46)=5.74, $p<0.05$.

2.2.3 Discussion

In the Tasks condition, as is standard for task-cueing experiments, participants were instructed to construe the problem as selecting the appropriate categorical rule to apply on the basis of the cue presented. In the CSR condition, these standard instructions were combined with a strong hint to try and learn stimulus-specific S-R and CS-R mappings. In this initial experiment it was found that even this relatively weak manipulation was effective in changing the nature of performance. The Tasks group showed the typical large switch cost, modest response congruence effect, and reduction in switch cost (and congruence effect) with preparation under these conditions that would be expected on the basis of findings reported in the literature on task switching (Monsell, 2003; Kiesel et al, 2010, Vandierendonck et al, 2010). The CSR group showed a smaller switch cost, no reduction in switch cost with preparation, and a much larger congruence effect. The initial performance of the CSR group was poorer, perhaps because they were simultaneously grappling with task-level and stimulus-level mappings, but they caught up by the end of the training phase. It looks as if the CSR participants rapidly mastered and used the simple S-R mappings for congruent stimuli, as performance on them was fast and asymptotic after three block pairs, and they showed no switch costs for these stimuli after just one block pair. Performance on the incongruent stimuli showed much more gradual improvement with practice, and continued to exhibit a switch cost, but this is not inconsistent with associative learning of the CSR mappings.

The transfer test, in which the stimuli were changed, also suggests that the CSR group were more reliant on associative retrieval by the end of training. The Tasks group show only a modest perturbation of performance on the transfer blocks pair, while the CSR group show a greater and more persistent impairment in performance, with switch costs reappearing for the congruent stimuli, and increasing substantially for the incongruent stimuli, though not all these trends are reliable.

Two conclusions can be drawn. First, associative learning of S-R and CSR associations can contribute to performance in a task-cueing paradigm with a small set of cues and stimuli presented over many trials, and using them provides – eventually – a reasonably effective strategy for producing good performance. However in this first attempt (by Monsell et al) to explore the contribution of CSR learning it certainly cannot be claimed that the CSR group were reliant on this associative strategy alone. They had been instructed in task-set level strategy as well, and the mere addition of a hint to use an associative learning strategy is a relatively subtle manipulation. Participants may well have varied considerably in the mixture of strategies they were using. Hence, in my subsequent experiments, training and instruction regimes were used that were designed to produce a better segregation of task-based and CSR-learning strategies.

The hint to learn associations did not merely accelerate an improvement in performance with training that would have happened anyway; it changed the pattern of performance qualitatively, reducing switch costs (especially for congruent stimuli), increasing the congruence effect, and abolishing the reduction in switch cost with preparation. Hence, a second tentative conclusion is that participants in our Tasks group, and by extension other task-cueing experiments, were not merely retrieving the response associated with each cue + stimulus combination, contrary to the proposal of Logan and Bundesen (2003). The data instead suggest that participants use the cue to select a categorical task-rule, and then apply it to the stimulus. Providing time for preparation allows the rule to be selected and readied for application in advance of the stimulus. This provides an additional answer to Dreisbach and Haider’s (2008) question of what benefit a task-set level rule provides when there are lower-level associations available: the task-set strategy affords effective preparation in conjunction with a reduction of the otherwise complex mapping of stimuli to responses to a simpler hierarchical relationship. And even when participants have no time to prepare before the stimulus, the hierarchical, two-step, strategy — select categorical rule, select appropriate S-R mapping — appears to be effective and seemed quite natural to our participants.

However, because in this initial study it is likely that there was a mixture of strategies in the CSR group, this experiment can only suggest features of the data diagnostic of associative versus task-set responding. For example, the possibility that the CSR group learned and used only S-R, and none of the CSR, associations cannot be ruled out — i.e. it is entirely possible that performance on the incongruent stimuli was based on application of categorical task rules. The fact that the switch cost for the incongruent trials in the CSR condition was as large (75ms) as the overall switch cost seen in the Tasks group (75ms) would be in line with this hypothesis. A cleaner manipulation of strategy is needed and this is what my first follow-up experiment aimed to do.

2.3 Experiment 1

In Experiment 1, the CSR group was directed to use stimulus-specific response rules, with no explicit mention of, or training on, the categorical task rules before participants were run in what we have been calling block pair 1. They were simply provided with a piece of paper listing the CSR mappings (in the same form as in the right half of Figure 2.1) and told to treat the first two blocks as practice blocks to acquire these rules. This also meant that when participants were transferred from one set of stimuli to the other set, they had not previously responded to these digits in this task (unlike in the Monsell Experiment). The Tasks group received exactly the same amount of exposure to the stimuli as the CSR group, but with standard task-cueing instructions. So for their first block pair, like the CSR group they received a piece of paper that contained the task rules (left half of Figure 2.1), and they also had no exposure to the transfer stimuli.

It remains possible for participants shown a list of stimulus-specific rules to infer that categorical odd/even and low/high rules contingent on the cue can be used, and for participants given task-cueing instructions to adopt a CSR strategy. I therefore debriefed participants on what they thought they were doing, and replaced participants in each group until we had two groups of 16 whose reports of their introspections matched the instructions in force for their group.

2.3.1 Method

2.3.1.1 Participants

The participants were 43 psychology undergraduates (Age range 18-35; average age 20.4 years, 7 male) at the University of Exeter. Participants took part for course credit and a bonus payment, which was contingent on their performance (average payment £2.11, range £1.75-£3.00).

2.3.1.2 Apparatus and stimuli

The experiment was programmed in Matlab (2008b) with Psychphysics Toolbox extensions version 3.0.8 (Brainard, 1997; Pelli, 1997), and run on an iMac computer. The participants were seated 50 cm from the 20″ screen and tested individually. Cues were coloured squares or diamonds of side 200 pixels (~6.1° of visual angle). The blue shapes had RGB values of (115, 194, 251) whilst the pink shapes had RGB values of (255, 192, 203). The set of stimulus digits used in the initial 10 block pairs was 1,4,7,8 for half the participants, 2,3,6,9 for the other half; in the transfer blocks the other set was used. Each stimulus digit was displayed in the center of the cue in Courier bold font size 60 (~1.3° of visual angle).

2.3.1.3 Design and Procedure

The design and randomization of the main part of the experiment were almost identical to those of the Monsell Experiment, with the small difference that which four digits were congruent and which were incongruent was counterbalanced by swapping the response mappings for half the participants for the low/high task. The procedure used differed from the Monsell Experiment only in the following important respects. First, there were no longer practice blocks. Second, participants in the CSR condition were at the outset given the CSR combinations as a list (as on the right side of Figure 2.1) and told to learn these mappings for the individual stimuli; they were told nothing about tasks. Participants in the Tasks group were given standard task-set instructions and provided with an instruction sheet in the form of a flow chart (as on the left side of Figure 2.1). Third, at the end of each block after the first pair of blocks, a score was calculated by adding 5 for each error to the mean RT in centiseconds for that block; participants were asked to try and beat (undercut) their previous score with that CSI to earn a bonus point worth 25p. Fourth, at the end of block pair 5, participants were given a piece of paper outlining, in the condition-appropriate fashion, the correct responses for the transfer stimuli. Participants kept this piece of paper for block pair 6, before returning it to the experimenter and then completing the final block pair.

Finally, participants completed a post experiment questionnaire, asking them how they approached the experiment, and in particular whether, in the case of the Tasks group, they had used the specific categorical task rules given them (in which case they were included) and whether, in the case of the CSR group, they had induced such categorical task rules at any point in the experiment (in which case they were excluded).

Although one might worry about the asymmetry of these exclusion criteria, the asymmetry is inherent in the theoretical relationship between introspective reports and the proposed properties of two systems: associative and controlled. If participants are using tasks in a controlled effortful manner it seems reasonable to assume that they can introspect on this and report if they were indeed making the effort to use these tasks (leaving aside the prospect that they might lie about their own introspections). Similarly if participants in the CSR condition induce the task-sets and use them to help their performance one would expect them to have conscious access to this process and access to whether or not they were intentionally using them. In contrast participants in the CSR condition who did not induce the task-sets may provide a rationale for a performance that could be driven by an underlying associative system in a variety of different ways (Nisbett & Wilson, 1977); they may not have introspective access to the way in which they were approaching the experiment.

Eight Tasks group participants were replaced because when asked “did you use the tasks to help you?” they said “no”, leaving 16 who said they did. Three CSR group participants were replaced because they mentioned using at least one categorical task rule when asked how they approached the experiment, leaving 16 who did not. The excluded participants will be discussed further below.

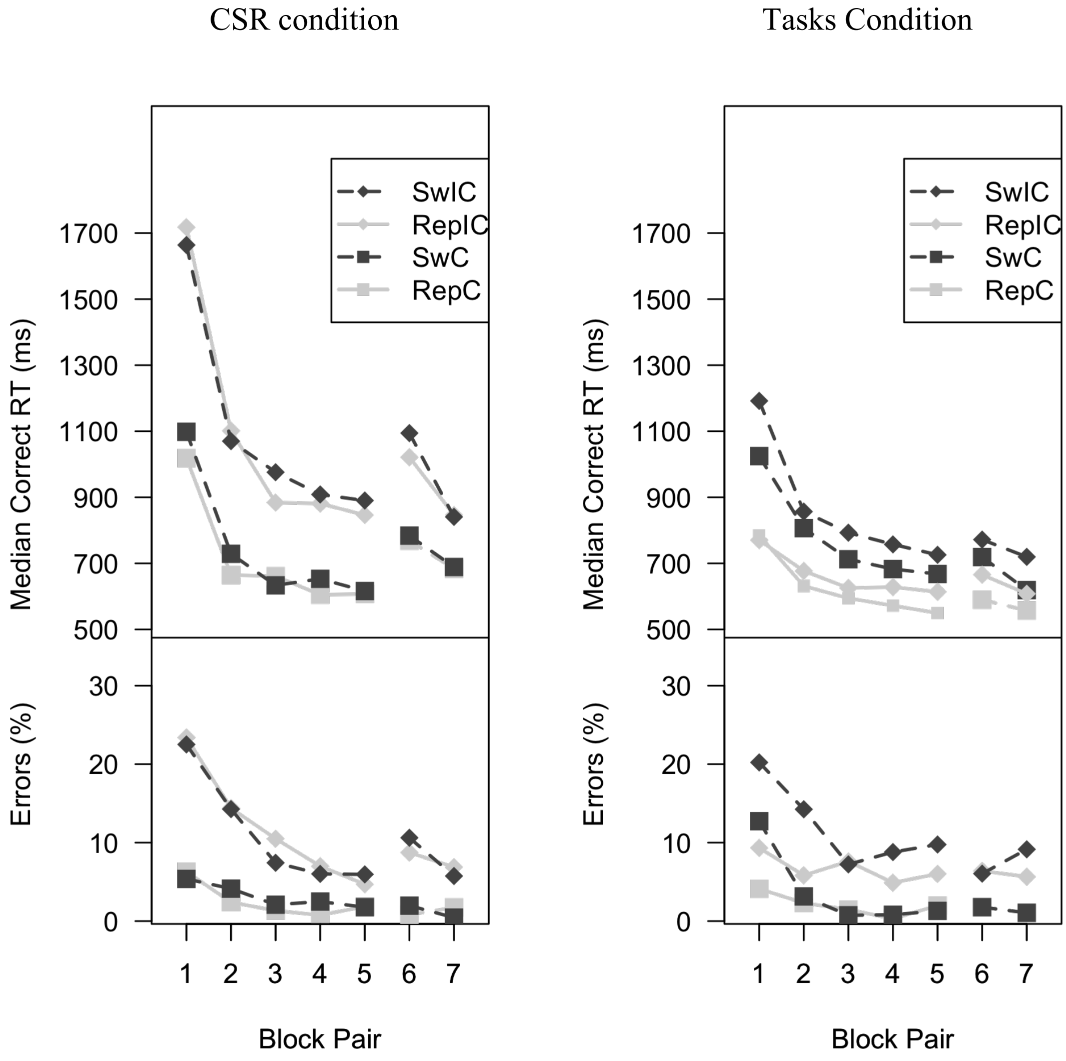


Figure 2.4 shows the means of median correct RTs and error % over successive block pairs in Experiment 1, plotted as in Figure 2.2.

2.3.2 Results

The data were analyzed much as for the Monsell Experiment. In Experiment 1, however, block pair 1 represents the initial encounter with these conditions, as there was no single task or task-cueing practice, and block pair 6 now represents the first encounter with the transfer stimuli. Both groups showed substantial improvement over the first 5 block pairs, with a marked improvement from the first to the second block pair. Once again performance was notably worse to start with in the CSR group, who had a relatively large number of stimulus-specific mappings to learn, than in the Tasks group, who had only the two sets of classification rules to learn. After five block pairs the difference was less marked (though still substantial), with the CSR group having a median correct RT of 740ms (3.57% errors), and the Tasks group 639ms (4.76%

errors). As in the Monsell Experiment, the CSR group's performance was much more impaired by transfer to a new set of stimuli in block pair 6, and the perturbation appears more persistent for that group. The analysis strategy was similar to that in the Monsell Experiment, except that the initial block pair, which now included the initial encounters with the stimuli and provision of instructional aids, was excluded.

2.3.2.1 Block Pairs 2-5.

Task switches and instruction. As Figure 2.5 shows, the Tasks group had a much larger RT switch cost (139ms) than the CSR group (29ms), F(1,30)=9.19, p<0.001. The switch cost in the Tasks condition was highly reliable, F(1,15) = 17.6, p<0.001, but the small effect in the CSR group was only marginally reliable, F(1,15) = 3.39, p=0.086. The error switch cost was also larger for the Tasks (1.95%) than the CSR group (0.15%), but this difference was only marginally reliable, F(1,30)= 3.28, p=0.080; there was a significant switch cost in the Tasks group, F(1,15)= 10.6, p<0.01, but not in the CSR group, F<1.

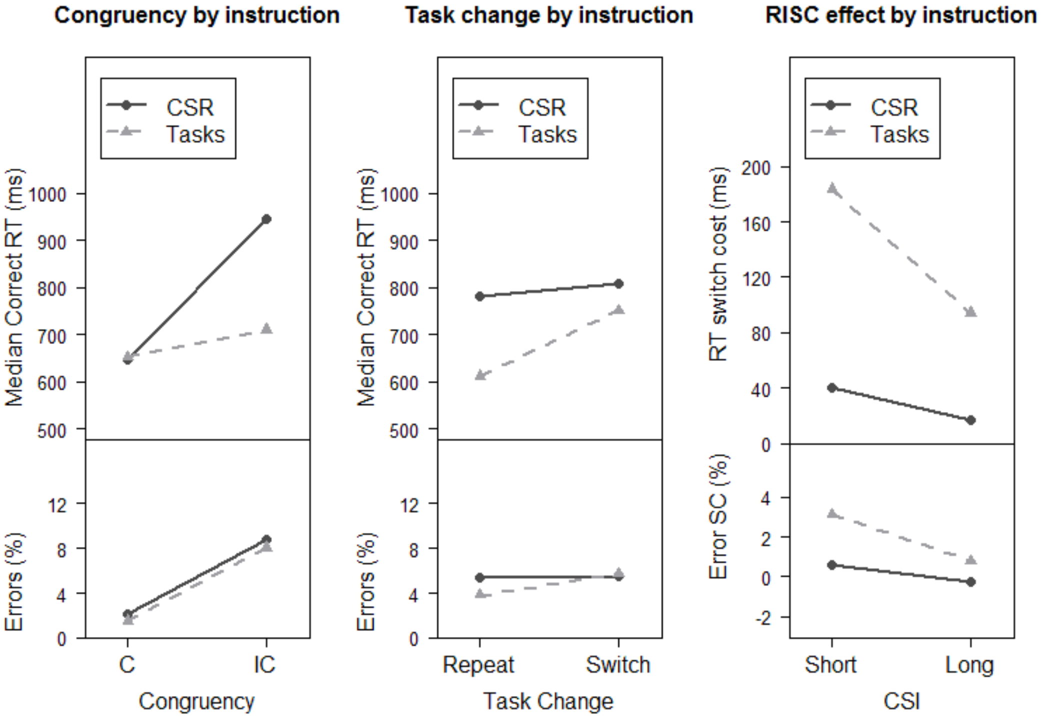


Figure 2.5 shows contrasts indicating three principal findings of Experiment 1, plotted as in Figure 2.3.

Preparation and instruction. As Figure 2.5 shows, preparation reduced the switch cost in the Tasks group from 184ms in the short-CSI blocks to 94ms in the long-CSI blocks, a reliable RISC effect, F(1,15)= 6.42, p<0.05. There was less evidence for such an effect in the CSR group, for whom the switch cost was 40ms in the short-CSI blocks and 17ms in the long-CSI blocks, F(1,15)= 1.91, n.s. However the three-way interaction was only marginally reliable, F(1,30)= 2.86, p=0.1. For the error rates, a similar pattern was present: for the Tasks group there was a marginally reliably larger switch cost in the short-CSI blocks (3.1%) than the long CSI blocks (0.8%), F(1,15)=4.05, p=0.063. For the CSR group the cost was 0.57% for the short-CSI blocks, and -0.26% for the long-CSI, F(1,15)<1. The 3-way interaction was not reliable, F<1.

Participants in the Tasks condition also showed a substantial overall preparation effect when just the repeat trials were considered, whereby they were faster in a long-CSI (531ms) than a short-CSI (692ms), F(1,15) =70.9, p<0.001. However, they also made more errors at the long-CSI (4.41%) than the short CSI (3.19%), F(1,15) =4.66, p<0.05. The CSR participants were also less accurate but only slightly faster at the long CSI (775ms, 5.67%) than at the short CSI (787ms, 5.06%). The interaction was not reliable, F<1.

Congruency and instruction. RTs showed a much larger effect of congruency (incongruent minus congruent) in the CSR group (differences: 299ms, 6.68%) than in the Tasks group (58ms, 6.54%); the interaction was highly reliable for RTs F(1,30)=21.5, p<0.001, but not error rates, F<1. In separate analyses of the two groups, the congruence effect was reliable for both the Tasks group, RT: F(1, 15) =12.6, p<0.001, errors: F(1,15)=26.3, p<0.001, and for the CSR group: RT: F(1,15) = 36.6, p<0.001, errors: F(1,23)=28.3, p<0.001.

Acquisition effects. As may be seen in Figure 2.4, overall performance evidently improved from block pair 2 to 5, and this was reliable for RTs, F(3,90)= 26.2, p<0.001 and the errors, F(3,90)= 10.5, p<0.001. The overall improvement in performance was not reliably different between the two groups.

The effects of task switch and the RISC effect were relatively stable across block pairs 2-5. However, the congruency effect in the CSR group changed across block pairs (as with the Monsell Experiment), as indicated by the interaction between block pair, congruency and instructions in the RTs, F(3,90)= 4.91, p<0.01, but not in the errors; see Figure 2.4. In separate analyses of the two groups there was a reliable interaction in the CSR group, RT: F(3,45)= 5.11, p<0.05; errors: F(3,45)= 5.42, p<0.01, but not in the Tasks condition, F<1 in both the RTs and errors. In the CSR group this reflects the fact that performance on the congruent stimuli was essentially asymptotic after about three blocks pairs, while performance on the incongruent stimuli continued to improve.

The effect of having time to prepare also differed across blocks and between the instructional groups in the errors, as indicated by the three-way interaction between CSI, block pair and instruction, F(3,90)= 2.76, p<0.05. Separate analyses showed this was due to a marginally significant interaction in the Tasks group, F(3,45)= 2.59, p=0.064 but not significant in the CSR group. The near significant interaction in the Tasks group was caused by participants initially (in block pair 2) making more errors in a short-CSI (7.12%) than a long-CSI (5.65%) but then by block pair 4 they are making more errors in a long-CSI (6.11%) than a short-CSI (3.42%).

Other significant interactions. In the Monsell Experiment, the switch cost was shown to be larger on incongruent than on congruent trials in both the Tasks and CSR conditions. Although not significant, the numerical pattern in the RTs in the present experiment showed a slightly larger switch cost for incongruent (Tasks: 147ms, CSR: 33ms) than congruent trials (Tasks: 130ms, CSR: 23ms). However, the error data did not show the same trend. In this experiment there was evidence in the errors of a difference between the two instructional conditions, F(1,30)= 16.1, p<0.001. Separate analyses revealed a significant interaction between switch and congruency in the Tasks group, F(1,15)= 10.8, p<0.01, and in the CSR group, F(1,15)= 5.36, p<0.05. In the Tasks group there were larger switch costs for incongruent (3.92%) than for congruent trials (-0.01%), but in the CSR group there was a larger switch cost for congruent (1.03%) than for incongruent trials (-0.73%). The interaction between CSI and congruency (significant in the Monsell Experiment), was not significant in Experiment 1, F<2, for either RTs or errors.

2.3.2.2 The transfer test

As in the Monsell Experiment the effect of changing the stimuli was analyzed by comparing performance in block pair 5 with that in block pair 6, using a mixed ANOVA with the following factors: task switch (repeat or switch), congruency, CSI, block pair (5 or 6) and instruction condition. Only factors that interacted with block pair will be discussed.

As expected, the CSR group was more affected by the transfer to new stimuli; their mean RT (error rate) increased by 177ms (1.97%) as compared to the Tasks group's 48ms (-0.94%). However, the interaction was only marginally reliable for both RTs, F(1,30)= 3.15, p=0.086, and the error rate, F(1,30)=3.72, p=0.063.

Transfer to new stimuli increased the size of the congruency effect in the CSR group (26ms, 4.78%) but not in the Tasks group (4ms, -1.48%). This was supported by a significant interaction in the errors, F(1,30)= 6.57, $p<0.05$, but not in the RTs, F<1. Separate analyses of the errors revealed a significant interaction between block pair and congruency only in the CSR group, F(1,15) = 7.17, $p<0.05$.

Under both instructional conditions, transfer to new stimuli changed the nature of the relationship between CSI and switching. In the errors participants went from having a larger switch cost in a short CSI (Tasks: 2.2%, CSR: 1.89%) than a long CSI (Tasks: 0.87%, CSR: -0.71%) to having a larger switch cost in a long CSI (Tasks: 0.95% CSR: 2.41%) than a short CSI (Tasks: -0.64%, CSR: 0.63%). There was a significant block pair by CSI by task switch interaction in the errors F(1,30)=4.99, $p<0.05$. This was not the case in the RTs, F(1,30)= 1.83, $p=0.186$, although participants under both conditions did show a larger difference between the size of the switch cost in a long and short CSI before transfer (Tasks: 135ms, CSR: 42ms) than after transfer (Tasks: 78ms, CSR: 4ms).

A further analysis showed that performance improved from block pair 6 (Tasks: 687ms, 3.83%, CSR: 917ms, 5.54%) to 7 (Tasks: 626ms, 4.20% CSR: 765ms, 3.70%) significantly in the RTs, F(1,30)=18.4, $p<0.001$ but not in the errors, F(1,46)=1.16, n.s. The improvement was greater for the CSR group, especially for incongruent stimuli, but the corresponding interactions were not reliable; most likely because there was considerable variability over participants within groups in how quickly they coped with the transfer test.

2.3.3 Discussion

The results of Experiment 1 show that participants in a task-cueing paradigm under instructions to learn CSR rules (and this time with no task-set instructions) produce a pattern of results quite different from participants under normal task-switching instructions. As in the Monsell Experiment, the CSR group showed large congruency effects, small switch costs and no reliable reduction in switch cost with preparation, while the Tasks group showed small congruence effects and large switch costs which reliably reduced with preparation. The participants in the CSR condition again rapidly mastered the congruent relative to the incongruent stimuli, while in the Tasks group performance on congruent and incongruent stimuli improved at about the same rate. Transfer to a new set of stimuli was notably harder for the CSR than for the Tasks group. The consistent pattern across both experiments encourages one to believe that it reflects genuine processing differences contingent on the participant's approach to the task rather than mere artifacts of procedure.

The interaction between switch and congruency in the CSR errors seems inconsistent with that observed in the Monsell Experiment. However, the RTs show a non-significant effect in the opposite direction, so I will not dwell on this finding, but simply acknowledge the inconsistency and note that this effect is found repeatedly later experiments. The lack of a CSI by congruency interaction in this experiment, even in the Tasks group, is another point of disagreement between the first two experiments. However, Kiesel et al (2010) note that this interaction has an inconsistent history in the literature and so although there may be some difference between the Monsell Experiment & Experiment 2 which explains this effect I will again not dwell on this side issue.

The transfer test was somewhat unsatisfactory in this experiment, partly because the introduction of new stimuli in block pair 6 was accompanied by an instruction sheet showing the new CSR mappings for the CSR group, and the new cues for the Tasks group. Thus, although the CSR group showed a greater perturbation in performance, this is neither surprising nor informative, as the CSR group had much more need to refer to the instruction sheet, and this is likely to have disrupted performance. Moreover, although large, the difference between groups was unreliable for RTs and only marginally reliable for error rates, possibly reflecting large individual differences in the need to refer to the instruction sheet. This issue is addressed in Experiment 2 in the next chapter.

It interesting to consider the performance of the replaced participants. There is a concern about the purity of these participants' data, in that although those in the Tasks condition reported abandoning using the task-sets and those in the CSR group reported inducing the task-sets it is not clear from their verbal reports at what point in the experiment this happened. Those participants in the Tasks condition who said they stopped using the task-sets (n=8) showed much larger congruency effects (243 ms, 12.9%) than those who maintained a Tasks strategy. They also showed a slightly reduced switch cost (122ms, 3.4%), and a lack of a reduction in switch (short CSI: 107ms, 2.8%; long CSI: 138ms, 4.01%). Although it is difficult to make conclusions from this data it is suggestive that those participants who did not say they were using the task-sets to help their performance showed performance more similar to that shown by the CSR group.

The case is less clear for the 3 participants in the CSR condition who induced the task-sets. They showed a slightly smaller congruency effect in RTs but not errors (227ms, 15.7%) and a slightly higher switch cost again in RTs but not errors (50ms, 0.1%) than for the rest of the CSR group. They also showed no suggestion of a RISC effect (short CSI: 42ms, -1.1%; long CSI: 57ms, 1.2%). Hence it is unclear if those participants who induced the tasks in the CSR group then display a pattern of behaviour more similar to those seen in the Tasks group.

2.4 General Discussion

These two experiments (like the other experiments presented later in this thesis) employed the contingencies of a task-cueing experiment with relatively small sets of four stimuli and four task cues. They showed that quite different patterns of performance are obtained when participants follow different instructions. I have contrasted the performance observed under standard task-set instructions, whereby participants apply one of two classification rules as specified by the cue, to the performance observed when participants are induced to learn and use cue + stimulus → response (CSR) mappings. In the two experiments, the task-instructed group showed the usual pattern of a large task-switch cost, a reduction in that switch cost when time was allowed for preparation, and modest congruence effects that changed little with practice after the first block pair. When transferred to a new set of stimuli, they also showed only modest perturbations of performance, similar for all combinations of congruence and switch/repeat, consistent with a transient adjustment as the existing rules are applied to new stimuli.

The two experiments induced CSR learning in different ways. In the Monsell Experiment, participants were merely given a supplement to the standard task instructions that encouraged them to learn CSR mappings. In Experiment 1, the CSR group was given a list of the CSR mappings to learn with no mention of tasks. In each case the CSR group exhibited a large congruence effect, a small but reliable (or nearly reliable) "switch cost" (as defined for the Tasks group), which was not significantly or consistently reduced by preparation. The CSR group also showed more rapid improvement on the congruent stimuli (which require the learning only of simple S-R associations) than on the incongruent stimuli (which require learning of conditional associations), and suffered a considerable decline in performance when transferred to a new set of stimuli, especially for incongruent stimuli.

It is worth commenting on the different results gained in the transfer test here and in Dreisbach, Goshke & Haider (2006)’s Experiment 1. Here we showed that in the CSR participants who were unaware of the task showed worse performance on transfer than the Tasks condition, whereas Dreisbach, Goshke & Haider (2006) showed worse performance in the Tasks condition than in their SR condition. Why might this contradiction exist? The most obvious difference is the use of bivalent as opposed to univalent stimuli. It may be that the conflict created by having bivalent stimuli makes them more difficult to learn about. Hence the effect of transferring to new stimuli is still present after the 8 presentations of the stimulus in a given task used in the experiments above, whereas for the univalent stimuli after the 10 presentations that Dreisbach et al (2006) used per block, performance has become much less perturbed. In other words one or two presentations of a univalent stimuli may result in asymptotic performance to that stimulus and so Dreisbach et al (2006)’s effect may be hidden in the size of their block. Here the bivalent stimuli are much trickier to learn and so the ~8 trials in a block in this experiment still demonstrate the perturbation.

I conclude the chapter by discussing in turn (i) the implications for theories of task-switching of the difference between performance in the Tasks and CSR conditions, (ii) whether the CSR group are learning associatively.

2.4.1 Implications for the compound-retrieval account of task-cueing.

The first conclusion we draw is that this outcome is problematic for the general class of theory of which Logan et al's (2003)'s model and its later elaboration (Schneider & Logan, 2005) are the notable explicit examples. According to such a theory, in a task-cueing experiment with standard instructions, participants do not select different, or maintain the same, classification rules through exercise of an "endogenous" control operation triggered by interpretation of the cue's identity. They merely retrieve the response maximally activated by the combination of cue and stimulus. They do not switch between two classification task-sets; they apply a single CSR task-set throughout. In our experiments we encouraged just such a strategy by using a very small stimulus set and giving instructions and training designed to encourage a compound retrieval strategy. And, it appears, participants can apply a CSR strategy successfully given a reasonable amount of practice, though it may initially be costly in errors. But, when they apply this strategy we see a pattern of performance quite different from that seen when participants use standard task-set instructions. Hence, it seems most unlikely that participants who are given standard task-set instructions, and produce the patterns of data usually reported in the task-switching literature, are using a CSR strategy. The one element of support that we are able to offer for Logan et al's position is that we do observe a switch cost, albeit a small one, under CSR conditions. It may be the case that these cues are not perceived as two separate cues per task (as only one dimension of the cue is relevant) and as such the switch cost in the CSR conditions may in fact represent, the effect of cue priming. Experiment 2 seeks to address this issue by using 4 cues which vary along one dimension with two assigned to each task.

A hierarchical task-set strategy (select classification rule then select response) not only appears to be a natural way for participants to construe the demands of a task-cueing experiment, it also confers some notable advantages. It obviates the need to learn a large set of stimulus-specific rules, it allows easy generalization to novel stimuli, it provides a basis for effective preparation when time and foreknowledge permit (the appropriate classification rule can be readied and perhaps the other suppressed), and (as suggested by Dreisbach & Haider, 2008, 2009) it defends against interference from currently irrelevant S-R mappings — as indexed here by the modest congruence effect in the Tasks relative to the CSR condition. And the generalization to new stimuli accommodates the observation that when task-cueing experiments use larger stimulus sets (e.g. Monsell & Mizon, 2006), or very large stimulus sets, such that stimuli never (Arrington & Logan, 2004) or only rarely repeat (Elchlepp, Lavric & Monsell, in preparation; Monsell & Mizon, submitted) the phenomenology is qualitatively similar to that observed with small stimulus sets.

This argument against the compound-retrieval account as a general account of task-cueing performance is tempered by two observations. First, it clearly is possible for people to adopt and efficiently use a CSR — or compound retrieval — strategy, at least under conditions where the stimulus set is small enough to permit effective stimulus-specific learning, where sufficient practice is given, and where effective preparation is not a requirement. Second, experimenters using the task-cueing paradigm, especially with small stimulus sets (as is quite common), need to be alert to the possibility that some of their subjects may be learning and using CSR learning even when they are given standard task-set instructions; in this case performance (or brain activation) would not be fully representative of the selection and application of task-sets.

To cope with large stimulus sets with few stimulus repetitions, or trial-unique stimuli with no stimulus repetitions, it is necessary for the compound-retrieval model to allow participants to retrieve the response to the combination not of specific stimulus and cue, but the stimulus's category and the cue (Arrington & Logan, 2004). If it is assumed that both parity and magnitude categories of a stimulus digit are automatically retrieved, then it is possible to entertain a compound retrieval account in which categories retrieved (e.g. "odd" and "low" for 3) combine with the interpretation of the cue (e.g. blue means odd-even) to retrieve the appropriate response. The problem with this idea is the differential effects of preparation we observe in our CSR and Tasks conditions. The Schneider and Logan (2005) model attributes preparation effects, including the RISC effect, to facilitation of cue encoding by a related cue on the previous trial (even when the cue changes). If such priming is invoked to explain the effects of preparation (especially the reduction in switch cost) for the Tasks condition, it must surely predict the same effects of preparation in the CSR condition, in which there is every indication that participants are indeed using a compound retrieval strategy. But there is less benefit of preparation overall in the CSR group than in the Tasks group and no benefit in reducing the switch cost in the CSR group. Hence the large and consistent preparation effects in the Tasks conditions of our experiments cannot be explained by facilitation of cue encoding. An alternative explanation of the preparation effect in the Tasks condition that could perhaps be accommodated by a compound retrieval account (in which both categories of the stimulus are retrieved and compounded with the cue) would be to assume that the cue biases attention towards the cue-relevant dimension of the stimulus (and more so the more time is available for the biasing to take effect). But this is exactly the sort of top-down biasing operation that Logan and Gordon (2001) envisage as requiring "endogenous" control in their ECTVA model of executive control. Such an adaptation of the compound-retrieval model would thus be a variant of a generalized task-set reconfiguration account.

2.4.2 Are participants in the CSR condition learning associatively?

So far we have entertained the possibility that the CSR participants learned to respond through low-level associative mechanisms. There is an alternative possibility. Maybe the participants are learning and using CSR rules, but in no less a "cognitive" or "symbolic" way than the Tasks group. Mitchell, De Houwer and Lovibond (2009), for example, have argued that there is no need to postulate an underlying automatic associative system, and that all learning takes place via the formation of propositional beliefs. While leaving intact the implications of our findings for task-switching, such an account would be counter to the hope we expressed at the outset, that our experiments might contribute to extending the domain of application of the idea that a hybrid of symbolic (or propositional) processing with low-level associative processes underlies human cognition.

On what basis might we distinguish a propositional and an associative account of the CSR group's performance? The next chapter will examine how an associative learning network performs under these conditions. However, a previously mentioned argument appeals to Stoet and Snyder's (2003a, 2003b, 2007) work with Rhesus monkeys, who after massive practice showed only small switch costs but a larger congruency effect than their human controls (who received standard task-set instructions and relatively little practice.) Of course, if one is willing to grant Rhesus monkeys the same sorts of propositional representation as humans, this observation has no force. It may then be worth investigating how animals such as pigeons -- usually taken to be purely associative beings (Mackintosh, 1988) — would perform in these situations.

One argument for a propositional account is that in their responses to our post-experiment questionnaires in Experiment 1, participants in the CSR condition articulated propositional rules. They usually noted that two of the numbers (the congruent stimuli) were each linked to only one response, e.g. "8 was always left". They would also outline the mappings for the incongruent stimuli, most often by outlining one (or two) of them and then explaining how various changes meant the opposite response. Explicit use of such rules would evidently be easier for the congruent stimuli as their rules are simpler and easier to learn and apply. No benefit for preparation would be expected, as a relevant rule could not be engaged until the stimulus was presented. The relatively rapid recovery from transfer to new stimuli in the CSR group (albeit slower than in the Tasks group) could be accounted for by the participant having already ascertained the rule structure required to solve the problem. However, the propositional account finds it harder to explain the small but significant switch cost, unless perhaps we assume that there is a hidden admixture of participants (sometimes) applying task-set rules in this group.

Of course the ability to articulate a propositional rule does not mean that performance is being driven by it. It has been argued that the product of associative learning is accessible to conscious awareness (Evans, 2003). So it may also be the case that although initial learning and perhaps even performance were driven by associative mechanisms the final product may still be able to be articulated as a verbal rule. It has also long been thought that skill develops through proceduralization of what is initially represented declaratively (e.g. Anderson, 1982; Fitts, 1964). So, even if propositional rules can be articulated, performance in the CSR group may nevertheless be mediated, after practice, by an associative procedure. These issues will be examined in further detail in later chapters.

A related issue may be raised for the Tasks group. Although their performance suggests use of a task-set strategy, they are exposed to the same opportunities for low-level associative CSR learning as the CSR group. Does the use of task-level classification rules block the automatic acquisition or expression of lower-level SR and CSR associations? Or does true associative learning take longer to provide an efficient "procedural" basis for performance than the single session of practice we have provided here? It may be noted that switch costs are relatively robust over multiple sessions (e.g. the two sessions used in all Rogers & Monsell's, 1995 experiments). This question will be addressed in Chapter 4.

The two experiments described in this chapter show that inducing a CSR strategy in the task-cueing paradigm leads to a pattern of performance quite different from that generated by use of the strategy of selectively activating a classification rule (or task-set) on the basis of the contextual cue, and atypical of the pattern usually seen in task-cueing experiments. The next chapter seeks to elucidate whether performance mediated by CSR learning is broadly consistent with an associative account by providing a computational model of associatively-mediated task switching. But learning and retrieving responses to compounds of cue and stimulus appears, on the whole, neither the natural nor the effective strategy for the standard task-cueing situation, as it takes longer to learn, does not generalize easily to new stimuli, does not allow for effective preparation, and does not shield against irrelevant information.

Chapter 3 Can an associative model explain performance in the CSR condition?

3.1 Introduction

I argued in Chapter 2 that performance in the CSR conditions of the Monsell Experiment and Experiment 1 could be a result of the participants learning associatively. However, it was also acknowledged that their performance may simply reflect use of an alternative strategy based on stimulus-specific propositional rules. If an associative explanation is in any way valid, I need to establish that it is indeed possible for an association-formation mechanism to produce the same pattern of data as the CSR condition. The experiments indicated that participants in the CSR condition showed large congruency effects which changed over blocks and small switch costs. This chapter first seeks to establish if an associative model produces the same pattern of data. However, it is also the case that in the Monsell Experiment and Experiment 1 the learning of the CSR mappings might have involved some interaction between associative and cognitive processes: in the Monsell Experiment the CSR subjects were first given standard task-cueing instructions, and in Experiment 1 the CSR subjects were explicitly told the CSR rules. I therefore conducted a third experiment similar to the Monsell Experiment and Experiment 1 except that now CSR participants had simply to learn the responses to cue+stimulus combinations by trial and error – the most appropriate experience for an associative learning system. For reasons discussed below there were now two distinct cues per task.

3.1.1 Issues with the previous experiments

One issue with the experiments presented so far concerns the cues used. Recent debate in the task switching literature (e.g. Monsell & Mizon, 2006; Schneider & Logan, 2006; etc) has stressed the need to have two distinct cues per task so that the task-switch cost is not confound with the cue-switch cost. In the experiments presented in the previous chapter this was only arguably (rather than definitively) the case. Although, for example, blue and pink diamonds were used to cue one task, it was the diamond aspect of the cue which indicated the task (or the mappings for the incongruent stimuli). As cue encoding effects are largely thought to be the driver of a switch cost in the compound retrieval model (Schneider & Logan, 2006; Logan & Bundesen, 2003) I will instead now model a case where definitively distinct cues are used. In the experiment to be presented below these were 4 different shapes. This may not completely resolve the issue. Schneider & Logan (2006) argued that the two cues assigned to one task would associatively prime each other via a task representation; if no tasks are represented as such then no associative priming should happen. However, it was argued in Section 1.3.4 that in an associative system two cues could become equivalent for the incongruent stimuli without the need of a task representation (Honey & Ward-Robinson, 2002). So although using two distinct cues per task reduces the extent to which a task switch is confounded with a cue switch there may be questions remaining about how the cues are represented, to which we will return below.

The other issue revolves around the way in which knowledge is acquired. In both the CSR groups in the experiments presented in the previous chapter, participants either knew the task-sets (the Monsell Experiment) or were told explicitly the correct responses for a given cue and stimulus pair (Experiment 1). It is difficult to know how reading and acquiring a declarative representation of the correct mappings would impact on an associative system that primarily learns through error correction. Would it simulate the trials? Would it now be faster to learn how to respond, as it “knows” the correct response? Or would performance be driven by explicit knowledge until the associative system’s performance is better? As there are no easy answers to these questions I chose to compare an associative formation model to a situation where participants were not informed of the correct responses at the start but had to learn them by trial and error: hence Experiment 2.

3.1.2 What model to use?

In Chapter 1 it was noted that there were reasons to believe that the best type of model for this situation is likely to be a connectionist model that has been modified so that it uses the best unit to carry each mapping, i.e. APECS (McLaren, 1997). What follows provides a brief recapitulation of the argument.

The mappings for the congruent stimuli were as outlined in Table 3.1. It is immediately evident that this structure should easily be learnt by any associative model as the digit stimuli predict the correct response irrespective of the cues.

		Task 1 Cues		Task 2 Cues	
		W (circle)	X (triangle)	Y (square)	Z (pentagon)
	A (1)	**L**	**L**	**L**	**L**
Stimuli	E (3)	L	L	L	L
	F (6)	R	R	R	R
	B (8)	**R**	**R**	**R**	**R**

Table 3.1 The associative structure of the congruent trials. L indicates a left R a right response. Boldface rows indicate example initially trained stimuli; the others were the transfer stimuli. The shapes and numbers in brackets give an example of the possible stimuli and cues that could be used in one counterbalancing of the experiment.

However, the incongruent trials, shown in Table 3.2, are more of a challenge for an associative model to learn. The problem here is that a single layer error-correcting model, e.g. Rescorla-Wagner (1972) would be unable to learn this structure as no one element completely predicts the correct response (unless it were adapted to have a configural input coding). A single layer configural model, e.g. Pearce's model (1987, 1994) would be able to learn the correct responses and due to generalisation between compounds which share an element it would find the incongruent trials harder than the congruent trials. For example, if I just consider the initially trained stimuli, WA-L would generalise to XA-L, YA-L, ZA-L, WB-R, WC-L & WD-R of which 4 give the desired L response and only 2 give the opposite response, whereas WC-L would generalise to XC-L, YC-R, ZC-R, WA-L, WB-R & WD-R of which only 2 give the desired L response and 4 now give the opposite R response. This would make the incongruent stimuli harder to learn.

		Task 1 Cues		Task 2 Cues	
		W (circle)	X (triangle)	Y (square)	Z (pentagon)
	C (4)	**L**	**L**	**R**	**R**
Stimuli	G (2)	L	L	R	R
	H (9)	R	R	L	L
	D (7)	**R**	**R**	**L**	**L**

Table 3.2 shows the associative structure of the incongruent trials in the same way as for the congruent trials in Table 3.1.

However, a model with purely configural encoding would not be able to explain the phenomenon of cue equivalence (Honey & Ward-Robinson, 2002; Hodder, George, Kilcross & Honey, 2003), which, as we are now considering a case with two distinct cues per task, becomes an issue. As previously mentioned, these authors trained rats or humans (respectively) with the same contingencies as the incongruent trials. They found that cues that indicated the same outcome from stimuli became equivalent, e.g. here W and X would become equivalent, as would Y and Z, in that there would be a greater degree of generalization between W and X than W and Y. Honey & Ward-Robinson (2002) found that a modified connectionist model was able to account for their data, which showed that performance to WX was much noisier (had a larger standard deviation) than performance to WY. They predicted this because W & X use the same hidden units to carry their mappings to both the outcome and no outcome and so when they are both presented both their hidden units are activated which leads to a large input to both an outcome and no outcome unit and hence noisy performance, as both the outcome and no outcome units will be highly active. In contrast W & Y do not share any hidden units and so presenting both of them leads to only a small amount of activation per hidden unit and so less input into the outcome and no outcome units, as so less noisy performance as neither the outcome or no-outcome units are fully active (see Section 1.3.4 for more details).

Hence, I have chosen a model from the same class as their chosen model, after my initial attempts at using a more traditional feedforward backpropagation (Section 1.4.1) network, with input, hidden and output layers were less than optimal. It was found that although a traditional feedforward backpropagation model could eventually learn the biconditional discrimination, it could not learn it within the time frame of the experiment. This is a well recognised failing of connectionist networks.

The chosen model is known as APECS (McLaren, 1994, 2011; Le Pelley & McLaren, 2001) and, as described in Section 1.4.2, has a good record in modelling human learning and memory. APECS has the basic characteristics of a backpropagation network (Rumelhart, Hinton and Williams, 1986), i.e. it is a feedforward error correcting system with input, hidden and output layers. This standard feedforward, back propagation network was adapted in two key ways in order to make it an APECS network, see Figure 3.1 for the architecture of the model. The key modifications were:

Learning algorithm and rates: The APECS learning algorithm allows the learning rates to change in an adaptive manner. On a given trial, the hidden unit with the largest positive error (with a large contribution for one output unit) is given a higher learning rate than the other hidden units. This effectively means that one (or a few) hidden unit(s) is (are) selected to carry each mapping from input to output.

Bias: The APECS group of models also includes an adaptive bias whose learning rate is varied to prevent catastrophic interference to old learning occurring when new information is learnt (McCloskey and Cohen, 1989). The adaptive bias prevents the same hidden unit being used by a different mapping and hence prevents the previous learning being over-written.

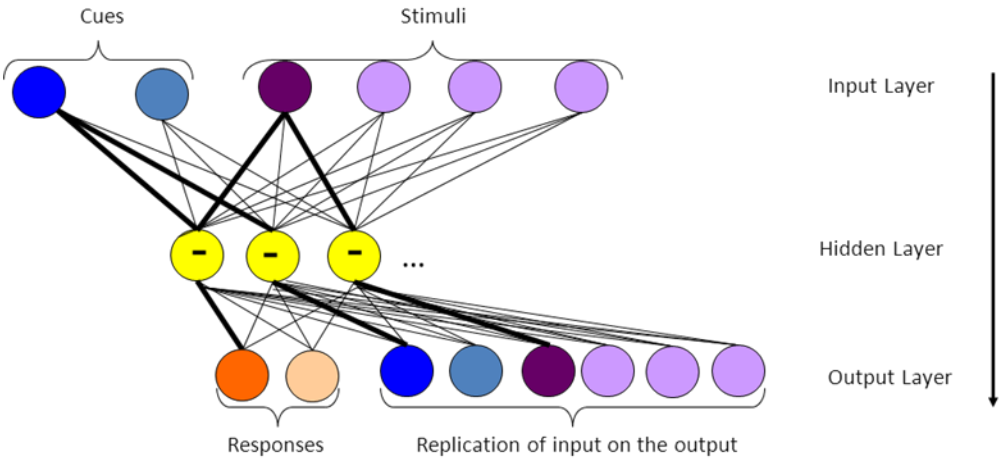


Figure 3.1 A diagrammatic representation of APECS. Note that darker units represent units activated on the trial shown. Darker units suggest a way in which these units could have found hidden units to carry their mappings, both to the correct output unit and to the unit replicating the cue on the output layer. Note that more input, hidden and output units were used in the actual simulations.

In addition to these two algorithmic changes there was also a structural change; the model was also trained to auto-associate the input layer with the output layer. That is, each unit on the output layer corresponded to a unit on the input layer and was trained to be active when that unit was present. The model also cycled 200 times on each trial. It then had a 200 cycle ITI where nothing was presented to the model and the target output was the resting level of 0.5. This was done as a simple way of acknowledging that a trial takes time to complete and that this time may not correspond to just one cycle of the network. However, as it was not supposed here that a cycle corresponded to an interval of time (see Chapter 6 for a further discussion of real time models) the temporal aspects of the experiment were not encoded in this initial simulation.

3.2 What does this model do in a task-cueing paradigm?

As previously noted it would be unfair to compare what this model can achieve with the experiments presented in the previous chapter, as it is unclear how this model should incorporate being told the mappings initially. This thesis takes the broad stance that there are cognitive and associative processes, and that although these may interact any discussion of how they interact is a matter for further research. So although one would expect the performance of the model to be broadly consistent with the performance in the Monsell Experiment and Experiment 1, there may well be other processes in these two experiments (and indeed later experiments) that cannot be captured by the model.

The model also cannot capture the effect of CSI, as it is a trial-based model with no means of encoding time. This means that any effect of the temporal disparity in the appearance of the cue and the stimulus will not be captured in this model. An attempt to address this issue, and other issues with this simplistic modelling approach will be considered in Chapter 6. The results of the present model will now be reported and then compared with the results of Experiment 2 where four distinct cues were used and participants in the CSR condition learnt by trial and error.

3.2.1 Model Method

The general approach taken to modelling in this thesis is to consider multiple runs through the model. For each run, a different set of starting weights was used, as was a different sequence of trials. The trial sequences were generated in an identical manner to those used for behavioural participants. The initial random variation in the weights is necessary for the model to be able to reach a solution at all. Given the variation in the initial starting weights, this causes variation in the precise nature of the solution to the problem. As there are subtle differences in the solutions (as one suspects is the case with human participants) the same statistical approach is necessary in order to ascertain which differences in the output of the models are genuine. In this instance the model was run 32 times. Running the model just 32 times was done for purposes of efficiency and in order to given an indication of the level of power required to illuminate the effects. Where a sample of 32 models did not lead to a clear-cut prediction about an effect that was deemed important because it appeared in the behavioural data or was theoretically meaningful, the sample size was increased to establish whether that effect would become reliable or not.

3.2.1.1 Sequencing

The trial sequencing constraints were the same as those for all the experiments mentioned in Chapter 1, as the same constraints were also used for Experiment 2. This means that the lengths of runs of trials on the same task were controlled within a block to approximate the expected distribution of run lengths for a binomial sequence with p(switch)=0.5. And each digit appeared equally often on switch, first repeat and second repeat trials in a run, and then randomly on other trials with the constraint that each digit appeared roughly equally often during the experiment.

3.2.1.2 Representation

The stimuli and cues were represented discretely with one input unit coding for each[5]. If a stimulus or cue was present the corresponding unit was given an activation of 1; if absent, 0. The responses were also represented discretely and the model was trained to a target activation of 0.9 for the correct response and 0.5 for the wrong one. As noted above, the model also learnt to replicate the inputs on the output layer, so that each unit on the output layer corresponded to a unit on the input layer and was trained to an activation target of 0.9 if the input unit was on and 0.5 if the input unit was off[6].

3.2.1.3 Structure

The network has three layers: 16 input units, 14 hidden units and 18 output units.

3.2.1.4 Learning parameters

The fast learning rate was set to 0.8 whilst the slow learning rate for the unselected units was 0.0005. For the bias the learning rate for selected hidden units was 0.5 and for others was 0.005. These parameters were the same as used in McLaren (2011) and McLaren, Forrest & McLaren (2012).

3.2.1.5 Output

The output of the model was examined by comparing the activation difference between the two output units which corresponded to the two responses and the target difference (target-trained). On this measure larger scores mean worse performance, i.e. less close to the desired discrimination.

---

[5] Other simulations show that coding the cues in a manner that would be more appropriate for simulating the experiments presented in the previous chapter (by using one unit to code a cue's colour and one to code its shape thereby encoding the relatedness of the cues) makes no substantial difference to the performance of the model, although the switch cost became only marginally reliable (as was in fact the case for the behavioural data); this could be a result of the slight difference in the associative structures formed.

[6] This replications means that increasing the amount of input, increases the amount the model has to learn. This could be viewed as an implicit parameter within the model.

3.2.3 Modelling Results

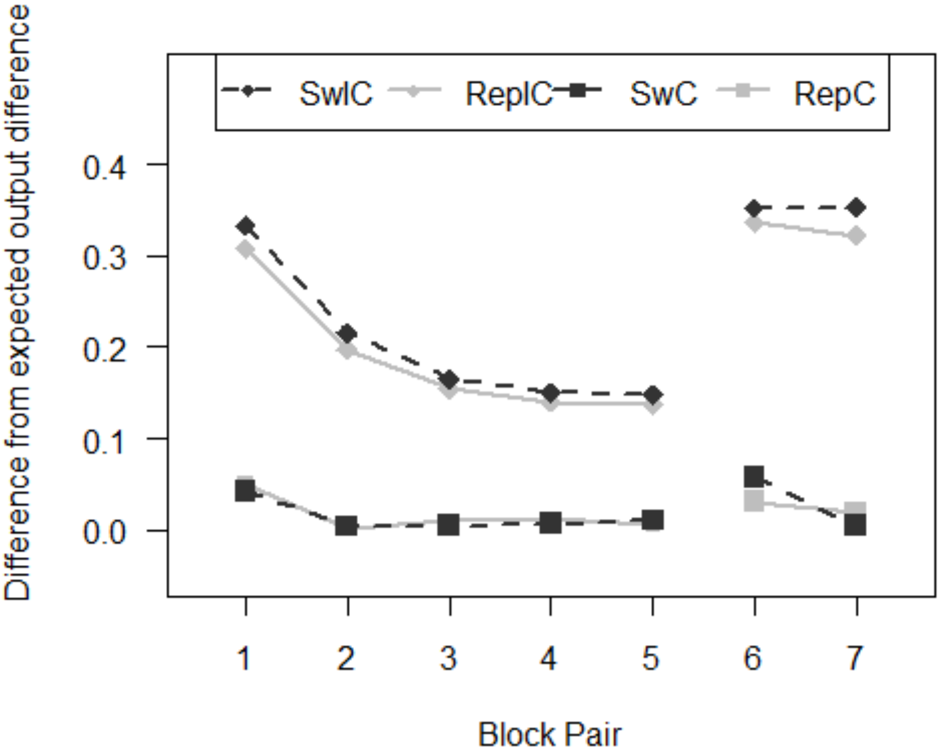


Figure 3.2 shows the performance of the model of Experiment 2 by block pair broken down by switch trials and congruency. Diamonds represent incongruent stimuli, squares represents congruent stimuli. Dashed dark lines represent switch trials and solid light lines represent repeat trials.

3.2.3.1 Block Pairs 2-4

From inspection of Figure 3.2 it can be seen that the pattern of results is indeed very similar to that seen in the Monsell Experiment and Experiment 1. The model shows a large congruency effect that decreases over time; a small switch cost that is largely present in the incongruent trials and a large effect of changing the stimuli in the transfer test. The results were analyzed using a repeated measure ANOVA, with the factors block pair (2,3,4,or 5), task switch, and congruency. Unlike the behavioural data there is only one measure here that is a continuous measures making it unclear what represents an error and so trials after an error (or indeed when an error is made) cannot be defined and so cannot be excluded from the analysis.

Task switches There was small but significant effect of "task switch" whereby performance was worse on switch trials (0.063) than repeat trials (0.056), F(1,31)=4.46, $p<0.05$.

Congruency There was a large and significant effect of congruency in the model, with performance worse on incongruent (0.116) than congruent trials (0.003), F(1,31)=153, $p<0.001$.

Switch by congruency. The switch cost was significantly larger for the incongruent (0.0130) than the congruent trials (0.000599), F(1,31)= 4.16, p=0.05. The switch cost was only reliable for the incongruent stimuli, F(1,31)= 4.99, $p<0.05$, and not for the congruent stimuli, $F<1$.

Acquisition effects. Overall performance reliably improved from block pair 2 to 5 (Figure 3.2), F(3, 93)= 18.2 $p<0.001$. The two-way interaction between block pair and congruency was significant F(3, 93)= 19.3, $p<0.001$ This interaction can be seen in Figure 3.2 which shows the congruent stimuli being learnt quickly whilst the incongruent stimuli take longer to learn.

3.2.3.2 Transfer to new stimuli

The transfer test was analysed by comparing performance in block pairs 5 and 6 using an ANOVA with the factors block pair (5 or 6), task switch and congruency.

There was a significant main effect of changing the stimuli, F(1,31)=185, $p<0.001$, which reflected worse performance in block pair 6 (0.157) than 5 (0.057). There was also a dramatic increase in the size of the congruency effect from block pair 5 (0.11) to 6 (0.242) that was supported by a highly significant interaction, F(1,31)=65.6, $p<0.001$. For completeness it is worth noting that as in the behavioural data performance in the model improved from block pair 6 (0.157) to block pair 7 (0.105), F(1,31)= 141, $p<0.001$.

3.2.4 Discussion

The model performed in a similar way to the participants in the Monsell Experiment and Experiment 1. It showed a small but reliable switch cost that was most in evidence on incongruent trials, and a large congruency effect that varied over block pairs. However, for reasons mentioned earlier the comparison to the Monsell Experiment and Experiment 1 is somewhat suspect, and Experiment 2 was run to provide a more appropriate comparison. I will resume discussion of the modelling after the results of Experiment 2 have been reported.

3.3 Experiment 2

This experiment was similar to Experiment 1, except that the CSR group now learned the mappings entirely by trial and error and two distinct cues were used per task.

3.3.1 Method

3.3.1.1 Participants

The participants were 46 psychology undergraduates (average age 19.0 years, 6 male) at the University of Exeter. Participants took part for course credit and a bonus payment, which was contingent on their performance (average payment £2.03, range £1.50-£2.50).

3.3.1.2 Stimuli

The task cues were a regular circle, triangle, square and pentagon with a standardized area of 40,000 pixels, filled with blue (RGB: 115, 194, 251); the digit stimulus was displayed in the center of the cue in size 60 Courier bold font (1.3° of visual angle). The two sets of digits used were 1,4,7,8 and 2,3,6,9. The computer and screen were as in Experiment 1.

3.3.1.3 Design and procedure

The design and randomization were identical to that of Experiment 1 except that there were now four distinct cues, with two of the cues signalling one task and two signalling the other and the cue changing on every trial. The same cue sequencing as in the Monsell Experiment and Experiment 1 was used, but this time the cues changed shape on every trial, and only some changes signaled a change of task.

The procedure was also identical to Experiment 1 except that no instruction sheet was provided; participants in the Tasks condition were given standard task-set instruction verbally and at the beginning of experiment on the screen, whereas participants in the CSR condition were directed to learn cue-stimulus → response connections on the basis of trial-by-trial feedback. Both groups were told the first two blocks were for practice. Hence in the CSR condition the participants had to learn how to respond by trial and error, whereas in the Tasks condition they could apply the specified task rules. As with the previous two experiments, after 5 block pairs the set of stimuli used was swapped to the other possible set of stimuli, and again participants were told they could treat the first two block pairs after the change as practice.

Participants were debriefed as in Experiment 1, and replaced if their reported strategy differed from that instructed. Six participants in the Tasks group who did not mention using the tasks in the questionnaire, and 7 participants in the CSR group who induced one (5 participants) or both (2 participants) of the tasks, were replaced, until there were 16 participants per group whose reports matched the instructions. The replaced participants will be discussed after the main experiment (Section 3.3.3.2).

3.3.2 Experiment 2 Results

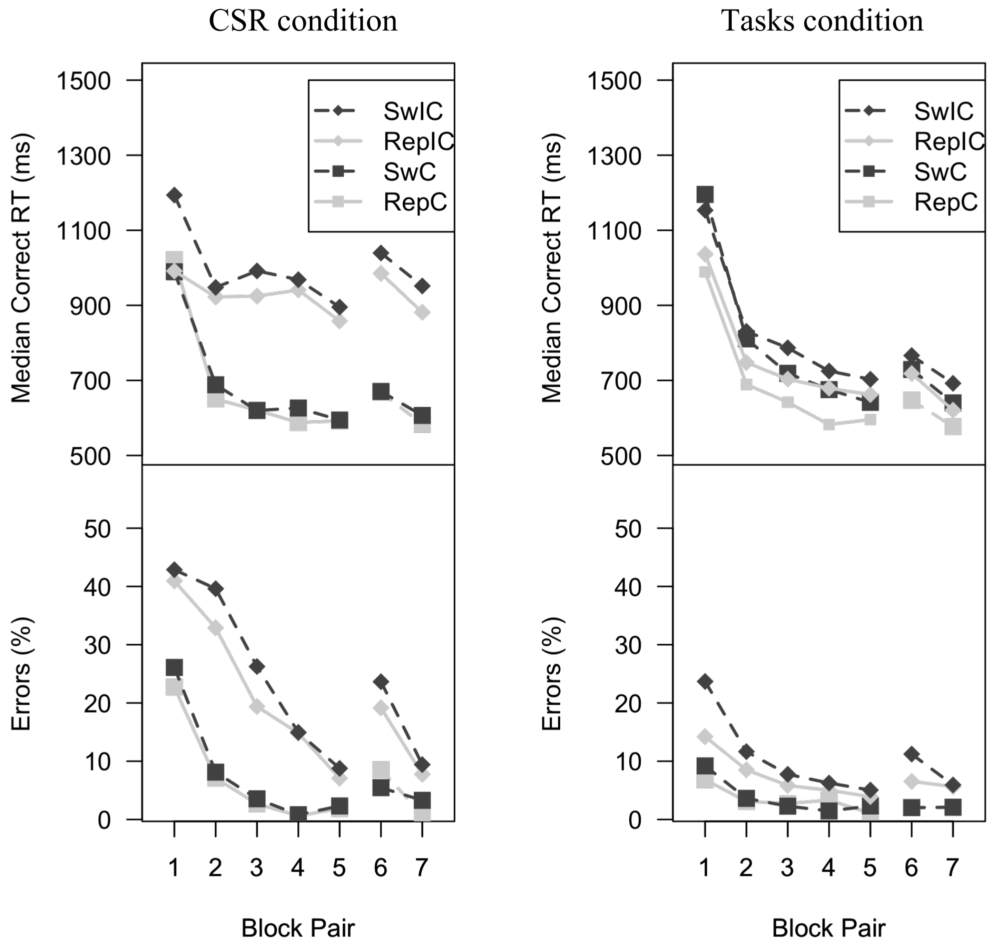


Figure 3.3 shows the means of median correct RTs and error % over successive block pairs in Experiment 3, for each combination of congruent/incongruent stimuli and switch/repeat trials. Darker dotted lines represent switch trials, lighter solid lines are repeat trials. Square symbols represent congruent trials, diamonds represent incongruent trials. The breaks in the lines mark transfer to a new stimulus set for block pairs 6 and 7. The left hand panel shows results for the CSR condition; the right hand panel shows the results for those participants in the Tasks condition.

3.3.2.1 Block Pairs 2-5

Performance is shown as a function of block pair in Figure 3.3. As the CSR group now had to learn 16 distinct CSR mappings from error-correction alone, it is not surprising that they did not reach a level of performance equivalent to the Tasks group by block pair 5. On average their median RT in block pair 5 was longer (735ms) than the Tasks group (651ms) and they made more errors (5.02%) than the Tasks group (3.09%); the difference stems largely from performance on the incongruent stimuli.

As for Experiment 1, a mixed ANOVA with the factors: task switch (repeat, switch), congruency (congruent, incongruent), CSI (long or short), block pair (2,3,4 or 5) and instruction condition (Tasks, CSR) was run on median correct reaction time (RT) and percent errors, and followed up by separate ANOVAs for the two groups where appropriate. Overall the CSR group were slower (777ms) and made more errors (11.9%) than the Tasks group (700ms, 4.62%). This was a significant difference for errors, F(1,30)= 25.6, p<0.001, but not RTs, F(1,30)= 2.47, p=0.13.

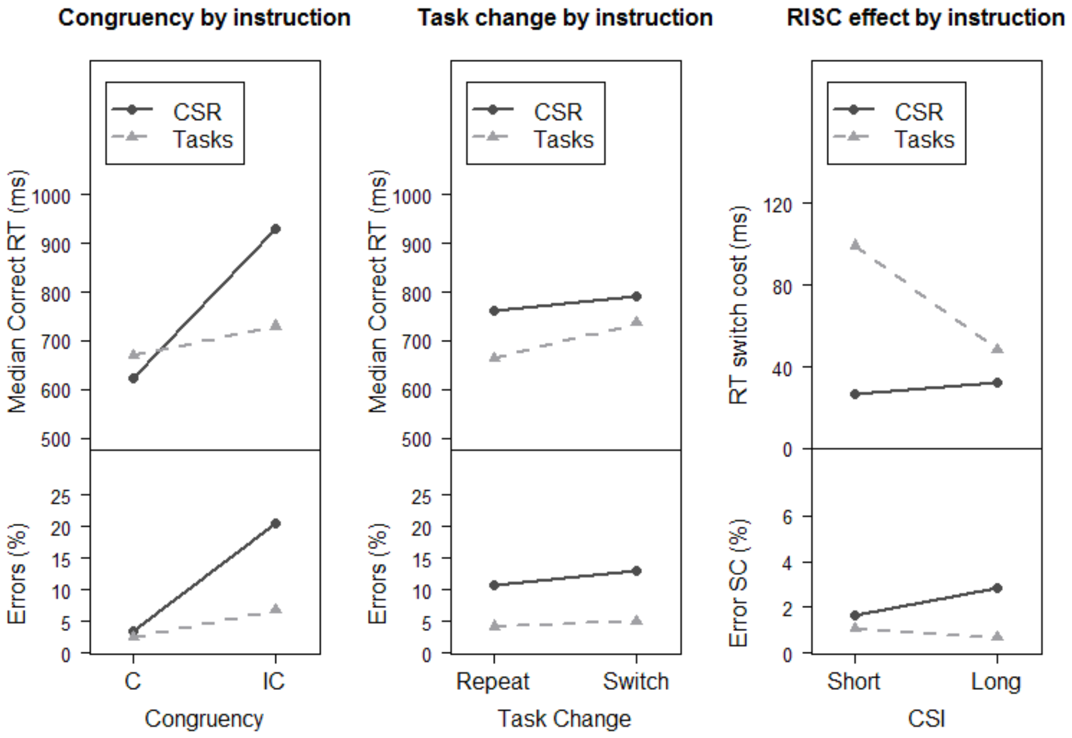


Figure 3.4 shows contrasts indicating three principal findings of Experiment 2. Reaction time contrasts are shown in the upper panels, error rate contrasts in the lower panels. The three sets of panels show interaction of instruction group with the effects of stimulus congruence (left), a task switch (middle), and the reduction in switch cost with increasing CSI (right).

Task switches and instruction. As with the Monsell Experiment and Experiment 1 there was a much larger switch cost in the Tasks group (73ms) than in the CSR group (30ms), F(1,30)=3.83, $p<0.001$ — see Figure 3.4. The costs were reliable for both the Tasks condition, F(1,15) = 15.1, $p<0.001$, and the CSR condition, F(1,15) = 5.58, $p<0.05$. For errors, the overall main effect of a task switch was reliable, F(1,30)=6.58, $p<0.05$ (repeat: M=7.98%, switch: M= 9.05%), but there was no reliable interaction with instruction, F(1,30)=1.31, p=0.26: the switch cost was 0.86% for the Tasks group and 2.24% for the CSR group

Preparation and instruction. As Figure 3.4 shows, preparation reduced the RT switch cost in the Tasks group from 99ms in the short-CSI blocks to 49ms in the long CSI blocks, F(1,15)=5.84, $p<0.05$. There was no such effect in the CSR group, for whom the switch cost was 27ms in the short-CSI blocks and 32ms in the long-CSI blocks F<1. The three-way interaction was marginally reliable F(1,30)= 3.27, p=0.084. Participants in the Tasks condition also showed a general preparation effect, whereby if only the task-repeat trials are considered they were faster with a long-CSI (527ms) than with a short-CSI (849ms), F(1,15) = 149, $p<0.001$. For the same contrast the CSR group was also faster with the long- CSI (745ms) than at the short-CSI (845ms), F(1,15) = 23.7, $p<0.001$, but the effect was considerably smaller. These interactions were not significant in the errors F<1.

Congruency and instruction. RT and error rates showed (Figure 3.4) a much larger effect of congruency in the CSR group (308 ms, 17.1%) than in the Tasks group (60 ms, 4.26%); the interaction was highly reliable both for RTs, F(1,30)=27.3, $p<0.001$, and in the error rates, F(1,30)=18.2, $p<0.001$. In separate analyses, the congruence effect was reliable for both the Tasks group, F(1, 15) =12.5, $p<0.001$, for RTs, and F(1,15)=13.3, $p<0.01$, for errors, and the CSR group, F(1,15) = 48.4, $p<0.001$, for RT, and F(1,23)=37.9, $p<0.001$, for errors

Acquisition effects. Overall performance improved from block pair 2 to 5 (Figure 3.3), and this was reliable for RTs, F(3,90)= 8.99, $p<0.001$. Participants in the CSR group this time were learning entirely by trial and error: hence the much higher initial error rate than in Experiment 1, and consequent steeper learning curve than the Tasks condition, as indicated by a significant interaction between block pair and instructions in the errors, F(1,30) = 18.3, $p<0.001$.

Again the Tasks group showed little change in congruence and switch effects over blocks, while the CSR group showed markedly different learning rates for congruent and incongruent stimuli, with near-asymptotic performance reached for congruent stimuli by about block pair 4, while the incongruent stimuli showed a more gradual but nevertheless substantial improvement over the whole range. This time, presumably because trial and error learning is harder, the improvement was manifest largely in errors, so that the three-way interaction between block pair, congruency and instructions was significant in the errors only, F(3,90)= 10.0,p<0.001. Separate analyses revealed a highly significant block pair by congruency interaction in the CSR condition, F(3,45)= 18.5, p<0.001, but not in the Tasks group, F(3,45)=2.34, p=0.09. Figure 3.3 shows that in the CSR condition performance on the incongruent trials was nowhere near asymptotic.

Other significant interactions. In agreement with the Monsell Experiment the switch cost was again larger for incongruent trials for the CSR group (39ms, 3.85%) than congruent trials (19ms, 0.63%), this was nearly significant in the errors, F(1,15)=3.80, p=0.07, but not in the RTs, F<1. For the Tasks group the error switch cost was larger for incongruent trials (1.82%) than for congruent:(-0.14%) and nearly reliable, F(3.47), p=0.082, but the RT cost was smaller for incongruent (63ms) than for the congruent trials (85ms), F(1,15)= 6.51, p<0.05. There was an overall significant interaction between task switch and congruency in the errors, F(1,30)= 7.02, p<0.05, but not in the RTs. This effect did not differ between the two experimental conditions in the error data or RTs.

The congruency effect was larger in the short CSI (Tasks: 82ms, CSR: 342ms) than the long CSI (Tasks: 39ms, CSR: 276ms), F(1,30)= 9.79, p<0.01, though the proportional reduction was much larger for the Tasks condition. The three-way interaction with instruction was not significant for either RT or errors, F<1.

3.3.2.2 The transfer test

Figure 3.3 also illustrates that performance was perturbed by the introduction of novel stimuli at block pair 6 (the transfer test). The perturbation again appears more marked and persistent for the CSR group, especially for incongruent stimuli. The effect of transfer was analyzed as for the Monsell Experiment and Experiment 1.

As expected, the CSR group was more affected by the transfer to new stimuli; their RT (error rate) increased by 106 ms (9.18%) as compared to the Tasks group's increase of 64 ms (2.18%). The interaction was significant for the error rate, F(1,30)= 10.5, p<0.01, but not the RTs, F(1,30)<1.

In the CSR group, transfer to new stimuli increased the congruency effect, by 58ms (8.59%), reflecting particular difficulty in learning a new set of incongruent stimuli, whereas in the Tasks group it decreased in the RTs (-10ms) but increased in the errors (4.43%). In the error rate there was a significant interaction between block pair and congruency, $F(1,30)=11.5$, $p<0.01$, but not in the RTs, $F<1$. Performance improved from block pair 6 (Tasks: 715ms, 5.28%, CSR: 841ms, 14.2%) to 7 (Tasks: 632ms, 4.07% CSR: 756ms, 5.33%) significantly both in the RTs, $F(1,30)=24.5$, $p<0.001$ and in the errors, $F(1,46)=33.4$, $p<0.001$. The CSR group reduced their error rate significantly more than participants in the Tasks group, $F(1,30)=19.4$, $p<0.001$— but had much more room for improvement.

3.3.3 Discussion

The CSR group, who learnt by trial and error, showed a similar pattern of results to the CSR groups in the previous experiments and to the model outlined earlier, i.e. large congruency effects and small switch costs. Preparation reliably reduced the switch effect in the Tasks group but there was no reduction in the CSR group, though this interaction was only marginally reliable. (The model was in any case not designed to address CSI effects). In the CSR group, the learning curves for congruent stimuli rapidly reached asymptote while the learning curve for the incongruent stimuli did not, suggesting that learning was easier for the congruent stimuli.

The consistency of these differences across experiments suggest that they are not greatly dependent on the procedure used to induce CSR learning, or other details such as the type of cues utilized in the experiment. This lack of a real difference supports the notion that CSR learning is dependent on an automatic process. If a controlled process drove the effects seen then we might expect differences based on verbal instructions given or the cues used. It is also worth noting that the way in which the input is represented in APECS it is ambivalent to the content of the cues and the stimuli, e.g. whether the cue is a colour or a shape.

However, it is clear that in this experiment, compared to the first two, the CSR group were at a much greater disadvantage relative to the Tasks group, both when they started the experiment, and at transfer. They had no way of knowing which response to make to a given cue-stimulus combination until they had encountered that combination and been given feedback at least once, and then they had to remember the mapping. This doubtless explains the CSR group’s higher error rates, both in blocks 2 to 5 and during the transfer block. It also highlights one of the key advantages of using a task-set strategy in this paradigm: the ability to deal with novel stimuli because task-rules are categorical not stimulus-specific. In all three experiments there is clear evidence that task-sets allow this transfer more easily than previously learnt CSR mappings, though I note that there do appear to be some savings at transfer compared to initial performance even in the CSR condition.

3.3.3.1 How did the model "solve" the problem?

The qualitative agreement between the pattern of responding in Experiment 2 (and indeed to some extent in the Monsell Experiment and Experiment 1) and the predictions of the model are striking. The model, inasmuch as it can, shows the same effects that are present in the CSR group. The interesting question to address now is 'how did the model solve the problem?' One of the additional advantages of using an APECS model over a standard connectionist network is that it tends to form less distributed representations at the hidden layer. This means its solutions to the problem of mapping input to output can be more easily interpreted.

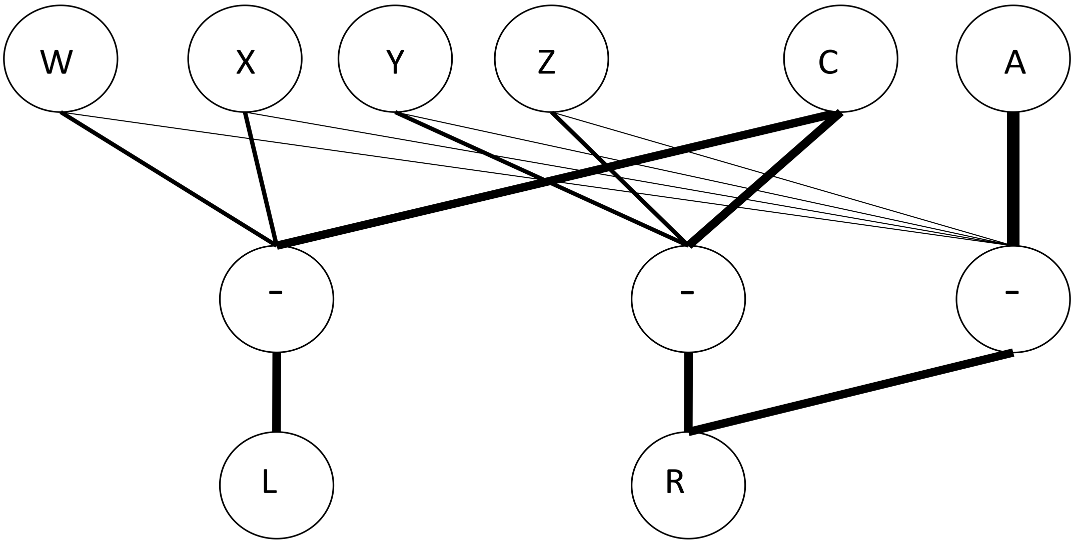


Figure 3.5 shows an idealised version of the typical structures formed between the input and hidden layer for a congruent stimulus A and an incongruent stimulus C. W & X are cues for one task and Y & Z are cues for the other. The thickness of the line indicates the relative strength of the links. B & D would use different units but in a similar vein to A & C respectively.

When examining the types of solutions created by the network cue equivalence can often be seen (see Figure 3.5 &3.6). Let us consider the more complex case of the incongruent cue and stimulus combinations first. For these, cues indicating the same stimulus->response relationship for the incongruent trials usually map onto the same hidden unit, e.g., the mapping of WC-L is carried by the same hidden unit as the mapping XC-L. For example in Figure 3.6, hidden unit 1 carries this type of mapping, as it strongly linked to cues 2 and 4 as well as incongruent stimulus 7. This is essentially the associative structure suggested by Honey & Ward-Robinson (2002) when they proposed how their hypothetical network would solve a biconditional discrimination with equivalent contexts. However, the solution is not always quite as clear cut as this – as we will see below.

In the mappings for the incongruent CS combinations both the cue and the stimulus play a role in activating the hidden unit in question, although usually slightly more influence is exerted by the stimulus than the cue. This is due to the cue equivalence developed during training. As the same hidden unit carries the mapping for a given stimulus for both equivalent cues, when the same stimulus is presented with the one cue this strengthens the mapping for the stimulus, whilst only strengthening the mapping from the currently active cue. Over a number of trials this leads to the stimulus having more influence on responding as its link has been strengthened more times than that of either cue.

A similar argument was put forward by Honey et al (2010) in their explanation of intradimensional/ extradimensional (ID/ED) set shift. In a typical ID/ED experiment participants are presented with AX+, BX+, AY-, BY-, in which A and B come from one dimension whilst X and Y come from a different dimension. They are subsequently presented with either CV+, DV+, CW-, DW- (ID) or CV+, DV-, CW+, DW- (ED), where C & D come from the same dimension as A & B, and V & W come from the same dimension as X & Y. Typically performance on ID set shifts is faster than ED set shifts (e.g. Mackintosh & Little, 1969; Roberts, Robbins & Everitt, 1988). Honey et al argued this was because during training X and Y developed stronger links to their respective hidden units as X and Y had more individual pairings where their mapping has been carried through the same hidden unit. This gives this dimension more influence over the hidden layer and then because they share more of their representation with V and W than C and D the former are better able to activate the hidden layer. In an ID set shift V and W then have an advantage because the links between V and W and the existing hidden units is more profitable as they can quickly learn the outcome, whereas in ED set shifts as the discrimination is based on the other component the links are disadvantageous.

Based on a similar argument, the congruent stimuli tend to have a mapping which is mostly based on the stimulus, so all of WA-L, XA-L, YA-L and ZA-L are carried by the same hidden unit but the activation of this hidden unit is much more dependent on the presence of the stimulus than of the cue (see Figure 3.5). This again is because all the cues are equivalent for the congruent stimulus and so, as the stimulus can occur with any one of four cues and use the same hidden unit, the stimulus' link gets strengthened more frequently. For example in Figure 3.6 hidden unit 14 carries the mapping for congruent stimulus 1.

This difference between the structure created by the congruent trials and the incongruent trials easily explains why the congruency effect emerges from the model. The structure for the congruent trials is much easier to learn and, given than most of the time the model "solves" the experiment by having just one hidden unit which carries all the mappings for a given congruent stimulus, the link between that hidden unit and the output unit receives many more trials on which it is strengthened than does the link from the incongruent stimulus and its (usually) two hidden units. In general the link from the congruent stimulus to its hidden unit and that hidden unit's link to the output unit should be stronger than for the incongruent stimuli as they have had more opportunities to be trained up. The congruent stimuli also experience less interference from other trials in establishing this stable position, as it is quicker for the stimulus to come to dominate performance and it is less subject to possible accidental unlearning due to trying to pair a chosen hidden unit with the opposing response. When learning the incongruent stimuli it is possible for its representation in one task to interfere with the other task and lead to unlearning – however, for the congruent stimuli there is only cue based interference and not stimulus based interference. So, like the experimental participants, the model finds the congruent stimuli a lot easier to learn and maintain than the incongruent stimuli.

Although these types of mapping, which are in agreement with the mappings outlined in Honey et al (2010), are one way in which the network solves the problem there are other slight variants. For example, particular incongruent or congruent trial types are sometimes carried by their own unique hidden units. This is simply due to the random nature of the starting weights which may have caused a given hidden unit to be particularly favourable for one particular pairing but not for any of the equivalent pairings. More commonly, sometimes WA and XA will use (or partially use) the same hidden unit as WC and XC, as this unit may have a low error term for this mapping, and thus the congruent trials will split up their mappings along the same lines as incongruent trials. Consider, for example, hidden unit 6 in Figure 3.6, which is shared between congruent stimulus 1 and incongruent stimulus 7. The sharing of hidden units between stimuli attenuate some of the disparity in influence between the cues and the stimuli, as this means it is possible to strengthen the cue-hidden unit link in the absence of one of the stimuli. The implications for the "switch cost" exhibited by the model of forming these different and standard structural solutions to "solve" the experiment will be discussed in more depth in Chapter 5.

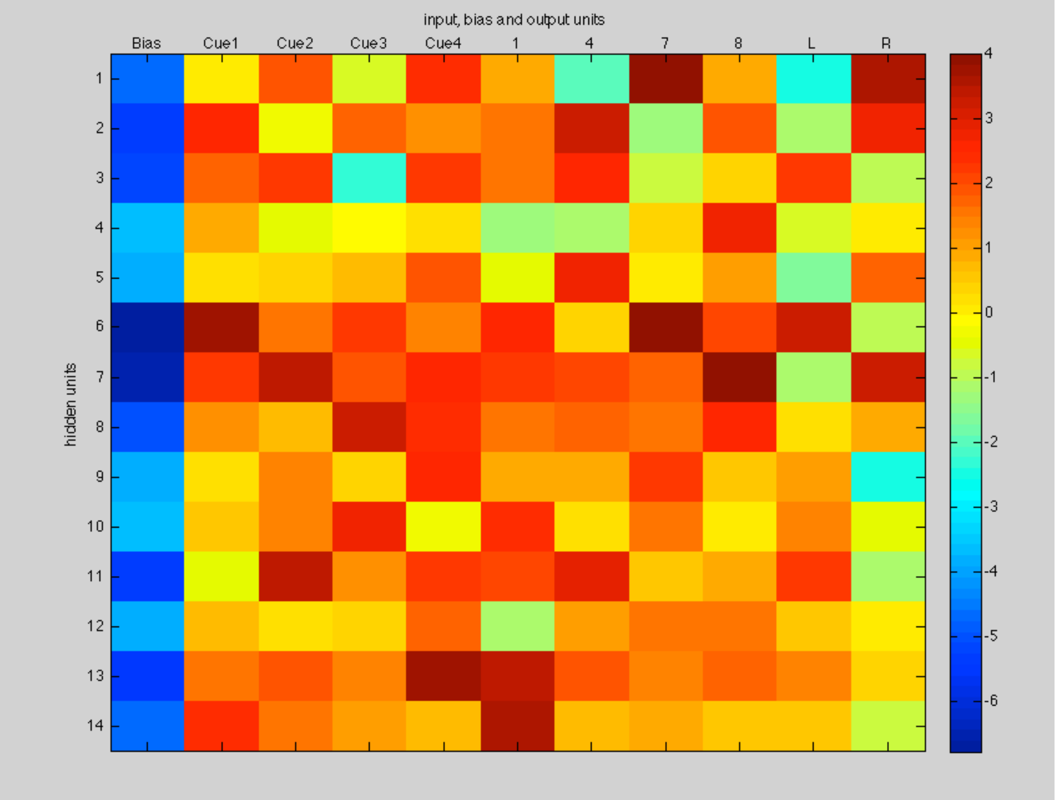


Figure 3.6 shows an example of the weights created by the end of block pair 5. Each row represents a hidden unit. The first column shows the negative bias that has developed for that unit, this is the value that the other inputs need to overcome in order for the hidden unit to become active. Columns 2-9 show the links to the various input nodes, whilst the final two column represent the link of the hidden unit to the left and right output units – in this case the network was trained such that 1 is always left, 8 is always right, 4 is left with Cue2 & Cue4 but right with Cue1 & Cue3, 7 is the right with Cue2 & Cue4 but left with Cue1 & Cue3. So it can be seen that hidden unit 1 represents the typical coding of an incongruent stimulus. Whilst hidden unit 14 represents a typical congruent mapping where the stimulus drives the link largely on its own regardless of the cue. The table also shows how the rest of the mappings, although following the general principles outlined above, are messier than in the model outlined by Honey & Ward-Robinson (2002).

The general agreement between the model and the data strongly suggests that associative learning is a viable mechanism for driving performance in the CSR group. This of course does not establish that this is the way performance is generated in the CSR group, but it serves as an existence proof that it could be.

Direct proof of the use of associative mechanisms rather than rule-based reasoning is hard to achieve. Rule-based accounts are often highly sheltered theories given the highly flexible nature of this type of symbolic system. This makes them difficult to disprove. It would also be naïve to think that participants in the CSR group in these experiments were not thinking or even hypothesis-testing whilst performing the experiment (see Chapter 7 for a fuller discussion of a possible rule-based alternative). There was nothing to prevent them from doing so and they were encouraged to perform at their best to gain the bonus payment. Indeed, some participants were clearly reasoning, as they stated they did not use the strategy they were given initially (and so were excluded). However, this makes the agreement between model and behavioural data even more impressive, as it suggests that even when participants probably engaged in processes other than those included in the model, the model still captures the significant features of their behaviour.

3.3.3.2 The "replaced" participants.

The participants we replaced because their introspective reports of their strategies did not match the strategies we intended to induce are also of interest. From Experiment 1 it was found that participants in the Tasks condition who did not claim to be using the tasks showed larger congruency effects, slightly reduced switch costs and no RISC effect. The data from the replaced participants in Experiment 2 agrees with this as the 6 participants who were replaced for not using the tasks showed larger congruency effects (187ms, 10.4%) and smaller switch cost (24ms, 3.8%) than the rest of the Tasks group. They also showed no RISC effect; with the switch cost at the long CSI (48ms) being larger than at the short CSI (0.3ms) in the RTs but not in the errors, where the switch cost at the long CSI (3.7%) was similar to the cost at the short CSI (3.9%). As with the equivalent replaced participants from Experiment 1, this suggests that for some participants at least, when they are instructed to use the tasks they opt for a more CSR-based strategy and so show a pattern of data more similar to those participants. The implication is that, in task-cueing experiments where CSR learning is relatively easy (e.g. one cue per task, small numbers of stimuli), the adoption of a CSR strategy by some participants could add unnecessary noise to data. It would seem worthwhile to debrief participants carefully in such experiments. And the use of small stimulus sets might best be avoided unless necessary for the issue under investigation. (Even with larger stimulus sets this issue may not be avoided as there may be a higher degree of generalisability between stimuli for which a given task rule would apply and so an associative alternative may still be present).

As with the results of the replaced participants from the CSR group of Experiment 1, the reciprocal case is harder to make: although Tasks participants may stop using the tasks whenever they please, in order for CSR participants to start using the tasks they first have to induce them. It is notable that Experiment 2 had more such participants (7 versus 4), maybe because participants in Experiment 1 were given a more prescriptive demonstration of how to approach the experiment (using the instruction sheet provided), whereas participants in Experiment 2 were simply told to learn by trial and error and so induce their own strategy. As with the replaced participants from Experiment 1, these participants showed smaller congruency effects (139ms, 12.5%) than the participants in the CSR condition. However, they did not show larger switch costs (24ms, 3.9%) or a RISC effect (long CSI: 58ms, 4.3%, short CSI: -9.2ms, 3.5%). It may be that, in both sets of replaced participants, although they induced and knew the tasks they weren't using them, or they induced them later in the experiment, such as during the transfer task when they were given new information. The next chapter examines the issues of whether or not participants who start off using CSR instructions can change to use the task-sets when instructed to do so and what the consequences of this are.

In conclusion this chapter has demonstrated that an associative model (APECS) can produce the qualitative pattern of switch costs, congruence, acquisition and transfer effects, shown by the CSR groups. It extended the empirical findings of the previous chapter and demonstrated that participants learning by trial and error with 2 distinct cues per task show the same general pattern or results as in the Monsell Experiment and Experiment 1. The strong agreement between the three experiments, which does not seem to depend on the amount of instruction given, also adds weight to an associative account of the CSR group's performance. The next chapter seeks to add more weight to this argument by using one of the properties of associative learning – the fact that it is automatic and can run in the background, whilst controlled processes cannot.

Chapter 4: Can participants change their approach?

This chapter reports an exploration of behaviour when participants are instructed to change between using Tasks and CSR strategies. In order to examine this I used a paradigm similar to the one outlined in the previous chapters, but with a lower switch probability. Hence I first establish that this variant of the task-cueing paradigm produced the same pattern of performance as in previous experiments and checked that APECS would also still produce the same pattern of data under these conditions. I then further extend the evidence suggesting that different processes are used by participants performing in the Tasks and CSR conditions by considering a *state trace* analysis (Bamber, 1979) of the data. I describe what happens when participants performing under one set of instructions are asked to change to the other strategy. This manipulation provides additional evidence on the nature of the processes contributing to CSR and Tasks. Specifically it suggests that the learning that mediates performance in the CSR condition is indeed automatic, whereas the Tasks strategy has to be consciously discovered or instructed.

4.1 Introduction

As outlined at the beginning of this thesis (1.1.1) human behaviour is commonly conceived of as being governed by two separate kinds of process that can be described as controlled and automatic. Dual process theories of learning and performance typically postulate that one (symbolic or propositional) system is controlled, resource limited and effortful whilst the other (associative) system is automatic, resource - unconstrained and effortless (e.g., McLaren, Green and Mackintosh, 1994, see Table 1.1). However, these theories have recently been challenged, with researchers questioning the need to postulate two systems at all (Mitchell, De Houwer and Lovibond, 2009; Newell Dunn and Kalish, 2011; Berry, Shanks and Henson, 2008). Whether or not two separate systems are conceptually necessary is one issue - but the main focus here is to ask whether anything is gained by characterizing these two classes of processes as controlled or automatic. This thesis has argued that performance in the CSR group could be driven by associative processes. One attribute commonly ascribed to associative learning is that is occurs automatically (McLaren, Green and Mackintosh, 1994). An implication is that while performance in the Tasks group is driven by controlled processes, the same automatic, associative learning processes may concurrently occur in this group as occur for the CSR group, given their identical stimulus and feedback experience. Hence Experiment 3 asked whether associative learning about the statistical structure of CSR experience can take place automatically, in the background, in a task-cueing experiment whilst the participant is actively engaged in controlled processing to select responses. It also asked the equivalent question for controlled, task-based performance - does later performance using the Tasks strategy show improvement after performance using the CSR strategy, or only after performance using the Tasks strategy? The results will be taken to indicate that the processes which mediate performance in the CSR and Tasks groups have different properties which divide along the lines of automaticity and control.

The paradigm used in Experiment 3 is an extension of the paradigm used in the previous experiments. The main difference was that instead of transferring the participants to novel stimuli, we instead transferred half of them to novel instructions, e.g. if they had originally been given CSR instructions they were transferred to Tasks instructions. The rationale for this manipulation assumes that participants are able to switch from one approach to the other because their use of task-sets to prepare and perform in the task is under voluntary control (De Jong, 2000). Although it is worth noting that most researchers in the task switching literature assume a role for both endogenous and exogenous processes (Kiesel et al, 2010, Monsell, 2003, Vandierendonck et al, 2010). The validity of the assumption that participants have complete control over their use of tasks will be discussed further below, but it is the assumption that there is some level of control which motivates the belief that participants could change their approach.

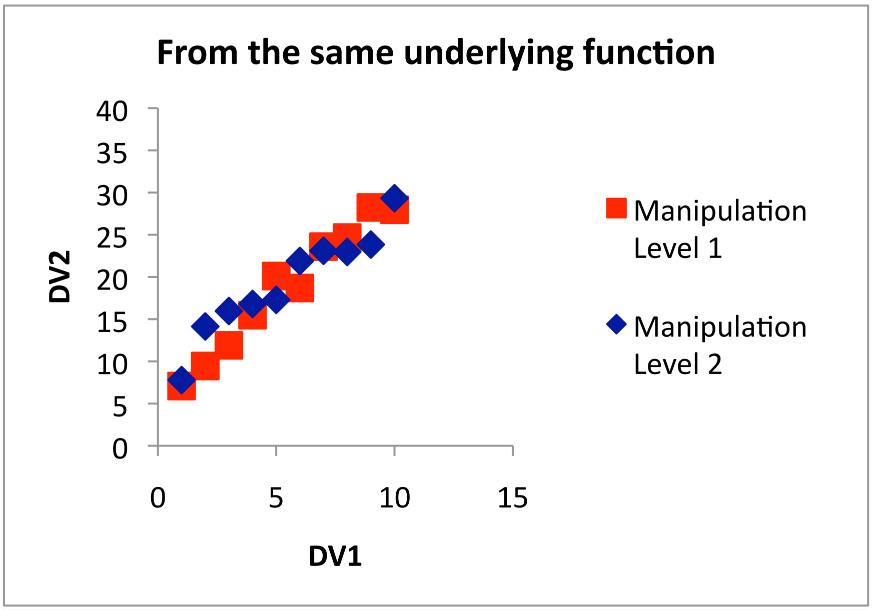
To examine whether our instructional manipulations were successful in altering the nature of processes mediating task performance, I once again examine two effects commonly found in the task cueing literature –the congruency effect (where responses are slower and less accurate to stimuli with different responses in the two tasks than when the response is the same) and the switch cost (where responses are slower and less accurate on trials when the task changes compared to the previous trial, than when the task repeats; Kiesel et al 2010; Monsell 2003). The previous experiments have indicated that under CSR conditions participants have a much larger congruency effect and a much smaller switch cost than participants under Tasks conditions. These two effects can therefore be used as markers to indicate whether participants are approaching the experiment using the processes engaged by Tasks or CSR instructions.

In addition to these two effects, as with the previous experiments, the reduction in switch cost (RISC) effect will also be considered. When participants have a longer time between the cue and the stimulus appearing, they can prepare for the task they are about to perform. This helps more on switch than repeat trials, and so typically participants using a Tasks approach have a smaller switch cost with a longer cue-stimulus interval (CSI). The Monsell Experiment and Experiments 1 and 2 found that this reduction is not observed when participants use CSR mappings. So if the instructional manipulation is successful this effect should closely correlate with the participant's instructional condition and will serve as an additional marker to measure the success of the manipulation.

4.1.1 State-Trace Analysis

I have been taking the differences in behavioural effects between Tasks and CSR conditions as markers of the approach in play, indicative of different processes occurring. But it remains possible that these differences reflect the use of the same processes modulated in such a way that an apparent dissociation emerges (Newell & Dunn, 2008). In order to examine this possibility, a different analytical technique was considered – state-trace analysis (Bamber, 1979). This is a more direct way of analyzing whether one or more processes are responsible for performance in two conditions. This is done by constructing a state-trace plot, in which two dependent variables are plotted against each other as a function of a "trace" variable which changes performance in a systematic way (e.g. block) and of a manipulation which is hypothesized to change the processes involved (in generating one or both dependent variables). If the manipulation modifies the relationship between the two dependent variables, the trace of the relationship between these two will be different at different levels of this manipulation, and distinct functions will appear on the state-trace plot. However, if the manipulation has no effect on the processes employed, the traces of the relationship between the two dependent variables should lie on one function (see Figure 4.1). This type of analysis has been used on paradigms as diverse as category learning (Newell, Dunn & Kalish, 2011), the face inversion effect (Loftus et al, 2004), and remember-know judgements in recognition memory (Dunn, 2008). I will use it here to assess whether one or more sets of processes need to be invoked to account for performance in Tasks and CSR conditions in the experiment. My prediction is that Tasks and CSR instructions will induce the use of two different sets of processes with different characteristics, producing distinct functions on state-trace plots, and that participants will be able to change from one set of processes to the other as a result of a change in instructional set.

I will also consider how best to characterize the processes involved in performance in task-cueing experiments under Tasks and CSR instructions in this experiment by considering how the transition from one set of instructions to another influences the measures that constitute my "markers". In this way I hope to confirm that participants can engage both controlled and automatic processing in task-cueing experiments.



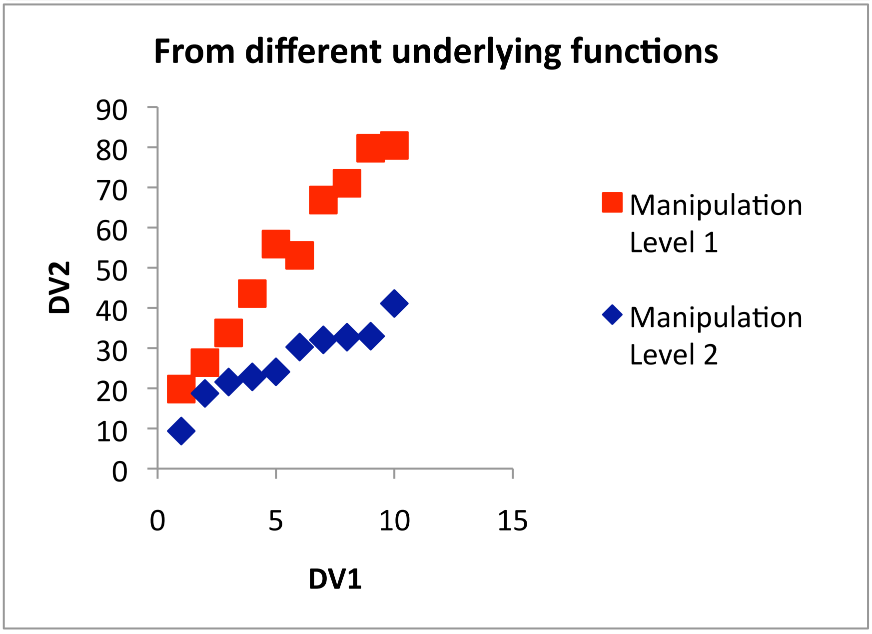


Figure 4.1 shows two illustrative and invented state-trace plots. These plots represent two hypothetical DVs plotted against each other. Red points come from one level of a hypothetical manipulation (such as instructions) whilst blue points come from another. The top one illustrates data that might be obtained if the same processes are involved at each level of the manipulation: the two traces of the variable lie on the same function. The lower plot illustrates a possible outcome if the manipulation taps into different processes: there are distinct functions for the two levels.

To summarise, the question directly addressed by Experiment 3 is: can participants change the way they approach the problem posed by the experiment simply by being instructed to do so? The expectation is that the answer to this question will be "yes": if the processes engaged in the Tasks condition are controlled processes, then by definition they are subject to our influence. But the processes more involved in the CSR condition may not be so amenable to instructional manipulation. Hence, we can speculate that when participants are using tasks they are also learning about the associative structure of the experiment in the background, and in this experiment look for evidence that this is, indeed the case. We can also ask if direct experience with using tasks is required to improve our ability to use them. The answers to these questions will provide more evidence on whether a dual or single process framework is required to give an account of human behaviour, and enable us to give a better account of the characteristics of the processes themselves.

4.2 Experiment 3

4.2.1 Method

4.2.1.1 Participants

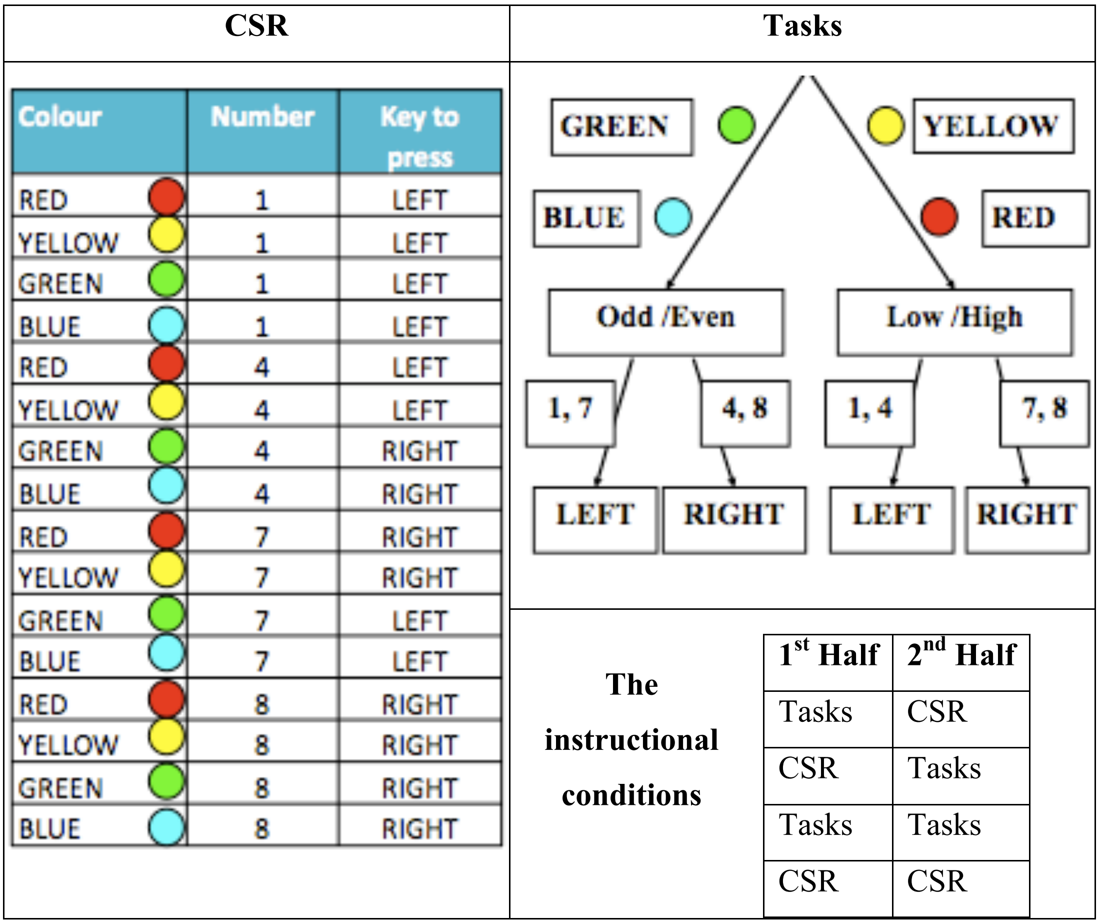
The participants were 90 undergraduates and postgraduates at the University of Exeter. Participants took part for payment contingent on their performance (payment was in the range: £5-£6.50).

4.2.1.2 Apparatus

The experiment was programmed in Matlab 2008a with Psychophysics Toolbox extension version 3.0.8 (Brainard, 1997; Pelli, 1997) and presented on an iMac with a 45cm screen. Participants were tested individually.

4.2.1.3 Design

Participants started under CSR or Tasks instructions, and their instructions were either changed halfway through or maintained through the second half. Thus participants were in one of four conditions generated by the factorial combination of the CSR and Tasks instructions as one factor with the first and second half of the experiment as the other (Figure 4.2).



CSR

Colour	Number	Key to press
RED 🔴	1	LEFT
YELLOW 🟡	1	LEFT
GREEN 🟢	1	LEFT
BLUE 🔵	1	LEFT
RED 🔴	4	LEFT
YELLOW 🟡	4	LEFT
GREEN 🟢	4	RIGHT
BLUE 🔵	4	RIGHT
RED 🔴	7	RIGHT
YELLOW 🟡	7	RIGHT
GREEN 🟢	7	LEFT
BLUE 🔵	7	LEFT
RED 🔴	8	RIGHT
YELLOW 🟡	8	RIGHT
GREEN 🟢	8	RIGHT
BLUE 🔵	8	RIGHT

Tasks

The instructional conditions

1st Half	2nd Half
Tasks	CSR
CSR	Tasks
Tasks	Tasks
CSR	CSR

Figure 4.2 shows examples of the instructions given to participants and the 4 possible instructional conditions.

The experiment also contained the same within-subjects factors as the Monsell Experiment and Experiments 1 and 2: task switch (switch versus repeat), response congruency (congruent versus incongruent), and CSI (long (1200 ms) versus short (100ms)).

The probability of switching task was one third (cf. Monsell and Mizon, 2006). With the exception of the first trial, the number of trials for a given task was equated in a block (24 out of 48 trials). There were now 49 trials per block. The stimuli appeared equally frequently in a given task, this was additionally constrained so that a given stimulus appeared twice on a switch trial and four times on a repeat trial in a given block.

4.2.1.4 Stimuli

The task cue was a yellow (RGB: 0,255,255), red (RGB: 255,0,0), blue (RGB: 0,0,255) or green (RGB: 0,255,0) circle, displayed centrally, of radius 112 pixels (~6.1° of the visual angle). Two of these colours were assigned to one task, and the other two to the other task; this was counterbalanced across participants. The stimuli used, which were counterbalanced across participants, were either the numbers 1,4,7,8 or 2,3,6,9 displayed in the centre of the cue in size 60 Courier bold font (~1.7° of the visual angle).

4.2.1.5 Procedure

Participants sat about 50 cm from the computer screen. They were given onscreen instructions that differed depending on whether the participant was in the Tasks or CSR condition initially. As for Experiment 1, blocks 1 and 2 were treated as practice blocks in which participants were provided with a piece of paper (see Table 4.1) reminding them how to approach the experiment and indicating the correct responses. The structure of these blocks was the same as in the rest of the experiment.

After Block 2 the piece of paper was removed and participants completed 8 blocks. After Block 10, depending on instructional condition, participants were either reminded of the instructions given originally or introduced to the other set. All participants received onscreen instructions and a piece of paper again detailing how they should now think about the experiment. Blocks 11-12 were thus treated as practice blocks. At the end of Block 12 the piece of paper was removed and a final 8 blocks completed, making 20 blocks in total.

Throughout the experiment, including the practice blocks, the CSI alternated by block between 1200ms or 100ms; the initial value was counterbalanced between participants. Participants responded using their index fingers on the “z” key for left and the “/?” key for right. The response keys were reversed for the high/low task for half the participants so that which stimuli were congruent and incongruent were counterbalanced. The response-stimulus interval (RSI) was the same in all blocks, so that for the short CSI the inter-trial interval (ITI) was 1600ms, whereas in the long CSI it was 500ms. If the participant made the wrong response, the message “WRONG!” was displayed for an additional 1500ms before continuation. If participants failed to respond within 4000ms “Time Out” appeared for 1500ms before continuation. At the end of a block the average RT, number of errors and bonus score (calculated as mean RT in ms divided by 10 plus 3 times the number of errors) for that block were displayed.

At the end of the experiment participants were interviewed to ascertain how they thought they had approached the experiment. Participants were only included in the analysis if they stated that when instructed to use the tasks they had actually used the tasks to help them, and if when instructed to approach the experiment using cue-stimulus-response mappings they had not used the task (or induced the task if it was the first half of the experiment).

4.2.2 Results

4.2.2.1 Participant exclusions

Instructional condition	Total Number Tested	Total Number Excluded	Reason for exclusion			
			Failed to use Tasks instructions	Failed to use CSR instructions	Did opposite instructions	Other
Tasks-CSR	33	17	6	9	1	1
CSR-Tasks	21	5	3	1	0	1
Tasks-Tasks	19	3	2	-	-	1
CSR-CSR	17	1	-	1	-	0
Total	90	26	11	11	1	3

Table 4.1 shows the number of participants excluded from each instructional condition and the reasons for these exclusions.

In total 26 participants were excluded (see Table 4.1). Participants were notably less likely to follow the experimenter's instructions in the Tasks-CSR group, from which 17 participants in total were excluded. Because of the large number of exclusions from this condition, the exclusions failed to conform to a uniform distribution across conditions, χ^2 (3, N=26) = 23.8, p<0.001. I shall return to this below.

4.2.2.2 Did performance in the first half conform to our expectations?

This experiment made a couple of subtle changes from the ones presented previously – most notably the probability of a switch trial changed from a half to a third. These changes mean it is important to consider first whether participants' performance in the first half of the experiment (before the instructional manipulation was introduced) were in line with the findings reported in the previous chapters. The results of the first half of the experiment were analysed in the same way as the results in the previous chapters. That is by using a mixed ANOVA with factors, block pair (2,3,4,5), CSI (long, short), congruency (congruent, incongruent), task switch (switch, repeat) and instructions (CSR, Tasks).

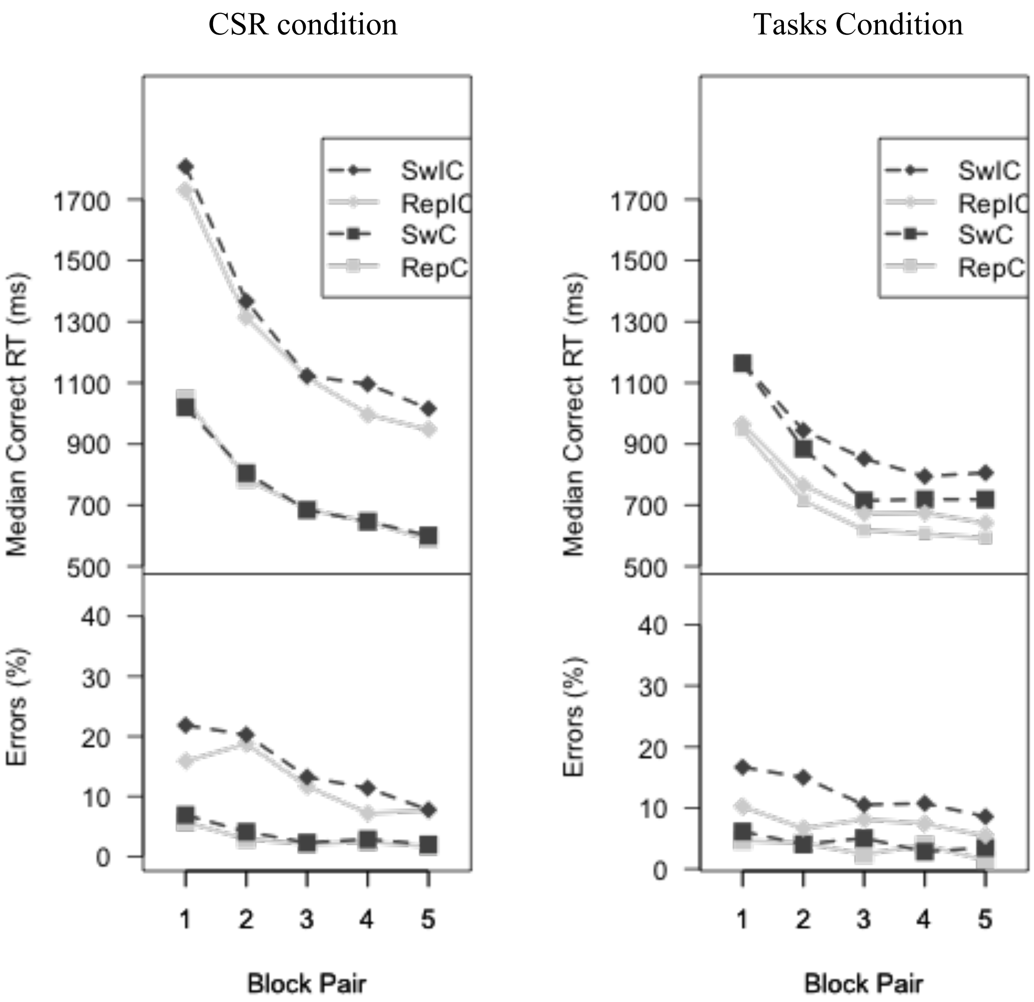


Figure 4.3 shows performance in the first half of the experiment over block pairs for participants in the CSR and Tasks conditions initially. Dark dashed lines represent switch trials while light solid lines represent repeat trials. Diamonds represent incongruent trials whilst squares represent congruent trials.

Tasks Switch and instruction. In line with our previous work, there was a larger switch cost under Tasks instructions (149ms, 1.8%) than CSR instructions (32ms, 1.1%), this was significant in RTs, F(1,62)=21.1, p<0.001 but not errors, F<2 (see Figure 4.3). The large switch cost under Tasks instructions was significant for both RTs, F(1,31)= 44.1, p<0.001, and errors, F(1,31)=8.24, p<0.01; the small switch cost under CSR instructions was also reliable for both RTs, F(1,31)=7.53, p<0.05, and errors, F(1,31)= 8.20, p<0.01.

Preparation and instruction The RISC effect was significantly larger under Tasks instructions (switch costs for short CSI: 201ms, 2.1%, and for long CSI: 96ms, 1.5%) than under CSR instructions (switch costs for short CSI: 24ms, 1.0%, and for long CSI: 40ms, 1.3%); this interaction was significant for RTs, F(1,62)= 17.7, p<0.001, but not errors, F<1. The RISC effect was only significant under Tasks instructions for RTs, F(1,31)= 22.4, p<0.001.

Participants in the Tasks condition also showed a general preparation effect, whereby if only the task-repeat trials are considered they were faster with a long-CSI (554ms, 4.6%) than with a short-CSI (784ms, 5.3%), F(1,31) = 83.3, p<0.01. For the same contrast the CSR group was also faster with the long- CSI (858ms, 6.6%) than at the short-CSI (912ms, 7.0%), F(1,31) = 8.76, p<0.001, but the effect was considerably smaller. These interactions were not significant in the errors F<1.

Congruency and instruction As we had expected based on our previous work, the first half of the experiment shows a significantly larger congruency effect under CSR instructions (441ms, 9.8%) than Tasks instructions (70ms, 5.4%) in both RTs, F(1,62)=99.9, p<0.001, and errors, F(1,62)= 8.10, p<0.01, see Figure 4.3. The congruency effect was significant under both CSR instructions in RTs, F(1,31)= 168, p<0.001, and errors, F(1,31)=70.4, p<0.001; and Tasks instructions in RTs, F(1,31)=19.9, p<0.001 and errors, F(1,31)= 43.8, p<0.001.

Acquisition effects Overall performance improved from block pair 2 to 5 (Figure 4.3) more for the CSR group than the Tasks group, RTs: F(3,186)= 3.54, p<0.05, Errors: F(3,186)=3.45, p<0.05. Participants in the CSR condition started with a much higher RT and error rate and so their decline was steeper than the Tasks condition. There was an improvement from block pairs 2 to 5 in the RTs both the Tasks condition, RTs: F(3,93)= 12.7, p<0.05, and the CSR condition, RTs: F(3,93)= 21.5, p<0.001. There was a reliable overall acquisition effect in the errors only for the CSR condition, F(3,93)=13.5, p<0.001, and only a marginally reliable effect for the Tasks condition, F(3,93)=2.67, p=0.053.

In line with previous experiments there was also a significant interaction between block pair, congruency and instructions, RT: F(3,186)= 5.41, p<0.01; errors: F(3,186)=4.67, p<0.01. As before this reflected a significant interaction of block pair and congruency in the CSR group, RT: F(3,93) = 5.74, p<0.01; errors: F(3,93)= 10.7, p<0.01, but not in the Tasks group (see Figure 4.3).

Other Significant effects In the Tasks condition there was a larger switch cost in the incongruent trials (159ms, 4.3%) than the congruent trials (134ms, 0.8%) but this modest interaction was significant only in the error rate, F(1,31)=9.29, p<0.01, but not for the RTs, F<2. In the Tasks condition the switch cost was significant for the incongruent stimuli in both the RTs: F(1,31)=38.0, $p<0.001$, and errors: F(1,31)=10.4, $p<0.01$ but it was only significant in the RTs for the congruent stimuli, F(1,31)= 31.1, $p<0.001$ and not for the errors, F(1,31)=1.52, p=0.227.

In the CSR condition there was a much larger switch cost in the incongruent trials (57ms, 1.8%) than the congruent trials (7ms, 0.6%); the difference was significant only in the RTs, F(1,31)=4.25, $p<0.05$, but not in the errors, $F<2$. In the CSR condition the switch cost was only significant for the incongruent stimuli, RT: F(1,31)=6.46, $p<0.05$; errors: F(1,31)=5.41, $p<0.05$, and not for the congruent stimuli, RT: $F<1$; errors, F(1,31)=1.69, p=0.203.

In the Tasks group there was a smaller congruency effect in the RTs for the long CSI (50ms, 5.2%) than a short CSI (86ms, 6.1%), this was significant in the RTs, F(1,31)=9.85, $p<0.01$, but the reverse effect was not significant in the errors, $F<2$. In the CSR group where there was no real difference between the long CSI (430ms, 10.5%) and the short CSI (453ms, 8.9%).

To summarise the data from the first half of the experiment was in broad agreement with the data presented in the previous two chapters. Those participants who began under the Tasks conditions showed the usual small and stable congruency effect and large switch cost which reduced with time to prepare. Those participants who began under CSR conditions, as in the previous experiments, showed large congruency effects, which changed with practice in a way which reflects the congruent stimuli quickly reaching asymptotic performance and the incongruent stimuli taking longer to learn. They also showed a small but reliable switch cost, which did not reduce with time to prepare and was limited mainly to the incongruent stimuli.

4.2.2.2.1 Modelling of the CSR behaviour, reprised.

Although the CSR group's behaviour was similar to that seen in the previous experiments, it is possible that the change in trial sequences resulting from a switch probability of one third rather than a half would affect how the associative model performed. In order to check this, the model from the previous chapter was run with the new sequencing constraints.

The same trial sequences were used as in Experiment 3, although only the first 5 block pairs were modelled. All the parameters of the model were kept constant from the previous chapter. The output of the model was examined in the same way as in the previous chapter by comparing the activation difference between the two output units and the target difference (target-trained).

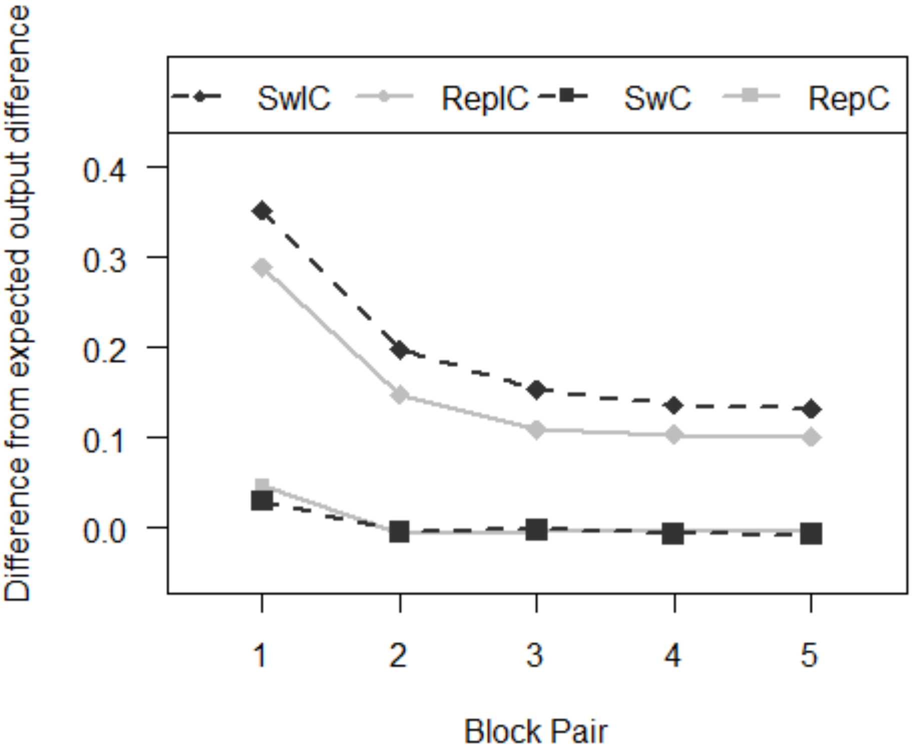


Figure 4.4 shows the model's performance for Experiment 3 by block pair by congruent/incongruent (light/dark respectively) and switch/repeat (dark and dashed/light and solid respectively).

4.2.2.2.2 Modelling Results

As with the results of the first half of the behavioural experiment the results were analyzed across block pairs 2-5 (as block pair 1 was treated as a practice block by the participants) using an ANOVA with just the factors, block pair, congruency, and switch, since the model cannot model CSI or the Tasks condition.

"Tasks switches" There was a significant effect of "task switch" in the model where performance was worse on switch than repeat trials (switch: 0.075, repeat: 0.055), F(1,31)=49.5, p<0.001 — see Figure 4.4.

Congruency There was a significant effect of congruency in the model where performance was worse for incongruent than congruent stimuli (congruent: -0.005, incongruent: 0.135), F(1,31)=168.5, p<0.001.

"Switch" by congruency. There was an overall "switch cost" only on incongruent trials, and it was significantly larger (0.0395) than for congruent trials (-0.00116), F(1,31)= 10.4, p<0.01. The switch cost was only significant for the incongruent stimuli, F(1,31)=57.6, p<0.001, and not for the congruent stimuli, F<1.

Acquisition effects. Overall performance reliably improved from block pair 2 to 5 (Figure 4.4), F(3,93)= 44.3, $p<0.001$. The two-way interaction between block pair and congruency was significant F(3,93)= 43.3, $p<0.001$ This interaction can be seen in Figure 4.4 which shows the congruent stimuli being learnt quickly whilst the incongruent stimuli took longer to learn.

The results of the model show a similar pattern to the modelling results in the previous chapter: a small but reliable switch cost restricted to the incongruent trials, and a large congruency effect which changes over time as the congruent trials are learnt faster than the incongruent trials. Evidently the change in the probability of a switch has had little effect on the overall pattern of the results.

As with the results of the previous chapter these results are in broad agreement with those seen in under CSR conditions, for the aspects of the experiment the model is able to encode. However, although it would be tempting to do a full comparison between Experiments 2 and 3 and the models of these experiments this would not necessarily be a valid comparison. The participants in Experiment 3 were given a piece of paper outlining the mappings as an attempt to more successfully manipulate which tactic they used, whilst those in Experiment 2 were made to learn by trial and error. The model of Experiment 3 nearly shows a bigger congruency effect (Model of Exp. 2: 0.113, Model of Exp. 3: 0.14), F(1,62)=3.46, $p=0.068$ and a much larger switch cost (Model of Exp. 2: 0.007, Model of Exp. 3: 0.02), F(1,62)=8.55, $p<0.01$, whilst the behavioural data suggests a larger congruency effect in RT but not in errors (Experiment 2: 308ms, 17.1%, Experiment 3: 441ms, 9.8%) but the same size switch cost (Experiment 2: 30ms, 2.24%, Experiment 3: 32ms, 1.1%). It is difficult, however, to infer whether this presents an issue for the model or is in some way driven by the use of the piece of paper.

I have now established that our markers for Tasks and CSR instructions are at least initially behaving as expected, based on previous behavioural data and, in the case of the CSR group, an associative model. Now I explore how these markers changed when the instructional set was altered.

4.2.2.3 Did participants modify their strategy as instructed?

The evidence just reported is consistent with participants using different processes under Tasks and CSR instructions in the first half of the experiment. In order to further test this, two state-trace analyses were conducted, to see whether this showed distinct processes and whether, when participants experienced a change in instruction, their data shifted from one function to the other. The errors quickly reach floor in this experiment they do not tend to change smoothly over blocks and so would not be a suitable trace variable (Newell, 2008) – hence a state-trace analysis will only be conducted on the RTs.

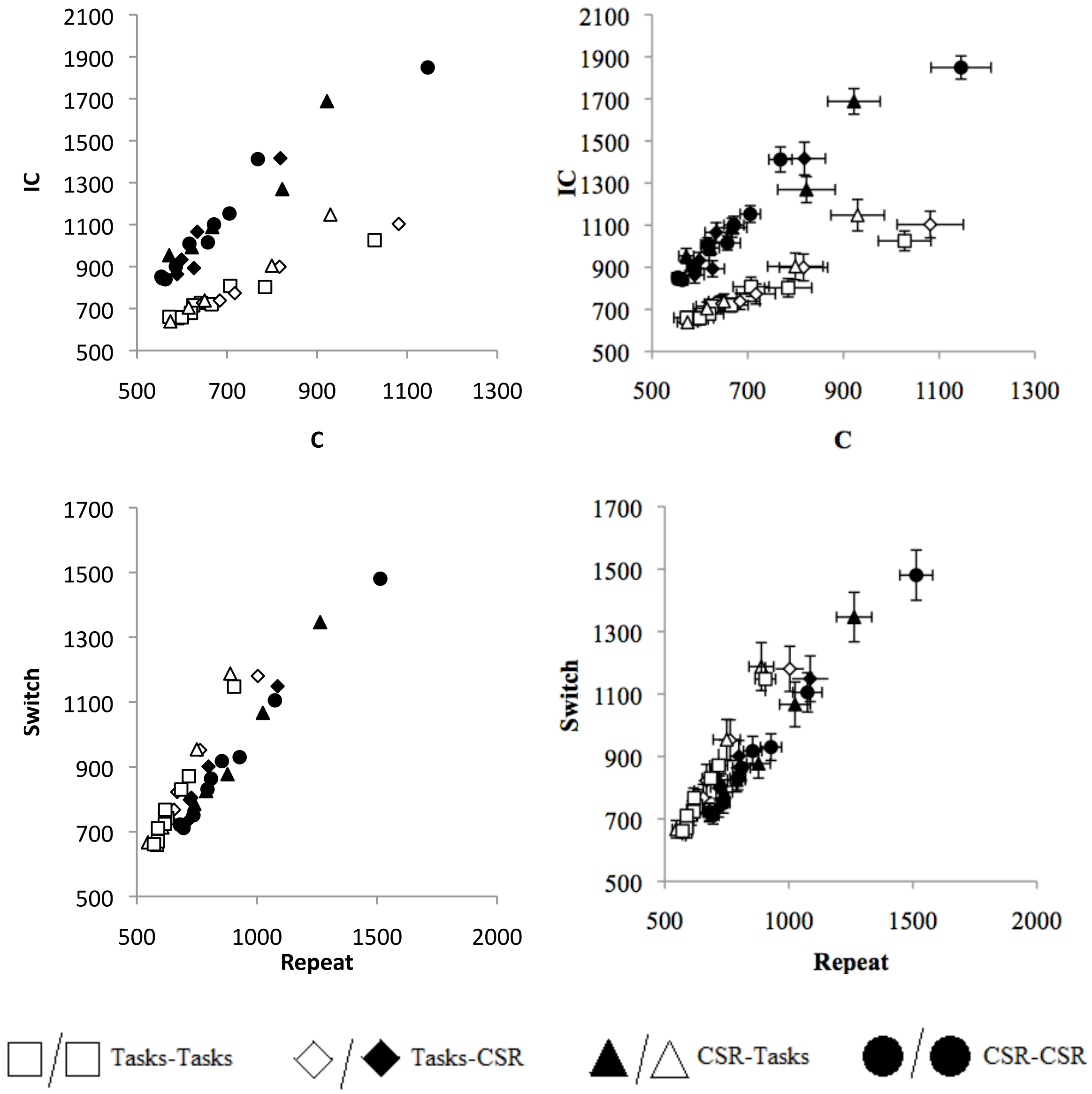


Figure 4.5 shows state trace plots. In each plot each point represents the mean median RT for a block pair (1-10). The shape of the point indicates the instructional group (squares: Tasks-Tasks, diamonds: Tasks-CSR, triangles: CSR-Tasks, circles: CSR-CSR) whilst whether the point is filled or open indicates the current instruction set. The panels at the top show incongruent RT against congruent RT and the panels on the bottom show switch RT against repeat RT. The panels on the left show the state-trace plots without error bars, whilst the panels on the right have error bars indicating standard errors over participants.

4.2.2.3.1 State-trace: Congruency

A state-trace plot was constructed by plotting incongruent RT against congruent RT for each instructional group for each block pair (including practice blocks). The plot shows a clear separation based on the instructions the participant currently had in play (point filled/open) in the relationship between congruent and incongruent trials over block pairs (Figure 4.5). The clearly separate functions for the two instructional sets suggests that the (un-excluded) participants did respond to the instructional sets by using different processes, and applied the instructed strategy in the second half even when they had been using a different strategy in the first half.

To test whether the separation of functions was reliable a stepwise multiple regression was run with the aim of predicting incongruent trial RTs. Congruent RT, current instruction type (CSR or Tasks) and instructional group (entered as 8 separate regressors, one for each half of each instructional condition) were the predictors. Compared to a model with just congruent RT predicting incongruent RT, introducing the current instructions variable as a predictor significantly improved the fit of the model with the R^2 going from 46.4% to 89.9%, $F(1,37)=159.663$, $p<0.001$. The high R^2_{adj} value 89.4% and the fit of the model, $F(2,37)= 164.811$, $p<0.001$, suggest the model with current instructions and congruent RT provided a very good prediction of incongruent RT. (The regression equation was estimated as Incongruent RT= 1.419×Congruent RT+361.784×current instructions-588.352.)

As it was found that current instructional condition was a reliable predictor, separate regressions for each instructional condition were run, with congruent RT and instructional group as predictors. In both cases these produced good models with just congruent trials as a predictor: CSR: R^2_{adj} 93.5%, $F(1,18)=272.459$, $p<0.001$, Incongruent RT=1.866×Congruent RT-171.257; Tasks: R^2_{adj} 90.9%, $F(1,18)=191.677$, $p<0.001$, Incongruent RT=0.969×Congruent RT+95.935[7].

---

[7] Although there are reasons for trying to predict incongruent performance from congruent performance, as one would expect more processing to be involved in incongruent trials, it is worth noting that this analysis also works when trying to predict congruent performance from incongruent performance. In other words, a regression run to predict performance on congruent trials from incongruent trials, current instruction type (CSR or Tasks) and instructional group (coded as above), showed that a model with current instructions was a more predictive model than a model with incongruent performance alone, R^2 going from 46.4% to 85.0%, $F(1,37)=95.3$, $p<0.001$. The high R^2_{adj} value 84.2% and the fit of the model, $F(2,37)= 104.847$, $p<0.001$, suggest the model with current instructions and incongruent RT provided a very good model of congruent RT. The relationship was estimated as Congruent RT= 0.598×Incongruent RT-231.136×current instructions+465.583. Again, it was found that when dividing the data up by current instruction type, good models of congruent performance were produced based just on incongruent performance without reference to group: CSR: R^2_{adj} 93.5%, $F(1,18)=272.459$, $p<0.001$, Congruent RT=0.503×Incongruent RT+128.505; Tasks: R^2_{adj} 90.9%, $F(1,18)=126.010$, $p<0.001$, Congruent RT=0.963×Incongruent RT-28.9.

4.2.2.3.2 State-trace: Switch

A similar analysis examining the relationship between switch and repeat RTs was constructed, plotting switch RT against repeat RT for each instructional group for each block pair (Figure 4.5)**.** This plot also shows clear separation in the trace over block pairs based on the current instructions the participants were using, similar to the result obtained with congruency.

A stepwise multiple regression to predict RT on switch trials, by entering repeat RT, current instruction type (as 8 separate regressors, one for each half of each condition) and instructional group as predictors. Compared to a model with just repeat RT predicting switch RT, introducing current instruction type as a predictor significantly improved the fit of the model with the R^2 significantly changing from 88.2% to 95%, $F(1,37)= 51.0$, $p<0.001$. The high R^2_{adj} value of 94.8% and the fit of the model $F(2,37)= 355$, $p<0.001$ suggest that this model provided a very good fit to the data. (The regression equation was estimated as switch RT= 1.057×repeat RT - 115.818×current instructions+222.544.)

As current instructions were found to be a significant predictor, separate regressions were run for Tasks or CSR instructions. It was also found that separate regressions for each instructional condition produced good models, Tasks: R^2_{adj} 96.2%, $F(1,18)=476.776$, $p<0.001$. Switch= 1.359×repeat-99.312; CSR: R^2_{adj} 97.7%, $F(1,18)=811.101$, $p<0.001$. Switch= 0.96×repeat+75.693[8].

Thus, we have good evidence for different sets of processes driving performance in our two different instructional conditions, Tasks and CSR. Now I look at how changing from one set of instructions to the other affected performance on this task. I

---

[8] Although there are reasons to be trying to predict switch performance from repeat performance, as one would expect more processing to be involved in switch trials, it is worth noting that this analysis also works when trying to predict repeat performance from switch performance. In other words a regression run to predict performance on repeat trials from switch trials, current instruction type (CSR or Tasks) and instructional group (coded as above), showed that a model with current instructions accounted for more variance than a model with switch performance alone, R^2 going from 88.2% to 96.0%, $F(1,37)=72.4$, $p<0.001$. The high R^2_{adj} value 95.8% and the fit of the model, $F(2,37)= 445.487$, $p<0.001$, suggest the model with current instructions and switch RT provided a very good model of repeat RT. The relationship was estimated as Repeat RT= 0897×Switch RT+113.994×current instructions-174.017. Again it was found that when dividing the data up by current instruction type, well fitting models of repeat performance were produced based just on switch performance and not on group: CSR: R^2_{adj} 97.7%, $F(1,18)=811.101$, $p<0.001$, Repeat RT=1.019×Switch RT-58.125; Tasks: R^2_{adj} 96.2%, $F(1,18)=476.776$, $p<0.001$, Repeat RT=0.709×Switch RT+95.203.

focus for the most part on between-group comparisons using data from the second half of the experiment. One reason for this is that I expect some changes from the first to second half of the experiment to occur even in those groups that do not experience a change of instructions (i.e. practice effects). Using a group that had the same experience in the first half of the experiment, but no change of instructions for the second half, controls for this. The other reason is that by comparing groups, we are able to see if performance in a group that has experienced a change of instructions is comparable to that of other groups in either the first or second half of the experiment, as we shall see. I consider the "markers" for task performance in due course, but start by looking at absolute RTs and error rates.

4.2.2.4 Effects of previous experience in a different instructional condition.

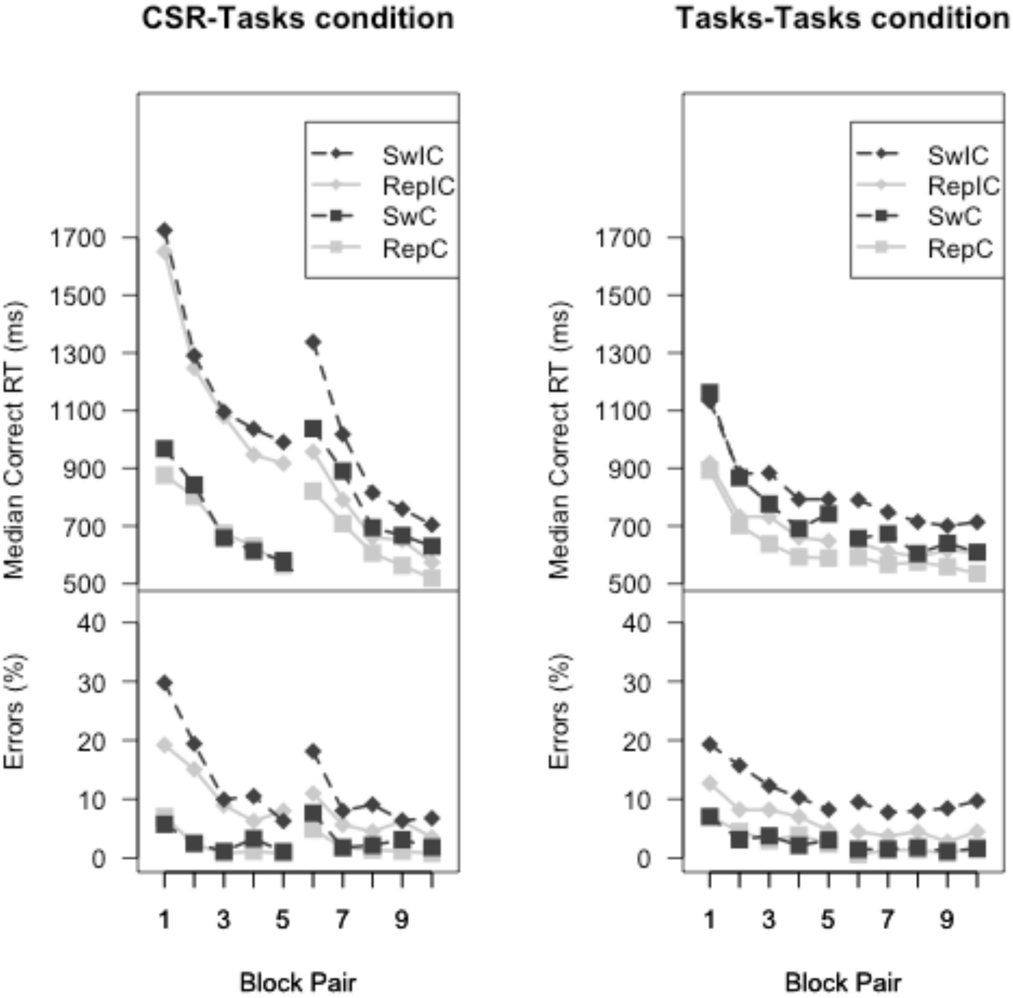


Figure 4.6 shows the performance of the CSR-Tasks group and the Tasks-Tasks group over block pairs for the whole experiment. Dark/dashed lines indicate switch trials, whilst light/solid lines indicate repeat trials. Diamonds indicate incongruent trials whilst squares indicate congruent trials.

4.2.2.4.1 CSR-Tasks

A contrast was run to test whether participants in the CSR-Tasks group, when changed to Tasks instructions (block pair 6: 763ms, 10.1%), picked up at roughly the same level of overall performance as both the Tasks-Tasks (block pair 1: 792ms, 10.6%) and Tasks-CSR (block pair 1: 858ms, 8.3%) groups started at, i.e. contrasted the CSR-Tasks group's performance against the Tasks-Tasks and Tasks-CSR groups' performance, see Figure 4.6. There was no significant difference which suggests that the CSR-Tasks group performs the same after they start to follow the Tasks instruction as they would have without 10 blocks following the CSR instruction; RT: F(1,45)=0.042, n.s. Errors: F(1,45)=0.075, n.s. As the first block pair and block pair 6 were practice blocks and therefore prone to a lot of variance the block pairs 7 and 2 were also checked (CSR-Tasks: 851ms (4.2%); Tasks-Tasks: 792ms (8.2%); Tasks-CSR: 858ms (6.7%)), the same contrast also produced a non-significant result; RT: F(1,45)=0.065, n.s., Errors: F(1,45)=1.09, n.s. Thus, I am not able to reject the hypothesis that when switched to Tasks instructions from CSR instructions, participant's performance is the same as when first beginning the experiment under Tasks instructions.

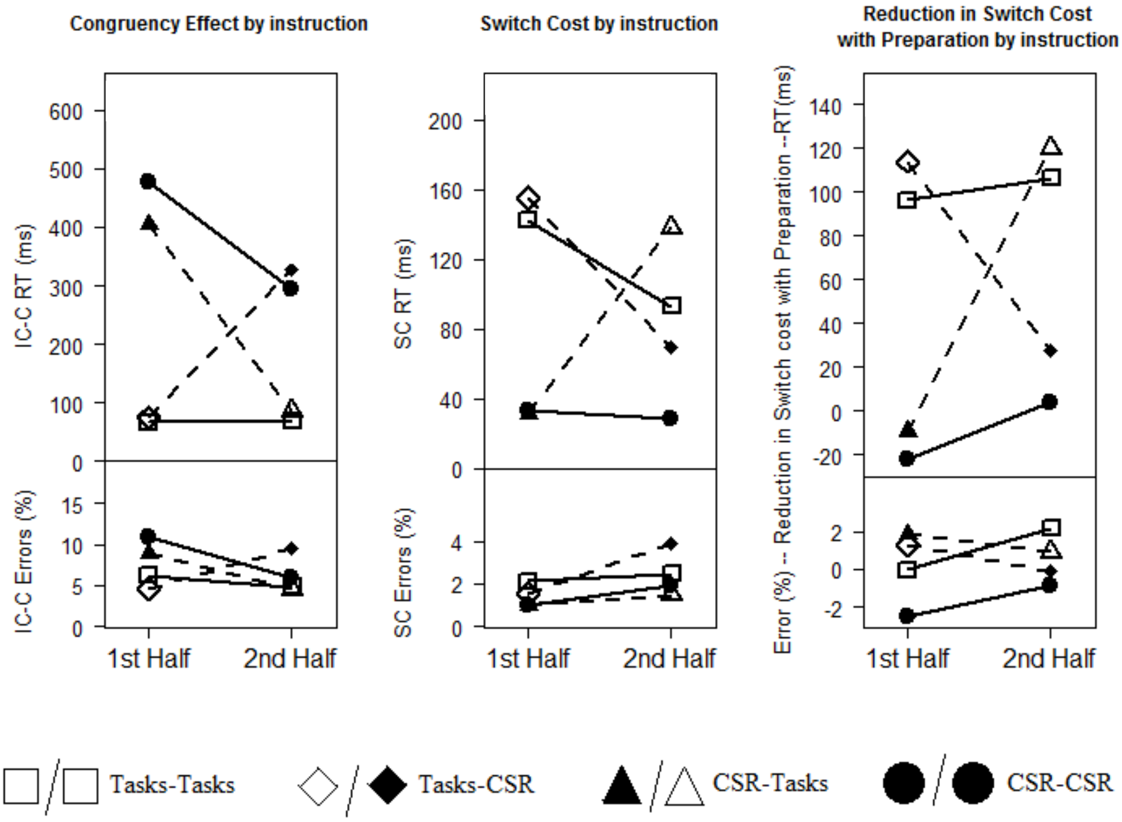


Figure 4.7 shows the size of the congruency effect, switch cost and RISC effect by instruction by half of the experiment. The shape of the point indicates the instructional condition: squares represent the Tasks-Tasks condition, circles the CSR-CSR condition, diamonds the Tasks-CSR condition and triangles the CSR-Tasks condition. Filled points represent times participants are using CSR instructions, whilst open points represent times when participants are using Tasks instructions. A dotted line represents a change in instruction, whilst a solid line represents a continuation of the same instructions.

There was evidence that the instructional manipulation changed the pattern of performance as expected. Participants who changed to the Tasks from CSR had a larger switch cost in the RTs the second half (170ms, 2.4%) compared to the CSR-CSR condition (34ms, 1.9%), F(1,30)= 24.0, $p<0.001$, and a smaller RT congruency effect: Tasks-CSR (113ms, 5.4%), CSR-CSR (307ms, 6.1%), F(1,30)= 17.3, $p<0.001$, see Figure 4.7. These contrasts were not significant for errors. Participants in the CSR-Tasks condition in the second half of the experiment showed a switch cost in the RTs more similar to the switch cost seen in the first half of the experiment for the Tasks-Tasks group, and the effect for congruency was for it to reduce to the same level as that seen in the Tasks-Tasks group, see Figure 4.7. These results confirm the impression that performance in the CSR-Tasks group in the second (Tasks) half of the experiment was similar to that under Tasks instructions in the first half of the experiment in other groups.

4.2.2.4.2 Tasks-CSR

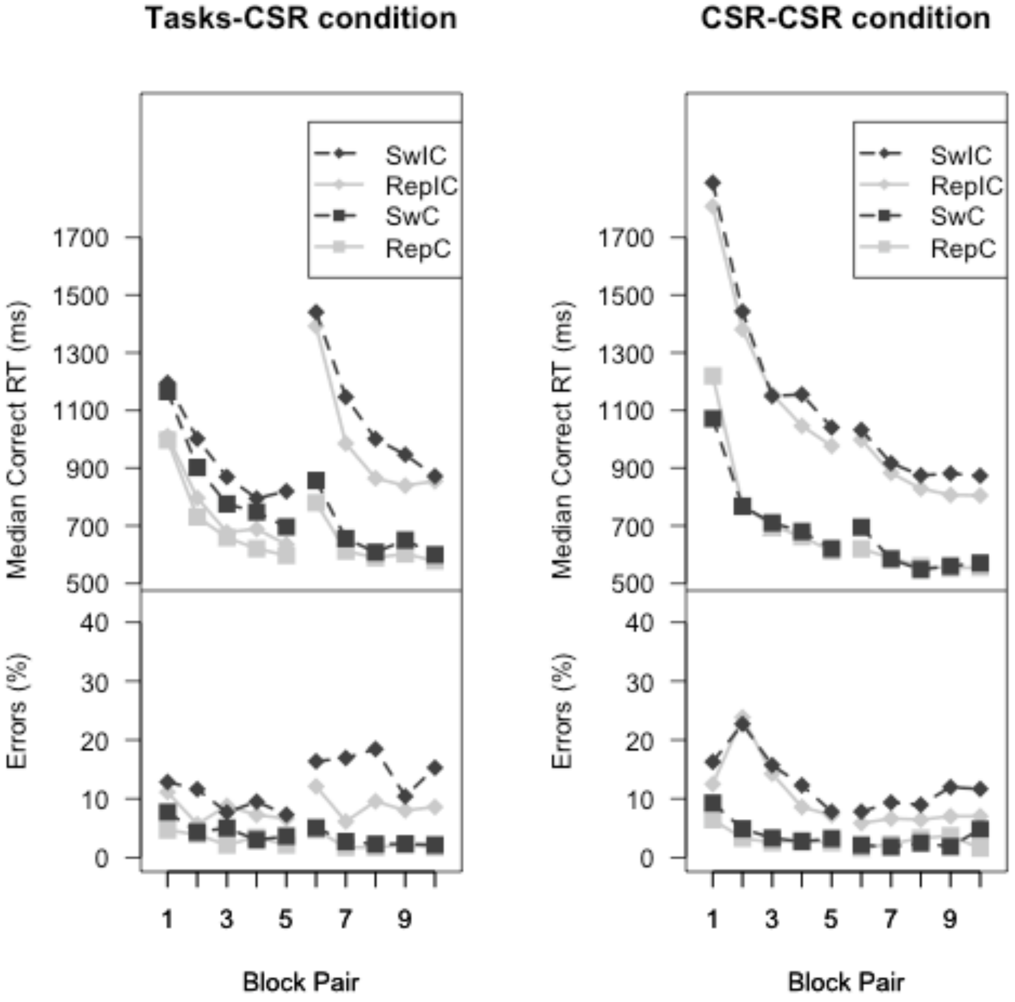


Figure 4.8 shows the performance of the Tasks-CSR and CSR-CSR groups over block pairs. Dark/dash lines indicate switch trials, whilst light/solid lines indicate repeat trials. Diamonds indicate incongruent stimuli and squares indicate congruent stimuli.

A contrast was run to test whether participants in the Tasks-CSR group (block pair 6: 1117ms, 9.6%), when they changed to CSR instructions, performed at a similar level to the initial performance of the CSR-CSR (block pair 1: 1497ms, 11.0%) and CSR-Tasks (block pair 1: 1305ms, 14.2%) groups, i.e. contrasted the Tasks-CSR group's performance against the CSR-CSR and CSR-Tasks groups' performance, see Figure 4.8. There was a significant difference in the RTs, F(1,45)=8.29, p<0.01 indicating substantially faster performance in the group that changed to CSR instructions after half a session of performing under Tasks instructions, than in the other groups starting under

CSR instructions. The difference in error rates was in the same direction, but not reliable, F(1,45)=1.40, p=0.24. Performances in the second block pair in the relevant half-session were also compared with the same contrast (Tasks-CSR: 850ms (6.9%); CSR-CSR: 1090ms (13.2%); CSR-Tasks: 1046ms (9.8%)), and again gave a significant contrast in the RTs, F(1,45) = 5.59, $p<0.05$, and a near significant contrast in the errors: F(1,45)=3.17, p=0.082. Hence, although 10 blocks experience of performing under CSR instructions did not confer a detectable advantage in performing under Tasks instructions, 10 blocks experience of performing under Tasks instructions did confer a substantial advantage in using CSR mappings, compared to groups starting from scratch.

Further evidence that the change in instruction set worked in this group is seen in the modulation of congruence effects. The congruency effect in the RTs in the second half of the experiment was much larger (381ms, 10.0%) than that for participants in the Tasks-Tasks group (72ms, 5.3%), F(1,30) = 24.0, $p<.001$. The switch cost was numerically smaller in the RTs in the Tasks-CSR group (Tasks-CSR: 68ms, 3.9%; Tasks-Tasks: 96ms, 2.8%) but not significantly so, F(1,30)=1.91, p=0.177; the opposite difference in errors was also non-significant, F(1,30)=1.95, p=0.173

The Tasks-CSR group had a congruency effect in the RTs in the second half (326ms, 9.6%) that was more similar to the CSR-CSR group in their second (307§ms, 6.1%), $F<1$, n.s., than in their first half (523ms, 10.0%) F(1,30)=4.41, $p<0.05$. Thus the evidence seems to indicate that performance in the Tasks-CSR group in the second half of the experiment more closely paralleled that of the CSR-CSR group in the second than in the first half of the experiment.

4.2.2.4.3 Reduction in Switch Cost with preparation

In the CSR-Tasks group there was a significant increase in the size of the RISC effect from the 1st half (-9.4ms, 1.8%) to the 2nd half (120ms, 1.0%) in RTs, F(1,15)=18.9, $p<0.001$, but not in the errors, $F<1$. In the Tasks-CSR group there was a near significant decrease in effect size from the 1st half (114ms, 1.6%) to the 2nd half (27.4ms, 1.6%) in RTs, F(1,15)= 3.42, p=0.08, but not in errors, $F<1$. Hence participants shifting to Tasks from CSR instructions begin to show substantial benefits of preparation in the reduction in switch cost, and participants switching to CSR from Tasks show a marked reduction in such benefits. This too attests to the success of the instruction manipulation in shaping performance.

4.2.3 Discussion

The results of the first half of Experiment 3 were in agreement with those of the Monsell Experiment and Experiments 1 and 2 and the results of the modelling I have reported; the small changes in switch probability and cue type were inconsequential. When using CSR instructions, participants demonstrated a larger congruency effect and a smaller switch cost than in the Tasks condition. Additionally participants in the Tasks condition reduced their switch cost with a longer CSI unlike those the CSR condition. The implications of these findings were considered in Chapters 2 and 3. A novel contribution of the present experiments was the state-trace analysis, which provided compelling additional evidence that different underlying processes are involved in the two instructional conditions.

The effects that require greater discussion in this chapter are those that occurred due to participants changing from one set of instructions to the other. What types of processes were responsible for the switch cost and congruency effects under CSR and Tasks instructions, to what extent can these processes be described as automatic, and was there any evidence of any process running in the background?

4.2.3.1 CSR-Tasks

Let us first consider the participants who change to the Tasks condition having started on the CSR condition. They have had an equal amount of experience with the experiment to the Tasks-Tasks group, but whilst they have been responding in the first half they have not been explicitly using a Tasks strategy. The results show that participants who change to using the task-sets show performance levels similar to those seen in the first half for the Tasks-Tasks and Tasks-CSR condition, suggesting that it is necessary to have direct experience with using tasks to be able to improve performance with them. This suggests both that the processes engaged under CSR conditions are easily overridden by Tasks instructions and that no task-related processing was occurring under CSR instructions[9].

---

[9] It is, of course, logically possible that that there is some interference from the CSR strategy and some benefit from having been in the CSR strategy that are cancelling each other out (e.g. generic practice effects).

The CSR-Tasks group's trace of switch against repeat and incongruent against congruent for the second half falls on the traces of the other groups who were using Tasks instructions, suggesting that they had changed to processes that were now the same as other participants using the Tasks instructions. Additionally, we can also consider the RISC effect. In the CSR-Tasks condition there was no RISC effect for the CSR portion of the experiment, but when the Tasks instructions were introduced to them, they exhibited a RISC effect of similar size to those participants who had used Tasks instructions initially (see Figure 4.6). This suggests that using task-sets to prepare is relatively easy to engage.

4.2.3.2 Tasks-CSR

In contrast, now consider the behaviour of participants who change to the CSR condition from the Tasks condition. Participants who changed to using CSR instructions showed performance that was significantly better than the initial performance of the CSR-CSR group and the CSR-Tasks group, and more like that seen in the second half for the CSR-CSR group, suggesting that in this case it is not necessary to have direct experience of using CSR instructions to learn about the associative structure of the experiment.

The results for the Tasks-CSR group have additional implications for the development of task automaticity. These participants showed a (marginally significant) reduction in their RISC effect when they transferred from Tasks to CSR instructions. Combined with the uneven pattern of exclusions, with many more from the Tasks-CSR condition than the other conditions, there is some suggestion that once participants have learnt to use tasks, stopping doing so is hard. This may be because the participants have had enough experience under Tasks instructions that a cue comes to automatically retrieve a task-set. Indeed, a few participants who were excluded reported that they tried to stop using the Tasks but were unsuccessful. This suggests a role for exogenous task based processes as well as endogenous processes, as although there seem to be issues with the participant's ability to completely stop using the task-sets, there is also evidence to suggest that participants can, at least to some extent, stop using the task-sets. This lends support to proposed hybrid account of task switching, which allows a role for both exogenous and endogenous processes e.g. Vandierendonck et al (2010).

Taken together these results tell us something about the relationship between experience and learning. The CSR instructions seem to tap into the underlying associative structure of the task – as evidenced by their broad agreement with the modelling results. The data indicates that participants are also learning something about this underlying associative structure whilst they are operating under Tasks conditions. However, the reverse is not true, because participants clearly needed experience with using the task-set control strategy in order to improve their performance with it; merely having experience with the input-output relationships did not help participants to deploy the Tasks strategy. It would also seem (from the difficulties in abandoning the Tasks strategy) that the cue comes to retrieve some part of the task–set relatively automatically, suggesting an automatic system that not only picks up the statistical regularities present in the environment but also those generated by mental states. This asymmetry suggests the need to ascribe different properties to the processes explaining behaviour under CSR and Tasks instructions; an automatic, associative system capable of running in the background for the former and a controlled, effortful system that has to be deliberately engaged (of course supplemented by the automatic/associative system) for the latter.

It is worth acknowledging that the uneven exclusion pattern may lead to a sampling bias for the Tasks-CSR condition. It may be that participants who are able to stop using tasks are in some way different. These participants may be naturally better at learning CSR associations, and would thus be able to more easily change to this approach. Thus, the fact that their performance is better than the initial performance of the CSR-CSR and CSR-Tasks condition could be explained by this natural tendency, rather than the fact they have learnt the associations in the background. Although this may be the case, it is worth noting that the performance of the Tasks-CSR group was not different in the first half from the performance of the Tasks-Tasks group, which suggests that they are not simply better at the experiment as a whole. Further investigation would be needed to see if there are individual differences in people's ability to approach the experiment using CSR or Tasks instructions.

In summary, these results suggest that the associative structure of the experiment can be learnt in the background, but that in order to learn how to use the task-set control strategy, direct experience with it is required. There is also some suggestion of developing automaticity within the task-set control strategy through the learning of cue-task associations. The conclusion is that these results offer support for the distinction between controlled and automatic processes being a useful one. This ties in with the notion advanced in earlier chapters that the behaviour of the CSR group is driven by an automatic associative learning system. However, although there is evidence that these automatic associative processes can run in the background, whilst the participant is engaged using task-sets, it is also important to consider what else is being automated and what remains controlled.

Chapter 5: Why is there a "switch cost" under CSR conditions?

This chapter addresses the issue of why there is a small, yet reliable, switch cost in the CSR conditions in the previous experiments, even when I have removed participants who induced the tasks. In order to address this question it examines one way to manipulate the associative structures formed under CSR conditions. The design of the paradigm was changed from the Normal condition in which the biconditional discrimination (incongruent stimuli) shared cues with the component discrimination to one in which separate cues were used for the biconditional discrimination (congruent stimuli) and the component discrimination. The model outlined in Chapter 3 predicted that a much smaller switch cost would be generated in this new condition and that it would disappear when direct cue-stimulus repetition trials (i.e. trials on which both the cue and the stimulus repeated) were removed from the data. These predictions were assessed in two experiments.

5.1 Introduction

Given that the variable "task switch" in some sense has no meaning when applied to participants in the CSR condition, as they are supposed to be responding to cue and stimulus combinations, not switching between tasks, it may seem peculiar that this contrast shows a significant effect under CSR instructions in the majority of the experiments presented in this thesis (the Monsell Experiment, Experiments 2 & 3) and a marginally significant effect in the same direction in Experiment 1. However, as I noted in Chapter 3, an associative model (APECS) also produces a switch cost in the task-cueing paradigm used so far (see Table 5.1 for the design of the task-cueing paradigm). To rule out the idea that CSR participants (or some of them, some of the time) were covertly construing the situation in terms of tasks, it is worth investigating *why* the model produces a "switch cost", and hence illuminate *why* a significant "task switch" effect occurs in the behavioural data in what should be the absence of any task-switching.

This investigation may also help establish whether APECS is a good model of the CSR condition. Up until this point the model has simply been used to add weight to the argument that the CSR participants could be learning associatively. Our confidence in APECS as a good model of the CSR condition would be greatly enhanced if it makes non-obvious a priori predictions that are confirmed by experiment. Examining how the model "solves" the paradigm may help in understanding which part of this "solution" generates the "switch cost" and hence help lead to a prediction about how to manipulate the paradigm in such a way as to eliminate the "switch cost".

		"Task 1" Cues		"Task 2" Cues	
		W	X	Y	Z
	A	L	L	L	L
Stimuli	B	R	R	R	R
	C	L	L	R	R
	D	R	R	L	L

Table 5.1 shows the mappings present in the standard task-cueing paradigm with 2 cues per task and 4 stimuli. W, X, Y, and Z are the task cues (colours in these both Experiment 4 & 5), whilst A, B, C, D are the stimuli (numbers in Experiment 4, symbols in Experiment 5). A & B are congruent stimuli (component discrimination), whilst C & D are incongruent stimuli (biconditional discrimination). This is the same structure as in all the previous experiments.

One-way to examine how APECS "solves" the paradigm is to consider the associative structure it has formed by the end of training. In Chapter 3, I described the structures created by the model to "solve" the task-cueing paradigm outlined in Table 5.1. I noted that the congruent trials tend to (but do not always – see below) form mappings that are carried by one hidden unit each, where the stimulus has a much stronger link to the hidden unit than any of the cues. I also noted that for the incongruent trials, cues for the same "task" tended to use the same hidden unit to carry their mapping. In both cases this was because these cues (all four in the case of the congruent stimuli and the two for the same task in the case of the incongruent stimuli) are functionally equivalent and so the same hidden unit tends to be the *best* unit for carrying a given mapping. This is a similar solution to that implied by the model proposed by Honey & Ward-Robinson (2002), see Section 1.3.4.

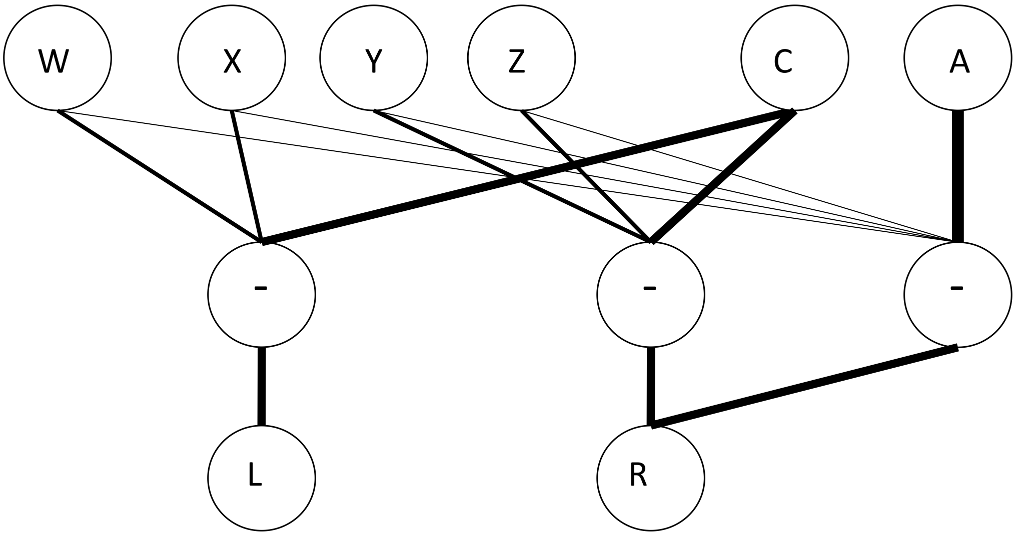


Figure 5.1 shows an idealised version of the typical structures formed between the input and hidden layer for a congruent stimulus A and an incongruent stimulus C. The thickness of the line indicates the relative strength of the links. B & D would use different units but in a similar vein to A & C respectively.

The associative structure of the solution for the incongruent stimuli means that even if the cue changes, but the incongruent stimulus repeats on a task repeat trial, then there will be a benefit as these two consecutive trials would use the same hidden unit to carry their mapping. This benefit exists because the link between that hidden unit and the output unit has more recently been strengthened. The fact that this link has been recently strengthened means that it has an advantage in activating the correct output unit over mappings that use different hidden units. As the same hidden unit would not be used consecutively for incongruent stimuli on a task switch this would lead to a repetition benefit. In other words there is a benefit in transitions between equivalent cue+incongruent stimulus combinations and these types of transitions can only happen on “task repeat” trials, whilst “task switch” trials are by necessity always transitions between non-equivalent cues. Note that the same is not necessarily the case for the congruent stimuli as all cues are equivalent; here there would just be an advantage for repeating congruent stimuli. This advantage would not map onto the “task repeat/switch” distinction in the way it does for the incongruent stimuli.

However, the model does not always learn to respond to the congruent trials by using just one hidden unit to carry all four possible mappings. This is because the congruent trials share cues with the incongruent trials. As noted in Chapter 3, this means that the congruent trials sometimes latch onto the same hidden units as those used by the response and "task" appropriate incongruent trial. For example for task cues W & X, A (a congruent stimulus) might share a hidden unit with C (an incongruent stimulus) that also requires a left response with those cues, see Table 5.1. For cues Y & Z, A might share a representation with D (a different incongruent stimulus which has the opposite response mappings to C) that requires a left response for these cues. Then, as with the incongruent trials, when there is a congruent stimulus followed by an incongruent stimulus on a task repeat trial that requires the same response, this acts as described above, using the same hidden unit to carry the mapping and so benefiting from the recent reinforcement of the connection between that hidden unit and that output unit. In other words the congruent stimulus is treated as being equivalent to the incongruent stimulus for this "task", and so the boundaries determining when it can benefit from using the same hidden unit fall along the lines of the division created with the variable "task switch".

Similar arguments have been used with APECS before. Indeed this was one of the mechanisms that Le Pelley & McLaren (2001) mentioned as being involved in the balancing act of backwards blocking (AB+| A+| B? see Section 1.4.2) – although they did not claim this mechanism drove the behavioural backwards blocking effect; rather, they claimed that the mechanism acted in the opposite direction resulting in the weakness of the backwards blocking effect. They argued that there was an advantage in having recently trained up the link between a given hidden unit and its output unit, as described above, as this would lead to overexpectation rather than backwards blocking. In other words training of A+ after AB+ would lead to a stronger link from the hidden unit that carries AB to the outcome. This would mean that presenting B after training with A+ would lead to higher expectation of the outcome than in a control condition when A+ was not presented.

However, Le Pelley & McLaren (2001) also showed that backwards blocking effects could be accounted for by APECS. Its ability to do this was largely driven by the way in which the bias developed during the ITI. When the model is receiving no input, due to the logistic activation function the hidden layer would have an activation of 0.5. This would be fed down to the output layer, and as the mapping had been recently trained, this would then activate the recently trained response. However, this would lead to a large negative error for that hidden unit, which would cause a negative bias to develop in order to counteract this over-expression. Thus recently activated hidden units, at the end of the ITI, would have a larger negative bias. This modification was incorporated into the model as the way of preserving the learning from interference. It means that it would now be much harder for any new stimulus to use the same hidden unit to carry the mapping, as the hidden unit now has a much larger bias, which would need to be counteracted in order for the activation to be able to propagate to the output layer. In relation to retrospective revaluation, this means that repeated presentations of A+ following AB+ leads to it being much harder for B to activate the hidden unit (as its link to the hidden unit has not been trained as much as A's link and now its link alone may not be enough to overcome the bias) and so responding to B decreases.

So why would the anti-backwards blocking effect dominate here, in a model that has been shown to give backwards blocking? There is one major difference between the way in which backwards blocking experiments are typically conducted and the task-cueing experiments described in this thesis. In a typical backwards blocking experiment, such as Le Pelley & McLaren (2001), the stimulus types are "blocked", that is the model/participant is presented with (multiple presentations of) AB+ and then subsequently with (multiple presentations of) A+. This means that the bias has multiple (usually consecutive) ITIs over which to build up, and so when B is then presented it has to contend with a significant bias built up over many ITIs, not just a single one. In the task-cueing paradigm all the different trial types are intermixed, meaning that the bias has fewer ITIs over which to develop between presentations of differing stimuli types that use the same hidden unit. As the learning rate for the bias is slower (0.1) than the learning rate for activated hidden units (0.8) this means that the increase in the connection between the hidden unit and the response dominates, when considering trial by trial transitions between stimuli. The development of the bias does, however, play some role as increasing the ITI does reduce the switch cost and removing the ITI from the model[10] leads to a greatly increased switch cost, due largely to the advantage for direct cue-stimulus repetitions (trials on which both the cue and stimulus repeat).

In addition to the "repetition benefits" described above there is also a mechanism for a "switch cost". In Chapter 3 it was noted that, in terms of the associative structures the model develops in response to the problem, the stimulus has a

---

[10] Note that removing the ITI also makes the model worse at learning the incongruent stimuli as the model is now less able preserve the previous learning, leading to greater interference between the incongruent mappings.

stronger connection to the hidden unit carrying the mapping than does the cue (Figure 5.1). This was because the stimulus occurred and was mapped through the same hidden unit more frequently than the cues, as equivalent cues use the same hidden unit to carry the mapping. Honey et al (2010) also noted that this differential strength in the links would occur in their model when explaining how it would solve IDS/EDS discriminations (Section 3.3.3.1). The differential influence of the cue and the stimulus means that, in the case of an incongruent stimulus, presenting just the stimulus, unlike presenting just the cue, will to some extent activate both the chosen hidden units for that stimulus, i.e. the one consistent with the cue which is also presented *and* the one carrying the relationship with the cues for the opposite "task". When the hidden unit for the opposite "task" is activated, it predicts the wrong outcome. This means that the hidden unit will receive a relatively high negative error score and so the connections to this unit will be weakened and inhibited by the growth of this negative bias. So the trial sequence WC-L, YC-R for example should be especially detrimental to the performance of the model, unlike the sequence WC-L, WD-R. There is much more stimulus-based interference in the model than cue based interference, and so presenting WC-L will lead to more unlearning/inhibiting of YC-R than WD-R. This means there is a cost to switching relative to repeating tasks, as mappings in the opposing task have to be re-learnt when the task switches back to them.

In support of this part of the explanation one could consider what happens when there is just one cue per task. With one cue per task one might expect a larger switch cost as now a cue switch is confounded with a task switch (Logan & Bundesen, 2003). However, as far as the model is concerned, the link from the stimulus and the cue to the chosen hidden unit should be nearly equal as they are both paired equally often with a given hidden unit. This would mean that there should be very little difference between the transition WC-L, YC-R and WC-L, WD-R as now the cues and stimuli would equally activate other hidden units they were using to carry their mappings. In support of this, if the model is run with just one cue per task, the switch cost is much reduced as compared to a model with 2 cues per task and allowing cue repeats, but is still significant. This is the case when only one cue per task is used even if direct cue-stimulus repetitions (trials on which both the cue and stimulus repeat) are excluded as there are other possible "repetition benefits" still available as the congruent and incongruent trials still share cues, which means they may also share hidden units to carry a mapping.

To summarise, the above suggests two possible drivers of the switch cost: the fact that equivalent cues are using the same mappings and the fact that these cues are shared between congruent and incongruent stimuli. To eliminate the two possible drivers of the switch cost there needs to be no cue equivalence and separate cues are needed for the congruent and incongruent trials. So I constructed a design where the incongruent trials have just one cue per task and this cue differs from the cues used for the congruent trials (Table 5.2). To clarify, consider that in Table 5.1 W & X are equivalent to each other, as are Y & Z in that they both predict the same outcomes for C & D. But in Table 5.2, neither W& X nor Y&Z are equivalent to each other, as they do not predict the same outcomes for a given stimulus, as they no longer occur with the same stimuli.

		"Task 1" Cues		"Task 2" Cues	
		W	X	Y	Z
	A	L	-	L	-
Stimuli	B	R	-	R	-
	C	-	L	-	R
	D	-	R	-	L
	E	L	-	L	-
	F	R	-	R	-
	G	-	L	-	R
	H	-	R	-	L

Table 5.2 shows new design in which there are now separate cues for the congruent stimuli, A, B, E, F, (component discrimination) and the incongruent stimuli, C, D, G, H (biconditional discrimination). The '-' indicates that trial type cannot occur.

If the above explanation of why the model produces a switch cost has any merit, then running the model with the design outlined in Table 5.2 should not produce a "switch cost". However, disallowing cue repeats – as I did in all the other experiments and simulations presented so far -might introduce problems with the new design. If no cue repeats were allowed then a "task repeat" trial would necessitate a change from a congruent stimulus to an incongruent stimulus. This may introduce unwanted sequential effects in this condition as compared to the standard condition used before. For this reason, the trial sequences now used allowed cue repeats and stimulus repeats.

This means that it is now possible on a “task repeat” trial but not a “task switch” trial to have a direct cue-stimulus repetition trial, i.e. one on which both the cue and the stimulus repeat. It is well known that direct cue-stimulus repetitions of the same trial type can possess an advantage in terms of the speed and accuracy with which one can respond. Hence, just allowing direct cue-stimulus repetitions may lead to a “switch cost” because this advantage can only occur on “repeat” trials. So for the modelling below the results will be considered both with these trial types present and with this type of transition excluded from the analysis.

5.2 Predictions of the Model

For ease of discussion, the structure outlined in Table 5.1 will be referred to as the “Normal” condition as it is the norm from previous experiments, whilst the modified version (Table 5.2) will be called the “Separate” condition as there are now separate cues for the different levels of congruency. It is worth noting that the number of stimuli in Table 5.2 is twice that in Table 5.1. This was done so that the total number of mappings to learn in each condition was the same – this leads to more directly comparable levels of performance. If only 4 stimuli are used in the Separate condition the model shows much faster learning simply because of experiencing more repetitions of a given cue- stimulus pair.

5.2.1 Modelling Method

The modelling was done in an identical way to the model of Experiment 2 in Chapter 3 and the model of Experiment 3 in Chapter 4, except for the sequencing used. The sequencing used was similar to that used in Experiment 3, in that there were 49 trials per block, it was constrained such that only a third of trials were “switch” trials and the number of times a given stimulus appeared in a given task on a “repeat” or “switch” trial was constrained. However, the cue sequence was generated differently, and the experiment was only 10 blocks long as it contained no transfer condition.

In order to generate the two conditions, first a sequence was generated using the “task switch” constraints above with 8 different stimuli. The cues were then assigned to the stimuli based on their congruency level, to generate the sequences seen by participants in the Separate condition, in which there was one cue per task per congruency level (Table 5.2). For participants in the Normal condition, where there were two cues per task, the four stimuli they were going to see were randomly assigned to that cue sequence with the constraint that within a block each stimulus appeared equally often in a task and twice as often on a “repeat” than a “switch” trial. This led to participants in the Normal condition experiencing the task with the structure in Table 5.1. The generation process meant that participants in each condition were yoked so that they saw exactly the same cue sequence. This means that when the comparison is being made between “switch” and “repeat” trials the same cue transitions are being considered. As mentioned above this cue sequence differed from the cue sequences previously used as it allowed cue repetitions to occur. The model was run 24 times in each condition.

5.2.2 Modelling Results

Figure 5.2 shows the modelling results for the two conditions. The striking difference between these two conditions seems to be in the size of the “switch cost” in the incongruent stimuli. For the Normal condition there is a sizeable “switch cost” whilst for the incongruent stimuli in the Separate condition it is much smaller. The manipulation seems to have reduced the switch cost considerably, as was anticipated.

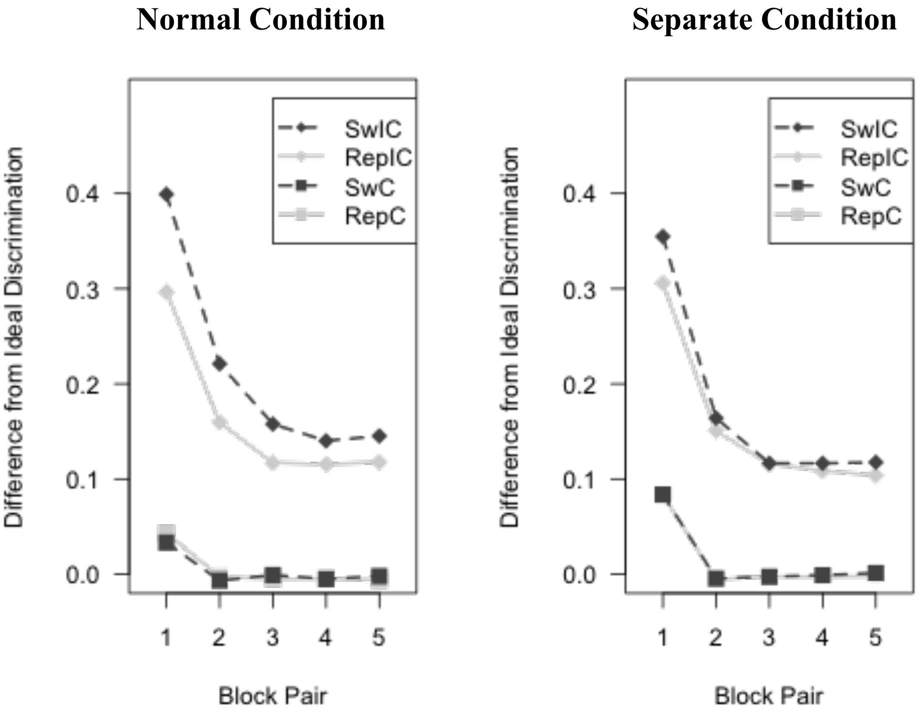


Figure 5.2 shows the results of modelling the Normal and Separate condition divided up by block, congruency and "task switch". Dark dotted lines represent "switch" trials whilst light solid lines represent "repeat" trials. Squares represent congruent stimuli and diamonds represent incongruent stimuli.

5.2.2.1 Modelling results including direct cue-stimulus repetitions

In order to assess whether the manipulation had produced a significant impact on the switch cost a mixed ANOVA was run on the data with the factors, "task switch" ("switch" or "repeat"), congruency (congruent, incongruent), block pair (1,2,3,4,5) and condition (Normal or Separate). This ANOVA was followed up with separate ANOVAs on each condition where appropriate.

"Task switches" and condition. There was a significantly larger "switch cost" in the Normal condition (0.025) than in the Separate condition (0.007), F(1,46)= 18.3, p<0.001. These "costs" were reliable in both the Normal condition, F(1,23)=49.5, p<0.001 and the Separate condition, F(1,23)=11.0, p<0.01.

"Switch" by congruency and condition. The "switch cost" in the Normal condition was significantly larger for incongruent trials (0.053) than the effectively zero cost on congruent trials (-0.003) and this difference was larger than the difference between the "switch cost" seen in the Separate incongruent trials (0.014) and congruent

trials (0.0), F(1,46)= 30.7, $p<0.001$. Separate analyses revealed that the "switch cost" was significantly larger in the incongruent trials for the both the Normal F(1,23)=105, $p<0.001$, and Separate conditions, F(1,23)=8.17, $p<0.01$. The switch cost was significant in both the Normal and Separate conditions for the incongruent stimuli, Normal: F(1,23)=89.6, $p<0.001$; Separate: F(1,23)=11.2, $p<0.001$, but not for the congruent stimuli, $F<1$.

Congruency and condition There was a significantly larger effect of congruency in the Normal condition (0.185) than the Separate condition (0.135), F(1,46)=10.5, $p<0.01$. The congruency effect was reliable in both the Normal, F(1,23)=297, $p<0.001$, and Separate conditions, F(1,23)=160, $p<0.001$.

Acquisition effects and condition. Overall performance reliably improved from block pairs 2 to 5, with the improvement being faster (initially) in the Separate condition than the Normal condition (Figure 5.2), F(4, 184)= 9.35, $p<0.001$. There were significant effects of block pair in both the Normal condition, F(4,92)=360, $p<0.001$ and the Separate condition, F(4,92)=573, $p<0.001$.

The three-way interaction between condition, block pair and congruency was significant F(4, 184)= 2.79, $p<0.05$. There were significant block pair and congruency interactions in both the Normal condition, F(4,92)=158, $p<0.001$ and the Separate condition, F(4,92)=112, $p<0.001$. This interaction can be seen in Figure 5.2, which shows the congruency effect initially being much larger in the Normal condition and then the difference between the conditions diminishing. This initial difference seems to be largely due to the congruent stimuli taking longer to learn in the Separate condition than the Normal condition.

Although there was a significant difference in the size of the "switch cost" between the two conditions the "switch cost" was not eliminated in the Separate condition. A possible reason for this was suggested above. By allowing cue repetitions (to avoid sequential effects that might be caused by alternating congruent and incongruent trials) the sequences used now include "direct cue-stimulus repetitions", that is trials on which both the cue and the stimulus repeat from the previous trial. These direct cue-stimulus repetitions can obviously only occur on "task repeat" trials. The model would be primed by just having learned the same cue-stimulus combination and so would be more accurate at generating the correct response. Hence this direct cue-stimulus repetition benefit may be driving part of the "task switch" cost seen in the model. In order to ascertain if this was the case, direct cue-stimulus repetitions were removed from the data generated by the model before further analysis.

5.2.2.1 Modelling results without direct cue-stimulus repetitions

It is worth noting that removing direct cue-stimulus repetitions does lead to a subtly different composition of possible trial transitions with respect to "task" switch/repeat trials in the Normal and Separate conditions. Specifically, removing direct cue-stimulus repetitions means that in the Separate condition it is not possible on a "task repeat" trial for the stimulus to repeat because any stimulus repetition would be a direct cue-stimulus repetition as there is only one cue per stimulus per task. However, it is still possible in the Normal condition to have a stimulus repetition, i.e. XA-L, WA-L on a task repeat trial because the same stimulus can occur with the other “task” cue. One may consider this to be the cause of any switch cost as obviously a stimulus repeat in the same task may afford an advantage. However, in the Separate condition it is possible to have a cue repeat with a functionally equivalent stimulus repeat, XA-L, XE-L, whereas this is not possible in the Normal condition as there are no functionally equivalent stimuli. Although superficially these transitions seem different, this may not in fact be the case if the interpretation of how APECS solves the paradigm has any merit. If this interpretation holds, then in the Normal condition equivalent cues may use the same hidden unit for a mapping and in the Separate condition equivalent stimuli may use the same hidden unit for a mapping (see below). This means that if there is a repeat advantage for such transitions then to this extent they are balanced across the conditions.

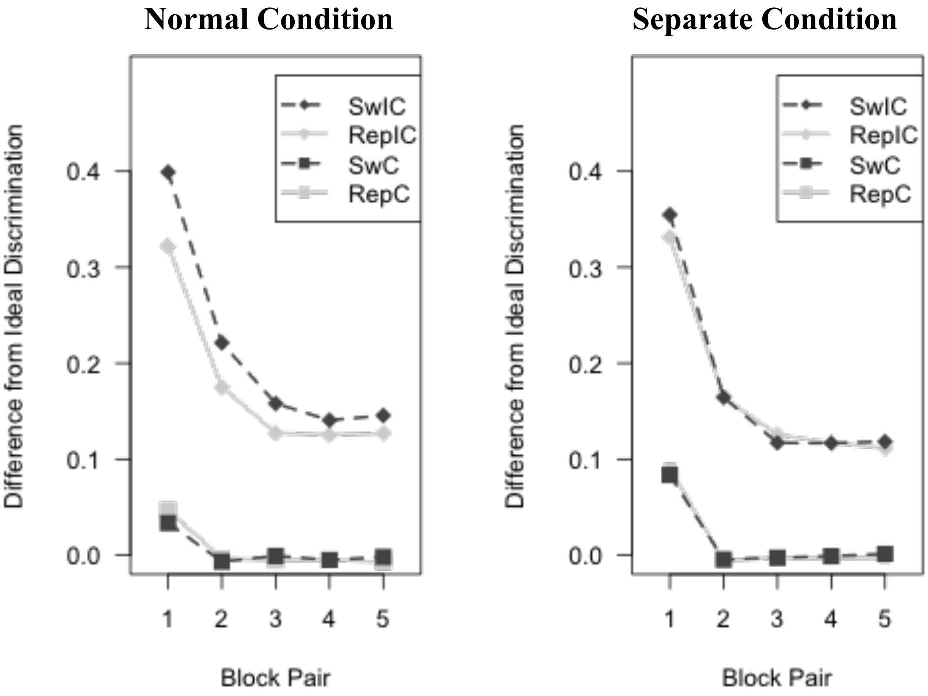


Figure 5.3 shows the results of modelling the Normal condition and the Separate condition when direct cue-stimulus repetitions are removed from the analysis. The graphs are plotted as in Figure 5.2.

By contrasting Figure 5.2 with Figure 5.3 one can see that the "switch cost" in the incongruent trials has reduced in both conditions. There is still a "switch cost" in the Normal condition but, as predicted, there now seems to be very little evidence of a "switch cost" in the Separate condition. The same ANOVA technique was used to analyse the data.

"Task switches" and condition There was a significantly larger "switch cost" in the Normal condition (0.018) than in the Separate condition (0.001), F(1,46)= 17.6, p<0.001. These costs were reliable in the Normal condition, F(1,23)=28.8, p<0.001, but were not in the Separate condition, F(1,23)=0.361, n.s.

"Switch" by congruency and condition. The "switch cost" in the Normal condition was significantly larger for incongruent trials (0.039) than the congruent trials (-0.003) and this difference was larger than the difference between the "switch cost" seen in the Separate incongruent trials (0.004) and congruent trials (-0.001), F(1,46)= 26.6, p<0.05. Separate analyses revealed that the "switch cost" was significantly larger in the incongruent trials for the Normal condition F(1,23)=64.3, p<0.001, but not the Separate condition, F(1,23)=1.49, n.s. The switch cost was significant for the incongruent trials in the Normal condition, F(1,23)=56.6, p<0.001, but not for the

congruent trials, F<1. The switch cost was not significant in either the incongruent or congruent trials in the Separate condition, F<1.

Congruency and condition There was a significantly larger effect of congruency in the Normal condition (0.191) than the Separate condition (0.140), F(1,46)=10.4, p<0.01. The congruency effect was reliable in both the Normal, F(1,23)=299, p<0.001, and Separate conditions, F(1,23)=155, p<0.001.

Acquisition effects and condition. Overall performance reliably improved from block pair 2 to 5, and was faster initially in the Separate condition than the Normal condition (Figure 5.3), F(4, 184)= 9.83, p<0.001. There were significant effects of block pair in both the Normal condition, F(4,92)=387, p<0.001 and the Separate condition, F(4,92)=600, p<0.001.

The three-way interaction between condition, block pair and congruency was significant F(4, 184)= 3.12, p<0.05. There were significant block pair and congruency interactions in both the Normal condition, F(4,92)=169, p<0.001 and the Separate condition, F(4,92)=115, p<0.001. This interaction can be seen in Figure 5.3, which again, as with the comparison including direct cue-stimulus repetitions, shows the congruency effect initially being much larger in the Normal condition.

5.2.3 Modelling Discussion

When direct cue-stimulus repetitions are removed from the analysis, the Separate condition no longer exhibits a detectable switch cost whilst a significant effect is still present in the Normal condition. This was as predicted based on the analysis of the associative structure created by the model given in the introduction. This result suggests that the reason the model shows a switch cost is related to there being two equivalent cues per "task" shared between congruent and incongruent stimuli. If human participants also show a similar pattern of results that would suggest that something about the way in which APECS solves the problem is similar to the way in which people learn how to respond.

As I had changed the precise nature of the sequencing it is first worthwhile to consider whether APECS solved the problem in the same way that I outlined in the introduction to this chapter. Figure 5.4 shows an example of the final connection strengths in the Normal condition. Notice how, in general, Cues 1&3 and Cues 2&4, which are equivalent to one another for the incongruent stimuli, tend to use more similar hidden units to carry a mapping. Figure 5.4 also shows the type of mappings outlined in the introduction (see Figure 5.1). For example, hidden unit 3 illustrates a hidden unit used to carry just a congruent mapping (8->R) and hidden unit 10 illustrates a hidden unit carrying just an incongruent mapping for one task (4 with Cues 1&3->R). Additionally, there is evidence of congruent stimuli hijacking incongruent representations. For example, hidden unit 14 was carrying an incongruent mapping (Cues 1&3 7->L) and was hijacked by a congruent stimulus (1->L). In conclusion, changing the sequencing as outlined above does not seem to affect the associative structures formed significantly.

An example of final connection strengths in the Normal condition

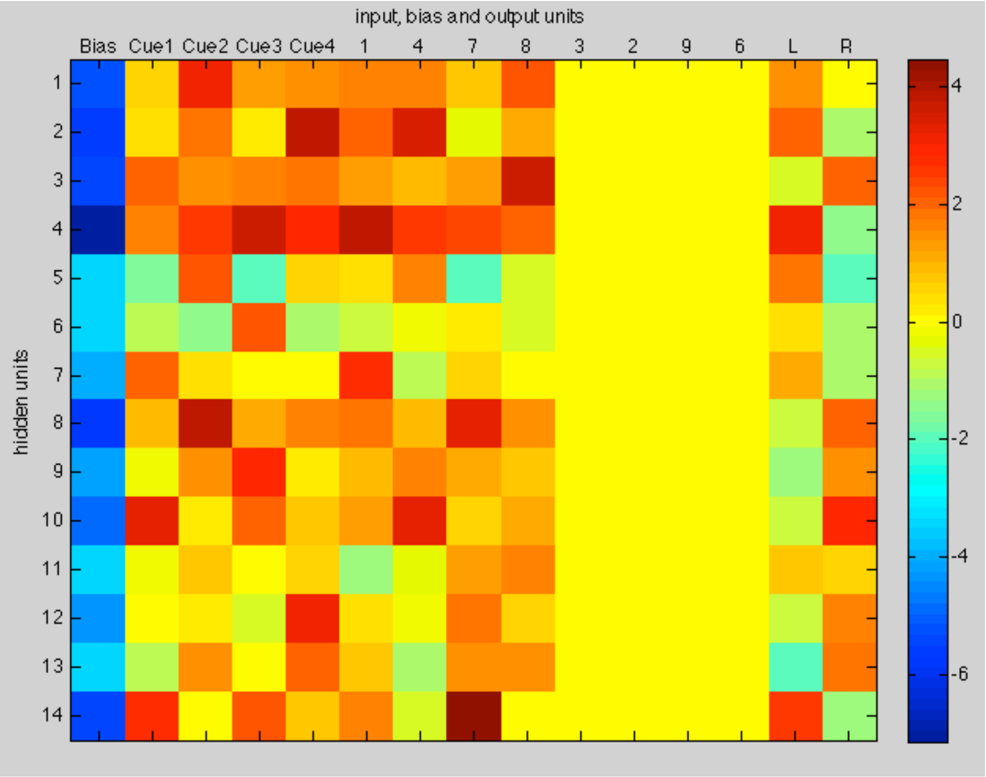


Figure 5.4 shows the strength of the final connections to the hidden units for an example model run in the Normal condition. The redder the colour, the more excitatory the link, whilst the bluer the colour, the more inhibitory the link. Where 1 always requires a left (L) response, 8 always requires a right (R) response and both of these numbers occur with all four cues. For 4 a L response is required for Cues 2 & 4, a R response is required for Cues 1 & 3, whilst for 7 a L response is required for Cues 1 & 3, whilst a R response is required for Cues 1 & 3. 3,2,6, & 9 are not used by this model but are used by other models for counterbalancing purposes.

An example of the final weights to hidden units in the Separate condition

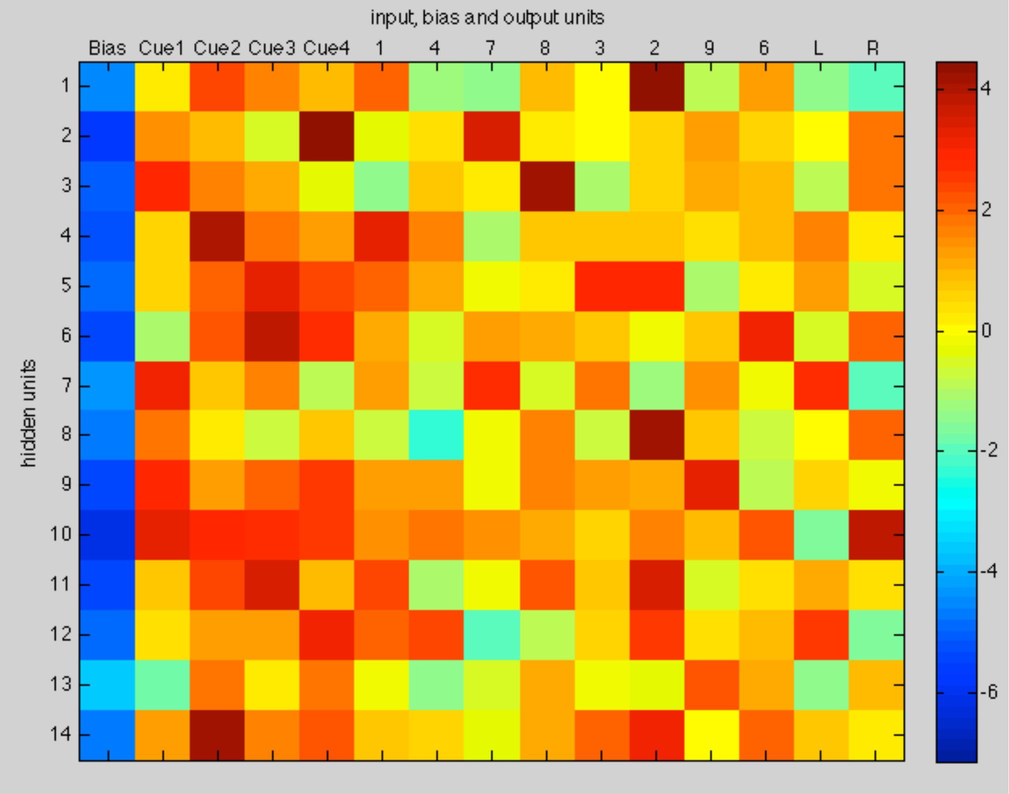


Figure 5.5 shows the strength of the connections for an example model run in the Separate condition. The redder the colour the more excitatory the link, whilst the bluer the colour the more inhibitory the link. Note that 1, 8, 3 & 6 only occur with Cues 2 & 3 and 4,7,2, &9 only occur with Cues 1 & 4. 1 & 3 always require a left (L) response and 8 & 6 always require a right (R) response. 4 & 2 require a L response with Cue 4 and a R response with Cue 1. 7 & 9 require a L response with Cue 1 and a R response with Cue 4.

It is interesting to consider how the model "solves" the Separate condition (see Figure 5.5). One can see that the mappings created for the Separate condition are different from those created in the Normal condition (Figure 5.4). In the Normal condition we found that for incongruent trials the stimulus had a stronger link to the chosen hidden unit than the cues (e.g. hidden unit 14 in Figure 5.4). However, in the Separate condition the mappings for incongruent trial now have stronger weights from the cue rather than the stimulus (e.g. hidden unit 2 in Figure 5.5). Why is there this structural difference? In the Separate condition, unlike in the Normal condition, there are now equivalent stimuli, rather than equivalent cues, i.e. C & G (4 & 2) are equivalent to each other. These equivalent stimuli, like equivalent cues, tend to use the same hidden unit to carry a mapping (e.g. hidden unit 12 in Figure 5.5). This means that, contrary to the Normal condition, the cue now uses a given hidden unit to carry its mapping more frequently, as when either of the equivalent stimuli are presented the cue is also presented, and so its link to the hidden unit gets strengthened. This leads to a slightly uneven distribution of influence between the cue and stimulus, which is opposite to that shown in the Normal condition.

Based on the argument outlined in the introduction, one might expect that the stronger cue to hidden unit connections may lead to a “switch benefit”, as now the trial type XC-L, XD-R is a more disadvantageous trial type than XC-L, YC-R. This is because presentation of XC will lead to more unlearning and a larger bias on trial type XD-R than YC-R. This difference will be considered in more depth in the discussion. However, an overall switch benefit was not observed in the modelling results. This is because there is still an advantage on “task repeat” trials even if direct CS repetitions are removed as, for the incongruent trials, if the equivalent incongruent stimuli occur sequentially within the same “task” then they use the same hidden unit to carry the mapping. As noted above, when the same hidden unit is used to carry a mapping there is an advantage because the link between the hidden unit and the output unit has been recently trained. So, although doubling the number of stimuli did equate overall performance as there were the same number of mappings, it also led to an associative structure that could generate effects relating to the variable “task switch”.

It is worth noting that in the Separate condition, congruent stimuli hijack the representations of the incongruent stimuli much more rarely (see Figure 5.5). As they no longer share cues, it is no longer the case that a hidden unit used to carry an incongruent stimulus mapping is a good candidate for carrying a congruent stimulus’s mapping: unlike in the Normal condition it cannot be activated by the congruent stimulus’ cue. However, one may look at Figure 5.5 and note that it is not strictly the case that congruent stimuli never hijack the same hidden unit as some units are clearly sensitive to both congruent and incongruent stimuli, e.g. hidden unit 4. This may be due to the relatively small number of hidden units in the network creating a need to share hidden units across dissimilar trial types. However, this would not be constrained as it is for the Normal condition in relation to the variable “task switch” as congruent and incongruent mappings from opposing “tasks” can now be shared by the same hidden unit, e.g. for hidden unit 4 it carries the mapping for 4 with Cue 4 and 1 with Cue 2&3, where only Cue 2&4 would be assigned as being from the same “task”.

5.3 Experiment 4

The model predicted that under Separate conditions there would not be a significant "switch cost" when direct cue-stimulus repetitions were excluded from the analysis of the data. An experiment was run in order to see if the structural manipulation would cause similar changes in the behaviour of human participants as it did in the simulation. As in Experiment 3, Experiment 4 used 4 distinct colours (red, blue, green and yellow) as cues and digits (1,4,7,8,2,3,6,9) as the stimuli.

5.3.1 Method

5.3.1.1. Participants

The participants were 32 psychology undergraduates (average age: 19.2 years, 10 male) at the University of Exeter, 16 assigned to each condition. Participants took part for either half a course credit and a bonus payment (24), which was contingent on their performance (average payment £1.63, range £1.00-£2.00) or for payment, part of which was contingent on their performance (average payment £3.66, range £3.25-£4.25)

5.3.1.2 Stimuli

The task cues and stimuli were as in Experiment 3. In other words the cue was a circle (6.7° of visual angle), filled with blue (RGB: 0, 0, 255), red (RGB: 255, 0, 0), green (RGB: 0, 255, 0) or yellow (RGB: 255, 255, 0); the digit stimulus was displayed in the center of the cue in size 60 Courier bold font (1.3° of visual angle). The two sets of digits used were 1,4,7,8 and 2,3,6,9 for the Normal condition and all these numbers for the Separate condition. The computer and screen were as in Experiments 1, 2, and 3.

5.3.1.3 Design

Trial sequencing was as outlined in the modelling section earlier in this chapter.

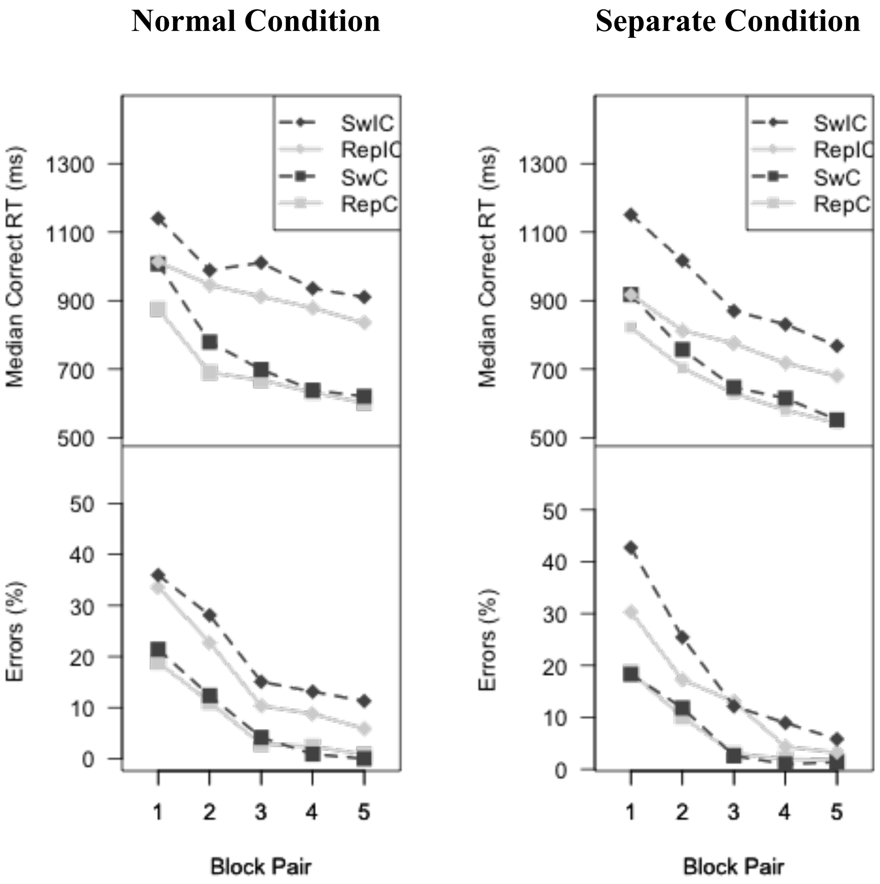
5.3.1.4 Procedure

Participants in both conditions were directed to learn cue+stimulus → response connections on the basis of trial-by-trial feedback, as in Experiment 2. The experiment was 10 blocks long and there was a CSI of 100ms in all blocks. As in previous experiments participants were interviewed and debriefed, however, in this case there were no exclusion criteria based on their questionnaire responses (see discussion for an explanation as to why this was the case).

5.3.2 Results

As with the modelling data, the data was first examined as a whole and then with direct cue-stimulus repetitions removed.

5.3.2.1 Results including direct cue-stimulus repetitions


Normal Condition

Separate Condition

Figure 5.6 shows the RT and error results of Experiment 4 including direct cue-stimulus repetition plotted as Figure 5.2

Figure 5.6 suggests the results of Experiment 4 are not in line with the predictions of the model. There is apparently a much larger switch cost in the Separate condition than the Normal condition. In order to analyse this effect a mixed ANOVA was run on the RT and errors with the factors, “task switch” (switch or repeat), congruency (congruent, incongruent), block pair (1,2,3,4,5) and condition (Normal or Separate). This ANOVA was followed up with ANOVAs on each condition where appropriate. As with the behavioural data in previous chapters, trials after an error were excluded from the RT and error analysis. Trials on which an error was made were also excluded from the RT analysis.

"Task switches" and condition Contrary to the predictions there was a smaller "switch cost" in the RTs in the Normal condition (68ms, 2.5%) than in the Separate condition (95ms, 2.6%), but this was not a significant difference in either the RTs or the errors. These costs were reliable in both RTs and errors of the Normal condition, RT: F(1,15)=35.2, $p<0.001$, Errors: F(1,15)=8.24, $p<0.05$, and of the Separate condition, RT: F(1,15)= 43.9, $p<0.001$, Errors: F(1,15)=9.73, $p<0.01$.

"Switch" by congruency and condition. The "switch cost" in the Normal condition was larger for incongruent trials (80ms, 4.4%) than the congruent trials (55ms, 0.6%). This difference was even more pronounced in the Separate condition where the "switch cost" for the incongruent trials (147ms, 5.4%) was much larger than for the congruent trials (42ms, -0.1%). There was a significant difference in this effect between the two conditions in the RTs, F(1,30)=4.9, $p<0.05$ but not in the errors, $F<1$. Separate analyses revealed that the "switch cost" was significantly larger in the incongruent trials than the congruent trials in the Separate condition, in the RTs, F(1,15)=29.7, $p<0.001$, and errors, F(1,15)= 14.5, p=0.092, but the effect was only significant in the errors for the Normal condition, F(1,15)=13.9, $p<0.01$ and not in the RTs, $F<1$.

Congruency and condition There was a numerically larger effect of congruency in the Normal condition (236ms, 11%) than the Separate condition (117ms, 9.2%), but this was not a significant difference between the conditions in either the RTs or errors $F<2$. The congruency effect was reliable in both the Normal, RT: F(1,15)=27.7, $p<0.001$, errors: F(1,15)= 43.4, $p<0.001$, and Separate conditions, RT: F(1,15)=112.7, $p<0.001$, errors: F(1,15)=50.3, $p<0.001$.

Acquisition effects and condition. There were no significant differences in the rate of acquisition between the groups, see Figure 5.6. However, there was a significant improvement from block pairs 1-5 in both the Normal condition, RT: F(4,60)=21.3, $p<0.001$, Errors: F(4,60)= 53.8, $p<0.001$ and the Separate condition, RT: F(4,60)= 18.3, $p<0.001$, Errors: F(4,60)= 39.5, $p<0.001$. Also there was some evidence of a significant change in the size of the congruency effect from block pairs 1-5 in RTs in the Normal condition, RT: F(4,60)=2.98, p=0.052, but not in the errors, $F<1$. In the Separate condition there was evidence of this effect in the errors: F(4, 60)= 11.0, $p<0.001$, but not in the RTs, $F<1$.

The results with direct cue-stimulus repetitions included are contrary to the results predicted by the modelling presented at the beginning of this chapter. However, the main prediction was about what happened when direct cue-stimulus repetitions were excluded from the analysis. Hence the analysis was repeated, but with direct cue-stimulus repetitions removed, see Figure 5.7.

5.3.2.2 Results excluding direct cue-stimulus repetitions

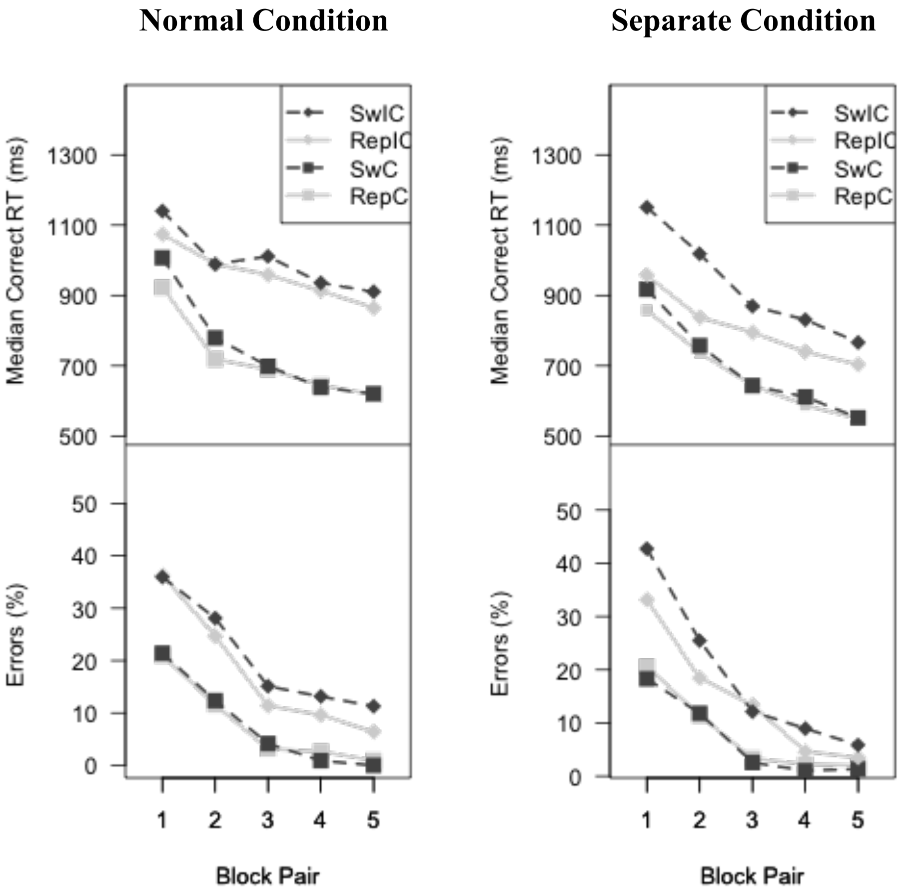


Figure 5.7 shows the RT and error results of Experiment 4 when direct cue-stimulus repetitions are excluded as in Figure 5.2.

"Task switches" and condition. In complete contrast to the predictions of the model there was a smaller "switch cost" in the RTs in the Normal condition (34ms, 1.4%) than in the Separate condition (71ms, 1.7%). The difference in cost between the conditions was significant in the RTs, F(1,30)=5.13, $p<0.05$ but not the errors. Separately analysed, the costs were reliable for the RTs in both the Normal condition, RT: F(1,15)=12.1, $p<0.001$, and the Separate condition, RT: F(1,15)= 28.4, $p<0.001$, but not for the errors in either condition, Normal, F<1, Separate F(1,15)=3.79, p=0.071.

"Switch" by congruency and condition. The "switch cost" in the Normal condition was larger for incongruent trials (37ms, 3%) than the congruent trials (30ms, 0.0%). This difference was even more pronounced in the Separate condition where the

“switch cost” for the incongruent trials (121ms, 4.4%) was larger than the congruent trials (21ms, -0.9%). The 3-way interaction showed a significant difference in this effect between the two conditions in the RTs, F(1,30)=7.56, p<0.05 but not in the errors, F<1. Separate analyses revealed that the “switch cost” was significantly larger for the incongruent stimuli than the congruent stimuli in the Separate condition, in the RTs, F(1,15)=25.6, p<0.001, and errors, F(1,15)= 13.0, p<0.01, but the effect was significant only in the errors for the Normal condition, F(1,15)=8.8, p<0.05, not in the RTs, F<1.

Congruency and condition There was a numerically larger effect of congruency in the Normal condition (245ms, 11.4%) than the Separate condition (181ms, 9.3%), but this was not significant in either the RTs or errors, F<2. The congruency effect was reliable in both the Normal, RT: F(1,15)=27.3, p<0.001, errors: F(1,15)= 45.4, p<0.001, and Separate conditions, RT: F(1,15)=115, p<0.001, errors: F(1,15)=51.2, p<0.001.

Acquisition effects and condition. There were no significant differences in the rate of acquisition between the groups, see Figure 5.7. However, a significant improvement occurred in block pairs 1-5 in both the Normal condition, RT: F(4,60)=23.6, p<0.001, Errors: F(4,60)= 57.8, p<0.001 and the Separate condition, RT: F(4,60)= 19.1, p<0.001, Errors: F(4,60)= 39.9, p<0.001. Also there was some evidence of a significant change in the size of the congruency effect from block pairs 1-5 in RTs in the Normal condition, RT: F(4,60)=2.90, p=0.05, but not in the errors, F<1. In the Separate condition there was some evidence of this effect in the errors: F(4,60)= 11.0, p<0.001, but not in the RTs, F<1.

5.3.3 Discussion

The results of the above experiment are for the most part in disagreement with the predictions I derived from the model. Participants in the Separate group showed a larger switch cost in their incongruent trials than those in the Normal group whilst the model predicted it should be smaller. The model also predicted that the switch cost in the Separate group should be solely caused by direct cue-stimulus repetitions. However, the switch cost in the Separate group comfortably survived removal of these direct cue-stimulus repetitions from the data.

Why such a large disparity between prediction and experiment? The answer appears to lie in the answers to the questionnaires that the participants in the Separate condition gave. 14 out of the 16 participants in the Separate condition claimed to approach the experiment in the same way. They induced one task rule[11], either high/low or odd/even, and realised that this rule applied to all cue values except for one colour for which the responses reversed (see Table 5.3 for an example). This approach to the experiment leads to a large switch cost as the participant is now changing from one rule to the exception – which in some ways is equivalent to changing between two "task-sets". Even though the need to change "task-set" now does not directly correspond to the defined task switch variable, more of these cases occur on switch trials than repeat trials. When an exception cue is presented on a switch trial, then the "task-set" must have changed. However, if an exception cue is presented on a "task-repeat" trial, then, although a proportion (just over half) of the time this is a change in "task-set", in some cases it has just followed an exception cue trial at which point the "task-set" has repeated. This difference in the proportion of changing to an exception on task repeat and task switch trials would lead to the large switch cost in the Separate condition.

In previous experiments some participants induced one or two of the task rules, but the majority of participants have not. Here, in the Separate condition, nearly all of the participants induced rules. Why the difference in proportion? Two factors could be involved. Firstly, participants have more information from which to extract the task

---

[11] It was for this reason that it made no sense to exclude participants based on rule induction as the majority of participants in one condition induced a rule.

rules, as they now experience 8 stimuli concurrently rather than 4. This would make the categories more obvious as they now have more examples of the stimuli from which to extract the common underlying factor. This equivalence between stimuli may be important in allowing participants to extract the underlying structure that is governing which response is correct.

		“Task 1” Cues		“Task 2” Cues	
		Blue	Green	Red	Yellow
	1	L	-	L	-
Stimuli	8	R	-	R	-
	4	-	L	-	**R**
	7	-	R	-	**L**
	3	L	-	L	-
	6	R	-	R	-
	2	-	L	-	**R**
	9	-	R	-	**L**

Table 5.3 shows a concrete example of the cues and stimuli used in Experiment 6. This shows that if participants learnt the rule lower than 5 is always left and higher than 5 is always right except for when the colour is yellow (the bold responses) and then the responses reverse, they would have an efficient and effective way of knowing how to respond in the experiment.

Secondly, it may be the case that inducing a rule and an exception is easier than inducing two rules. Although participants have sometimes induced one rule, then noted an exception with the original task-cueing structure (Table 5.1), in that case the exception rule is more complex as it contains two conditions not just one, i.e. the exception is if an incongruent stimulus appears on one of two colours. Hence more participants in the Separate condition may have induced this rule structure as it is easier to induce and simpler to use than the equivalent structures in the Normal condition.

Although this difference in the proportion of rule induction does not speak directly to the issues contained within this thesis, it does not disagree with the notion that there is an automatic associative system that gathers information and also a controlled cognitive system that uses rules. It also suggests that when one approach is notably more efficient that the other then this is likely to be the approach driving performance. Here in the Separate condition the rule and exception route is clearly a more efficient strategy for the majority of the stimuli, and so for the majority of participants this approach dominates.

Although this experiment did raise some interesting ideas and is incidentally supportive of the idea of participants being able to find and use rule-based strategies when these are effective, it was not a true test of the modelling at the beginning of this chapter. Thus, in Experiment 5, the same manipulation was run except that symbols were used rather than numbers so that there were no odd-even or high -low task-rules for the participants to induce. If the processes we have been examining throughout the thesis are associative in nature, then changing the numbers to symbols should have very little effect on the pattern of performance. But it will ensure that participants cannot induce rules and so have to rely on CSR associations, which I hypothesise are learned via a low-level associative system. It is this system that APECS was intended to model and not the rule based behaviour exhibited in Experiment 4.

5.4 Experiment 5

5.4.1 Method

5.4.1.1 Participants

The participants were 55 psychology undergraduates (Average age: 20.9 years, 8 male) at the University of Exeter. Participants took part for either half a course credit and a bonus payment (n=47), which was contingent on their performance (average payment £1.48, range £1.00-£2.25) or simply for payment (n=8, average payment £3.59, range £3.50-£3.75). 7 participants were excluded for either failure to obey the basic experimental instructions (5 participants), e.g. failing to press a key for the entire first block, or for being more than 3 standard deviations away from the overall average RT or error score (2 participants).

5.4.1.2. Stimuli Design and Procedure

This was identical to Experiment 4 except that the symbols ©,&,%,µ,fl,@,$,± were used in place of the numbers 1,2,3,4,6,7,8,9 respectively and there was now no CSI, so the cue and the stimulus appeared simultaneously.

5.4.2 Results

5.4.2.1 Results including direct cue-stimulus repetitions

Inspection of Figure 5.8 and Figure 5.9 shows that the results of this experiment now bear more similarity to the results produced by the modelling described at the beginning of this chapter. They both show a notably larger switch cost in the Normal condition as compared to the Separate condition.

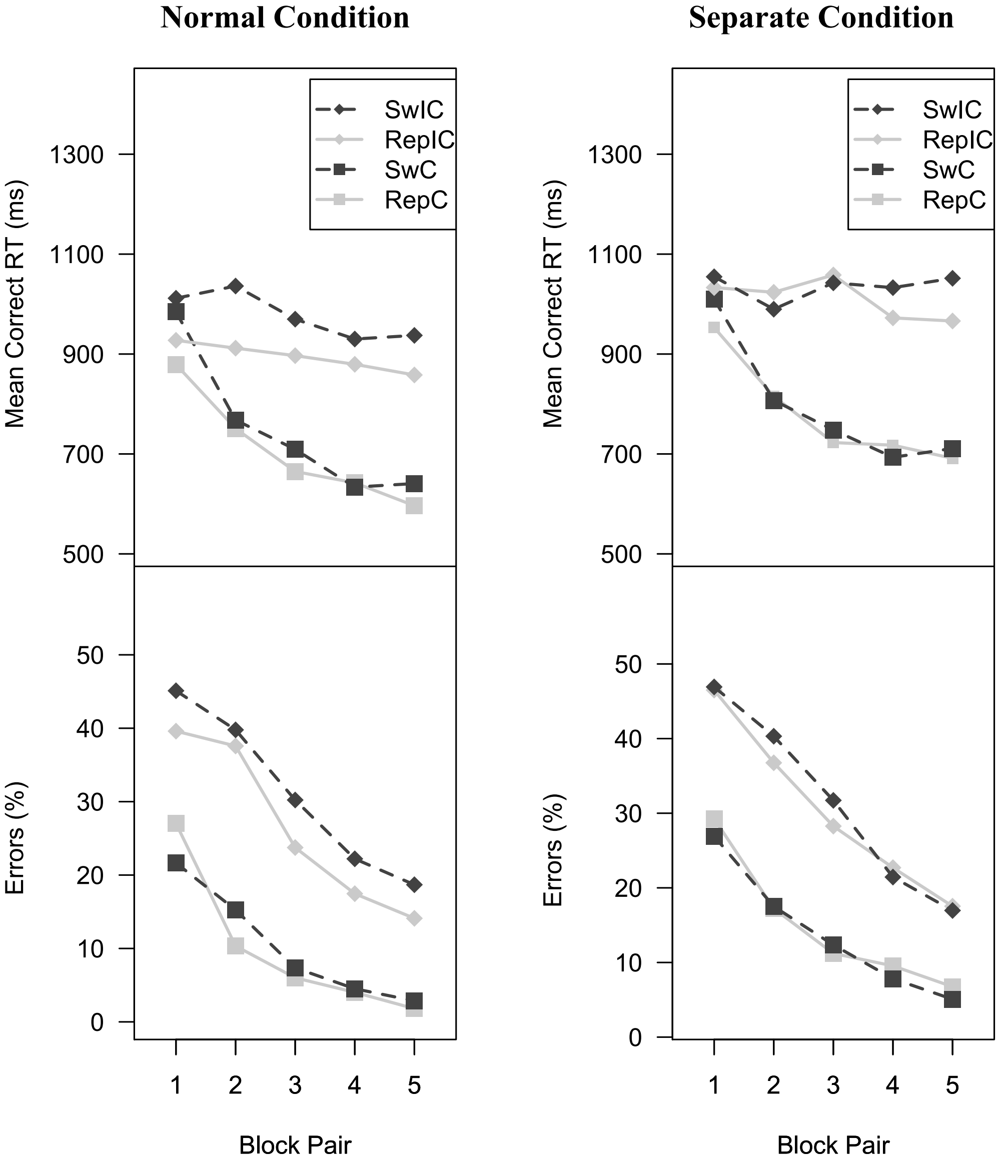


Figure 5.8 shows the RT and error results of Experiment 5 plotted as in Figure 5.2.

In order to examine if this was a significant effect, the data were first examined with no exclusion of direct cue-stimulus repetitions. A mixed ANOVA was run on the RT and errors with the factors, “task” switch (switch or repeat), congruency (congruent, incongruent), block pair (1,2,3,4,5) and condition (Normal or Separate). This ANOVA was followed up with ANOVAs on each condition where appropriate. Trials after an error were excluded from the RT and error analysis, and trials on which an error was made were also excluded from the RT analysis.

"Task switches" and condition. There was a significantly larger "switch cost" in the RTs in the Normal condition (58ms, 2.2%) than in the Separate condition (26ms, 0.8%), F(1, 46)=4.76, $p<0.05$, but this did not reach significance in the errors, F(1,46)=1.27, n.s. These costs were reliable in both RTs and errors of the Normal condition, RT: F(1,23)=23.6, $p<0.001$, errors: F(1,23)=5.18, $p<0.05$, but only in the RTs of the Separate condition, RT: F(1,23)= 9.44, $p<0.01$, errors: F(1,23)=0.71, n.s.

"Switch" by congruency and condition. The switch cost in the Normal condition was larger for incongruent trials (77ms, 4.5%) than for the congruent trials (39ms, -0.1%). Similarly, in the Separate condition the switch cost for the incongruent trials (31ms, 1.6%) was larger than for the congruent trials (21ms, -0.1%). However, the difference between conditions was not significant in either the RTs or the errors. Separate analyses revealed that the "switch cost" was significantly larger in the incongruent trials for the errors in the Normal condition, F(1,23)=5.22, $p<0.05$, and nearly so for the RTs, F(1,23)= 3.10, $p=0.092$, but not in either the RTs or errors in the Separate condition, $F<2$.

The "switch cost" was significant in the Normal condition in the RTs in both the incongruent, F(1,23)=15.1, $p<0.001$, and congruent trials, F(1,23)=11.6, $p<0.01$. The "switch cost" was also significant in the errors in the incongruent, F(1,23)=7.17, $p<0.05$ but not the congruent trials, $F<1$. In the RTs in the Separate condition there were near significant "switch costs" in both the incongruent, F(1,23)=4.09, $p=0.054$, and congruent, F(1,23)=3.40, $p=0.078$, trials. There were no significant switch costs in the errors in the Separate condition in either the incongruent or congruent trials, $F<2$.

Congruency and condition There was a numerically larger effect of congruency in the Normal condition (238ms, 18.1%) than the Separate condition (205ms, 14.6%), but this was not significant in either the RTs or errors $F<2$. The congruency effect was reliable in both the Normal, RT: F(1,23)=111, $p<0.001$, errors: F(1,23)= 60.0, $p<0.001$, and Separate conditions, RT: F(1,23)=78.3, $p<0.001$, errors: F(1,23)=96.3, $p<0.001$.

Acquisition effects and condition. There were no significant differences in the rate of acquisition between the groups, see Figure 5.8. However, a significant improvement across block pairs 1-5 was observed in both the Normal condition, RT: F(4,92)=12.8, $p<0.001$, Errors: F(4,92)= 59.0, $p<0.001$ and the Separate condition, RT:

F(4,92)= 8.76, p<0.001, Errors: F(4,92)= 47.7, p<0.001. There was also a significant change in the size of the congruency effect across block pairs 1-5 in both the Normal condition, RT: F(4,92)=8.92, p<0.001, Errors: F(4, 92)= 4.65, p<0.01 and the Separate condition, RT: F(4, 92)= 17.1, p<0.001, Errors: F(4, 92)= 2.72, p<0.05.

5.4.2.2 Results excluding direct cue-stimulus repetitions

The same analysis was repeated but with direct cue-stimulus repetitions removed.

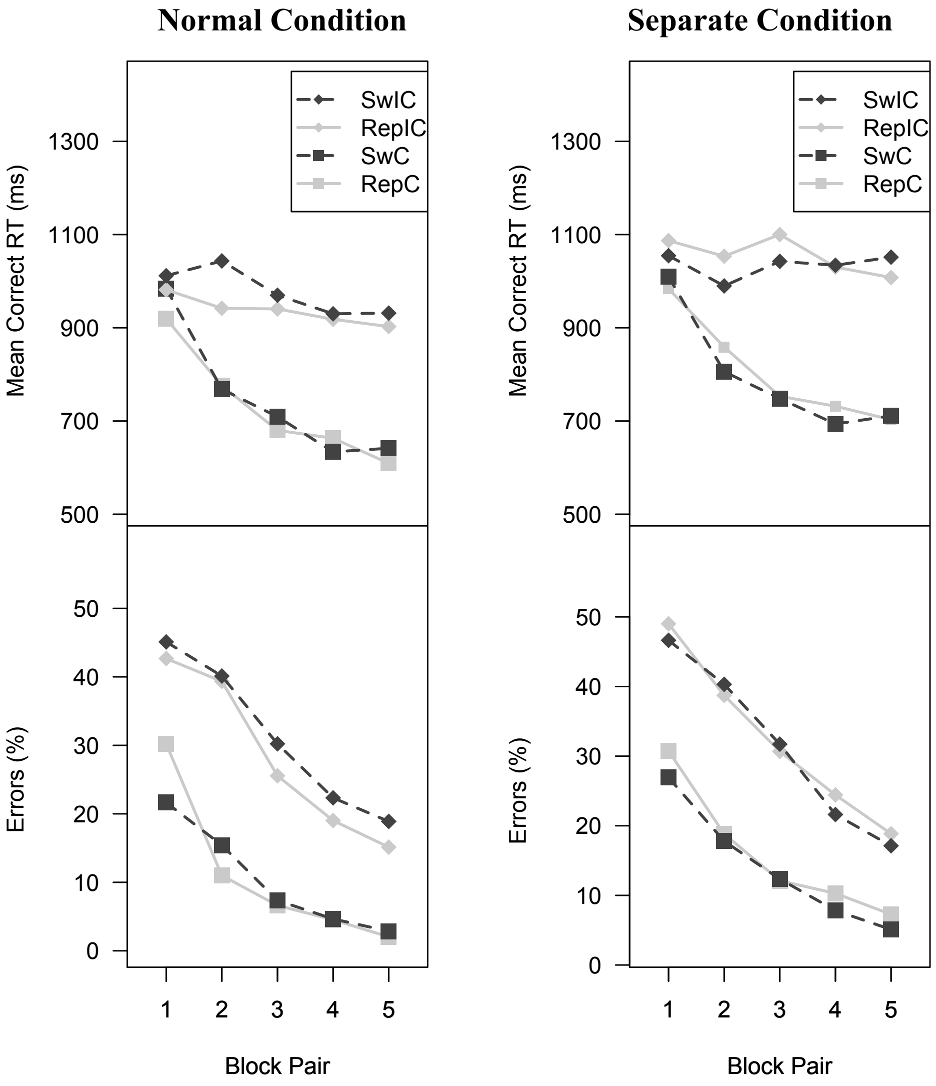


Figure 5.9 shows the results of Experiment 5 without direct repetitions, plotted as in Figure 5.2.

"Task switches" and condition. The "switch cost" in the data for the Separate condition (-9ms, -0.9%) was now no greater than zero, and it was significantly smaller for RT than in the Normal condition (25ms, 0.8%), F(1,46)=5.45, p<0.05, but this did not reach significance in the errors, F(1,46)=1.67, n.s. These costs were reliable only in the RTs of the Normal condition, RT: F(1,23)=4.59, p<0.05, but not in the errors,

F(1,23)=1.04, n.s. There was also no effect of "switch" in the RTs or errors of the Separate condition, F<2.

"Switch" by congruency and condition. The "switch cost" in the Normal condition was larger for incongruent trials (29ms, 2.7%) than for congruent trials (21ms, -1%). In the Separate condition the "switch costs" were negative in both the incongruent trials (-12ms, -0.4%) and the congruent trials (-5ms, -1.5%). The difference between the two conditions was not significant in either the RTs or the errors, F<2. The only evidence of a switch by congruent interaction was in the errors in the Normal condition, which showed a nearly significant interaction, F(1,23)=3.54, p=0.073. All other switch by congruency interactions were non-significant, Fs<1.

Congruency and condition There was a numerically larger effect of congruency in the Normal condition (253ms, 18.6%) than the Separate condition (214ms, 15%), but this was not significant in either the RTs or errors F<2. The congruency effect was reliable in both the Normal, RT: F(1,23)=115, p<0.001, errors: F(1,23)= 91.6, p<0.001, and Separate conditions, RT: F(1,23)=80.7, p<0.001, errors: F(1,23)=59.2, p<0.001.

Acquisition effects and condition. There were no significant differences in the rate of acquisition between the groups, see Figure 5.9. However, a significant improvement across block pairs 1-5 occurred in both the Normal condition, RT: F(4,92)=12.8, p<0.001, Errors: F(4,92)= 48.4, p<0.001 and the Separate condition, RT: F(4,92)= 7.75, p<0.001, Errors: F(4,92)= 59.2, p<0.001. There was also a significant change in the size of the congruency effect across block pairs 1-5 in the RTs in the Normal condition, F(4,92)=8.31, p<0.001, and this was nearly the case in the errors, F(4, 92)= 2.56, p=0.058. This interaction was significant in both the RTs and errors in the Separate condition, RT: F(4, 92)= 15.3, p<0.001, Errors: F(4, 92)= 4.56, p<0.01.

5.4.3 Discussion

The prediction made by the model about how the change in the associative structure of the task-cueing paradigm would affect the switch cost was confirmed by Experiment 5, in which precautions were taken to restrict the participant to a purely CSR strategy. There was a significantly smaller switch cost in the Separate condition than in the Normal condition, and this small switch cost was eliminated when direct cue-stimulus repetitions were removed. This suggests that when participants are approaching the experiment by learning CSR mappings then their behaviour is in agreement with the results of APECS. It also suggests that the switch cost observed in the Monsell

Experiment and Experiments 2 & 3 (and the nearly reliable switch cost in Experiment 1) may be due to the particular design of the task-cueing paradigm used – specifically the use of two cues per "task" and the sharing of cues between incongruent and congruent stimuli.

The main disagreement between the model's predictions and this dataset is the distribution of the switch cost in the Normal condition when direct cue-stimulus repetitions are included in the data. The model predicted a switch cost only for the incongruent trials. However, in the behavioural data in the Normal condition when direct cue-stimulus repetitions were included, there was evidence of a switch cost (39ms, -0.1%) in the congruent trials in the RTs. This was also nearly the case in the Separate condition when direct cue-stimulus repetitions (21ms, -0.1%) were included for the RTs but not for the errors. In some ways it is strange that the model does not produce a switch cost for the congruent trials when direct cue-stimulus repetitions are included. However, for the most part the model tends to map the congruent trials simply by the stimulus rather than by the cue+stimulus combination, meaning that for the model any stimulus repetition would induce the same level of priming in the model, so there should be no advantage for direct cue-stimulus repetitions over stimulus repetitions that could lead to a benefit for congruent repetition trials. Indeed when these trials are removed there is a slight switch benefit for the congruent trials (which is also seen in the behavioural data). This difference in the congruent trials switch cost could simply be a floor effect in the model, as it learns these trial types much faster than the human participants. But it could also be due to something the model is not capturing, such as a priming effect in the speed of perceiving the cue or stimulus, as Logan and Bundesen (2003) suggest.

5.5 General Discussion

This chapter took a model-based approach to considering why there is a small yet reliable switch cost in the CSR conditions presented so far in this thesis. Based on a consideration of the nature of the representations that the model forms, I hypothesised that two aspects of the design of the task-cueing paradigm may be generating the switch cost, namely the use of two equivalent cues per "task" and the fact that these cues were shared between the congruent and incongruent stimuli. Modifying the task-cueing design to eliminate these aspects of the design, and removing direct cue-stimulus repetitions, resulted in an abolition of the switch cost in the simulations generated by the model.

The predictions of the model were assessed in two experiments. Experiment 4 used digits as the stimuli, as in all the previous experiments reported in this thesis. The results of this experiment were contrary to the predictions of the model, as they suggested that the new design (i.e. the Separate condition) of the experiment led to a larger switch cost than the previous design (i.e. the Normal condition). However, examining the questionnaire data from this study revealed that the majority of participants in the Separate condition were using a task plus exception based strategy. As this thesis has previously established that using task-sets results in larger switch costs, and the modelling was not meant to apply to participants using task-sets, this may well explain the disagreement between the results of the experiment and the model.

In response to this analysis Experiment 5 was run. The key modification from Experiment 4 was the use of symbols instead of numbers, a manipulation that prevented participants in this condition from being able to induce task-sets. The data here was in broadly in agreement with the predictions of the model. This suggests that under conditions that successfully induce a CSR strategy (rather than some sort of rule-based strategy), when the associative structure of the task-cueing paradigm is manipulated in such a way that there are no longer two equivalent cues per task, and the cues are not shared between congruent and incongruent stimuli, then there is little or no evidence of a switch cost for participants using CSR instructions.

Being able to make and validate a non-obvious prediction from a model clearly increases the credibility of that model. This illustrates one of the key advantages of taking a model based approach, the ability to make and test concrete predictions. But this chapter has also highlighted one of the potential pitfalls of this approach, as the initial experiment to test the model failed to produce results that matched up with the predictions because the experiment, in some sense, did not properly engage the processes that were thought to be under test. Hence when making model based predictions it is important to ensure that the performance of the participants in the experiment is being governed by the same processes as those which are being modelled, otherwise it could be the case that models are unnecessarily falsified by data that there were not intended to model.

Chapter 6 Real-time associative modelling of the CSR condition

This chapter develops further the associative modelling applicable to the data presented in this thesis. So far I have used a static APECS model as a benchmark model of how an associative system could perform. In this chapter I first present an initial attempt to extend this model by creating a real-time version of an APECS model that is able to encode the temporal aspects of the task-cueing paradigm. This endeavour met with considerable success: after careful parameterisation it was able to account for the data from Experiment 2, but ultimately the model was found to then not be able to account for the data of the first half of Experiment 3 with the same set of parameters. In general this model was found to be too sensitive to small changes in its parameters and so lacks the necessary robustness to allow us to have confidence in its predictive powers, as the data suggest that it is unable to cope with even very small changes in the sequencing of the input it receives. It may be that this problem can be solved in the future, but for the moment my conclusion is that just creating a real-time version of the existing static model may not be the answer required to model my data.

A second line of enquiry described in this chapter is to consider an alternative associative model. The static APECS model considered so far is an example of a configural associative model. However, in the associative learning literature there is another class of associative model that takes an elemental approach. My chosen example of an elemental model is AMAN: the Attention Modulated Associative Network (Harris & Livesey, 2010, Thorwat, Livesey & Harris, 2012). This model was chosen as it has previously been shown to be able to learn both biconditional and component discriminations (Harris & Livesey, 2010), is real-time and was recently created in a freely accessible form (Thorwat, Livesey & Harris, 2012). In general I have found that AMAN is able to account for many aspects of the data from the CSR condition: the congruency effect, the change in the congruency effect over time and when used to model Experiment 2, the switch cost. Additionally, unlike the real-time APECS version I will present here, it showed a main effect of CSI, whereby it performed better on trials with a long CSI than a short CSI, which was the case in Experiment 2. However, there were aspects of the behavioural data that the model was unable to account for, notably the results of Experiment 5. My conclusion will be that no truly satisfactory real-time model of the associative processes involved in task-cueing paradigms run under CSR instructions is available, but that these preliminary attempts do indicate that there are grounds for optimism that such a model can be developed.

6.1 Real-time APECS

One of the main drawbacks of the version of APECS used up until this point in the thesis is that it is unable to encode temporal aspects of the experiment. Although the model did use more than one cycle through the network for each trial, this was not meant to represent the passing of time, rather to acknowledge that it may not be unreasonable to allow the model to adapt more than once per trial. If time were truly encoded by the model, the activations of the units would evolve dynamically. That is, activations would grow and decay over time, determined not only by the input they receive but also by their own previous activation. This gives the units a continuous activation function of time rather than the step function of time that occurred in the static model. Although not the only approach to creating a real-time model, this principle is used in other models, such as the task switching models of Gilbert & Shallice (2002) and Brown et al (2007) outlined in Chapter 1 and the AMAN model presented below (Harris & Livesey, 2010; Thorwat, Livesey, Harris, 2012).

Creating a real-time version of the model enables it to encode exact timings of the experimental trials, e.g. the amount of time that the stimulus and cue are present on the screen and the cue-stimulus interval. As the amount of time the cue and stimulus were presented to the participant was determined by their reaction time, the real-time model also needs to be able to decide when to "make a response". The static model had no way of managing this, as it only generated output levels and not responses. In order to generate a way of dealing with this problem, the model outlined below incorporates a decision mechanism based on the one developed by Wills and McLaren (1997) and Wills et al (2000). This involves having separate response units (see Figure 6.1) that compete until one of them gets over a threshold – at which point a response is made. This mechanism allows both RTs and error rate to be generated; as now there is a measure of the time it takes the model to reach a response, and also then a measure of whether it was the correct response that was decided upon.

A third critique of the static APECS model is that it is not a realistic demonstration of learning from feedback. When the output is taken from the model, the model has just received 200 cycles training it to map that outcome to the desired response. However, this is not what the participants experience. Participants receive no feedback, except from their own intuitions whilst they are making a response, it is only after they have made a response that they receive feedback telling them if they were correct. It is from this feedback that they have to learn how to respond correctly the next time. This means that participants are not learning the response when the stimulus and cue are still present but instead are learning about some lasting representation of the stimulus and cue. Making the model real-time allows the variation of the target outputs within a trial to vary in such a way that they are a better reflection of the participants' experience.

Below I give a brief description of a real-time version of the APECS model. Although not entirely unsuccessful, this attempt at a real-time APECS model does not represent a perfect solution to the problem. The difficulties with this approach will be discussed afterwards.

6.1.2 The structure of a real-time APECS model

This model is a real-time network based on the APECS (McLaren, 1994, Le Pelley and McLaren, 2003, McLaren, 2011) class of algorithms. The APECS algorithm used was subtly different from the static APECS model described in Chapter 3. The previous version of APECS chose a hidden unit to carry the mapping by giving the highest learning rate to the hidden unit with the overall highest error score - this part of the algorithm was kept the same. Thus the way in which the weights develop between the input layer and the hidden layer and the hidden layer and the output layer are governed by the same rules in the static version mentioned earlier and the real-time version described below. However, in the previous version a high learning rate for the bias was given to the units with very negative error scores, as well as very positive error scores where there was large contribution from one output unit. The real-time APECS algorithm differed in that now the higher learning rate for the bias was simply applied if the error score was negative. This leads to subtle differences in the way the learning of a particular mapping is preserved, however, the overall rationale of the model stays the same. The other difference was that, for simplicity, the input units were no longer replicated on the output layer[12].

---

[12] These differences occurred due to a matter of history, as attempts to create the real-time version of APECS pre-dated the development of the most recent static version of APECS used in the rest of this thesis. This modelling effort was done concurrently with the running of Experiments 1 &2.

This algorithm was then made real-time. Activations within the network were set to rise gradually and then decay exponentially over a number of cycles through the network. Thus, the activation of a unit was determined both by its input and its previous activation. For the input layer the following formula was used to update the activations.

$$Ai_t = \frac{Ai_{t-1} + k \times V}{1 + k \times V + d}$$

where Ai_t is the activation of a unit in the input layer on cycle t, k is a constant (1) that determines how fast unit activation reaches its asymptotic value, d is a constant (0.016) that determines how fast a unit’s activation decays and V determines a unit's asymptotic value. To model the external input to the model: V was given a value of 1 if the unit was receiving external input and 0 if it was receiving no external input. Values in parentheses are those used in the simulations reported here.

In the case of the hidden and output layers the following formula was used as the input to the units was itself variable.

$$A_t = \frac{A_{t-1} + m}{1 + m + m \times \mathrm{e}^{Input_t}}$$

where *m* is a different constant (0.01) that determines how fast these units reach asymptote and Input is the input to the unit from the units connected to it.

As a result of these changes, a trial now extended over many cycles of the network, which meant that RT could be modelled by the number of cycles through the network, with one cycle corresponding to 0.75ms.

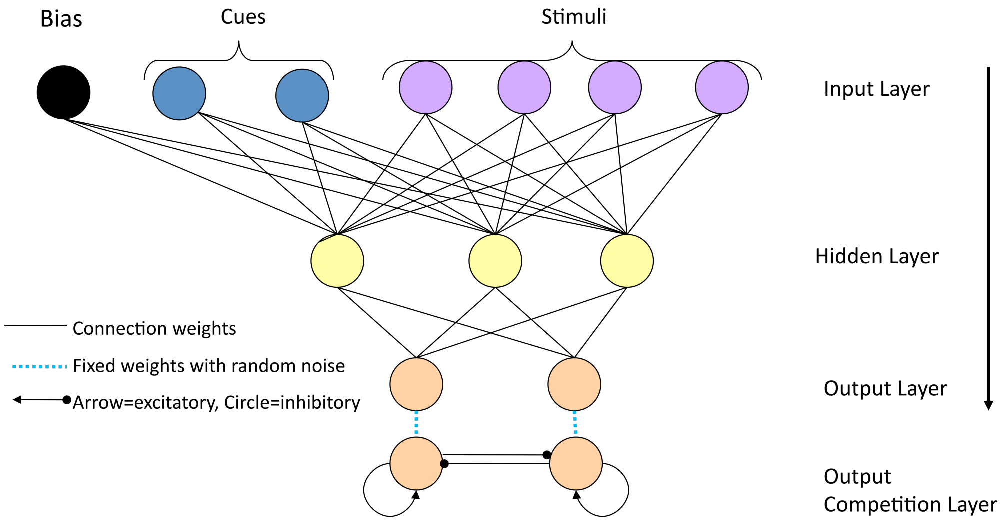


Figure 6.1 shows the architecture of the real-time APECS model.

Decision mechanism As mentioned earlier, this model also had a performance layer based on Wills and McLaren (1997) and Wills et al (2000). It consisted of two units, which took a direct copy of the activations in the output layer on each cycle and then competed with each other in a self-excitatory but mutually inhibitory manner, see Figure 6.1. This system means that the decision layer amplifies the differences seen on the output layer, so the activation difference between the two decision units becomes much greater than that between their corresponding output units. The network is judged to have made a decision when the activation of one of these units passes a threshold. In order to mimic participants' attempts to respond within the usual time restrictions, the threshold criterion that was used to judge when a decision was made was set to drift downward during the decision interval so that criterion for a response was lower towards the end of the 4 seconds allowed, effectively forcing a response in all but the most intractable cases.

6.1.3 Real-time APECS model of Experiment 2

The model was used to simulate the CSR condition of Experiment 2. The experiments were modelled using the same randomization procedures used to generate trial sequences for the participants. The stimuli were coded discretely using one input unit for each. The cues in Experiment 2 were modelled with a different unit representing each of the four shapes. One output unit represented a right response, whilst the other represented a left response.

Over the course of a trial the network went through four stages. When the task-cue was presented, input for the cue was gradually turned on. During this CSI the network was learning not to respond i.e. the target for both outputs was to be off, as participants have to learn to wait for the stimulus to appear before they can respond and thus the network is in some sense learning to not respond to the cue on its own.

In the next phase of the trial the network was also presented with the stimulus. The cue input was maintained. This phase of the experiment continued until the network had made a decision, i.e. one of the decision units had reached the threshold activation level or the trial timed out. During this time the network was trained with its target outputs set to the activation of the output layer, i.e. the network is allowed to follow its own hunches about what the correct response is. This was an innovation developed as a way round the problem of the network not being able to know the actual target outcomes at this stage and the desire to not want to artificially "turn off" learning during the stimulus-response interval. Learning to its own output layer means that if the network "knows" the correct response for a trial then this output will be facilitated. The number of cycles (i.e. the length of time) for which this continues is used to model reaction time, whilst whether the appropriate response is the one to surpass the threshold was used to model accuracy.

In the third phase, the network had just responded and receives feedback about what the correct response was. During this time the network was trained towards the correct target outputs (which were now known), and as this happened the activations of both the cue and the stimulus units decayed.

Finally when the activations in the input layer had decayed to such a level that the network had effectively no activations left from the previous trial, the network then received no input and was trained to not respond for the remainder of the inter-trial interval. The inter-trial interval was also made longer if the network had produced the incorrect response, in line with the extra feedback duration participants experienced on error trials. When the inter-trial interval ended the start of the next trial began.

The model was run 32 times on a complete sequence of trials of the same kind as given to participants in the experiment, each time with a different trial sequence and randomized initial starting weights of the network. The model included 12 input units, 10 hidden units and 2 output units. The RT data for the model is given in ms corresponding to cycles x 0.75. Note that I have not attempted a quantitative fit to the data by varying parameters to minimize any discrepancy between model and experiment. Instead I have focused on the patterns in the data produced by the model and how they compare to those in our empirical results.

In this model the learning rates for the weights between the input layer and the hidden layer were different from those between the hidden layer and output layer. The fast learning rate between the input layer and the hidden layer for chosen hidden units was 0.3, whilst the slow rate was 0.03. For the connections between the hidden layer and the output layer the fast rate for chosen hidden units was 0.32 whilst the slow rate was 0.001. This was done as it was found that this was needed for the model to be able to provide an account of Experiment 2. For the bias the selected hidden units had a learning rate of -150 times their error score and for unselected units the learning rate was 0.032. The learning rate for the selected hidden units bias was calculated like this to avoid the biases becoming too negative. These parameters were selected as they were found to produce viable data and they were kept constant for all the modelling results outlined below.

6.1.4 Real-time APECS simulation results for Experiment 2

Figure 6.2 shows the results of the real-time APECS model of Experiment 2, which shows a similar pattern of data to that found in Experiment 2. There performance of the model clearly improves from block pairs 1 to 5. There is also clear evidence of a large congruency effect, which changes over block pairs. There is also some evidence of a switch cost in the incongruent trials, at least in the errors. There is also clear evidence of a perturbation in performance when the new stimuli are introduced in block pair 6.

The same analysis was run on the real-time APECS model of Experiment 2 as was run on the behavioural data of Experiment 2 without the condition factor. To be more specific, a repeated measures ANOVA was run on the simulated median correct RT and percentage error rates, with the factors, task switch (switch, repeat), congruency, (congruent, incongruent), CSI (long, short) and block pair (2,3,4,5). As with the behavioural data trials after an error were excluded from the RT and error analysis and trials where an error was made were excluded from the RT analysis.

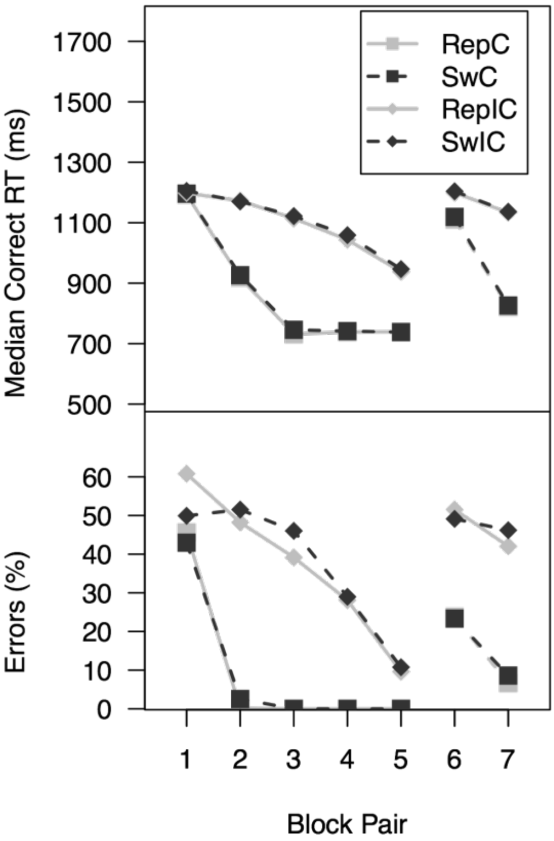


Figure 6.2 shows the results from the real-time APECS simulation of Experiment 2. Dark dotted lines represent switch trials, whilst light solid lines represent repeats. Squares represent congruent stimuli whilst diamonds represent incongruent stimuli.

6.1.4.1 Block Pairs 2-5

Congruency. Cycles and error rates showed a large congruency effect (286 ms, 32.5%), this was reliable for the cycles, F(1,31)= 103, p<0.001 and the error rates, F(1,31)= 782, p<0.001. This effect is consistent with the large and highly significant congruency effect found in the behavioural data.

Task switches. Also in agreement with the behavioural data for the CSR condition in Experiment 2, switch trials (931 ms, 17.5%) took slightly longer and contained more errors than repeat trials (925 ms, 15.7%). Although this effect is numerically much smaller than in the behavioural data (only 6ms and 1.8% errors) it is reliable for both RT F(1,31) = 17.8, p<0.001 and errors, F(1,31)= 6.63, p<0.05. The switch cost was larger in the incongruent stimuli (7ms, 3.5%) than the congruent stimuli (6ms, 0.2%) but only significantly so in the error data, F(1,31)=15.1, p<0.001.

Preparation. As with the behavioural data preparation did not reduce the switch cost, long CSI switch (7ms, 2.16%); short CSI (6ms, 1.48%), F<1. The model was overall faster in a short CSI (924ms) than a long CSI (932ms), but it made more errors with a short CSI (16.89%) than a long CSI (16.79%). Neither of these differences was reliable.

Acquisition effects. Overall performance evidently improved with block pair, this was reliable for both cycles F(3,93)= 102, p<0.001 and the errors, F(3,93)= 161, p<0.001.

The size of the congruency effect changed over time, for both simulated RT F(3,93) = 38.1, p<0.001, and error rate F(3,93) = 144, p<0.001. As for the behavioural data, the congruent stimuli were learnt more quickly than the incongruent and appeared to reach asymptotic performance around the third block pair.

6.1.4.2 Transfer Analysis

The model showed a marked and persistent perturbation in block pair 6 when the new stimuli are introduced (Figure 6.2). The effect of transferring participants from one set of stimuli to another was tested using a mixed ANOVA with the following factors: task switch (repeat or switch), congruency, CSI, block pair (5 or 6).

Transfer. As expected, there was a main effect of transfer, in which the model was worse after the stimuli had been changed (1199ms, 35.7%) than before (760ms, 1.84%), in the cycles, F(1,31)= 112, p<0.001 and the errors, F(1,31) = 1375, p<0.001.

The congruency effect changed when the stimuli were changed, it was much smaller after the stimuli had been changed (10.4 ms) than before (58.4ms) in the cycles data, F(1,31)= 6.91, p<0.05; the opposite was true of the error rates, where the congruency effect was larger after transfer (10.5%) than before (3.67%), F(1,31)= 6.38, p<0.05.

The effect of CSI was amplified when the stimuli were changed, before transfer the model was 2.32 ms faster in a short CSI than a long CSI, whereas after transfer this increased to 49.5ms, F(1,31) = 46.9, p<0.001.

6.1.5 Discussion of real-time APECS model of Experiment 2

It is worth noting that making the model real-time in this instance does not seem to have affected the overall pattern generated by the model. The model still produces a large congruency effect which first increases, then decreases in the course of training, and a small, yet significant switch cost. It also correctly predicted that the switch by congruency interaction was only present in the errors and not in the RTs, as in the behavioural data for Experiments 2. The model also, like the results of Experiment 2, did not show a reduction in the switch cost with time to prepare.

However, making the APECS model real-time did not produce a general preparation effect. This is in contrast to the behavioural data as there was a general preparation effect in Experiment 2. This may be related to the constraint that during the CSI the model learns not to respond. This would mean that having the cue present for longer is of no additional benefit, and possibly of some cost. It could also be that the fore period effect is outside the scope of the model as the main effect of CSI may be due to phasic alertness (Niemi & Näätanen, 1981; Nobre, Correa & Coull, 2007) which is not included in the model.

There are some other issues with the real-time APECS model. The main one is to do with parameter specificity. Unlike the static APECS model, this model is very sensitive to changes in its parameters. Its real-time nature means that it only works when a very delicate balance has been achieved. The units need to be active enough when the model makes a response and then decay slowly enough for any learning to occur. The learning itself needs to occur gradually to avoid oscillations but quickly enough so that it can learn from the decaying representations of the cue and the stimulus.

This is further complicated by the bias needing to change rapidly enough to preserve learning, but not so much that it prevents any new learning occurring. The decay of the response criterion also needs to happen slowly enough to give the activations time to build up, but not so slowly that the model "times out" on every trial. This is related of course to the rate of activation of the units, as this determines how fast the response units grow in activation and pass threshold. This fragility as a result of sensitivity to the parameters may not be a problem in itself if there were one set of parameters which could explain everything (after all it could be that evolution has tuned the model to the correct set of parameters), however, in this particular case the sensitivity leads to issues when it comes to making predictions with the model.

One issue one might consider is what happens when the constraints of the sequence are changed from a having a task switch probability of a half, as they were in Experiment 2 to a task switch probability of a third as they were in Experiment 3. This change seems simple enough for a model to cope with, indeed the static APECS model presented earlier, had no problem solving both sequences with the same set of parameters. However, with the same set of parameters the real-time model discussed above now fails to show a “switch cost”, when all the evidence so far suggests that a switch cost still occurs with a third switch probability and its previous static version was able to show a “switch cost”. It seems reasonable to expect the model to be able to account for both sets of data with the same set of parameters.

In conclusion, the evidence within this chapter suggests that it may be possible to produce a real-time APECS model which produces measures of both RTs and errors, and is able when tuned to a given situation to show the same basic effects as present in human participants, i.e. the model of Experiment 2 shows a similar pattern of responding to those seen in the human participants. It is also worth noting that this model is able to learn based just on the type of feedback the participants had – instead of the rather artificial way in which static feed forward back propagation systems tend to learn by knowing the correct outcome for each trial at same time as being presented with the input. However, this model is clearly not robust enough to be able to correctly cope with reasonable variation in task parameters, even just in the sequences presented to the model.

6.2 An elemental comparison

Despite the difficulties encountered when attempting the transformation into a real-time model, APECS seems to be able to account for much of the CSR data found so far in this thesis. However, APECS takes a theoretical stance about how an associative model would represent stimuli from the external world. It is worth considering if alternative assumptions could explain the results, whether these would lend themselves more easily to a successful real-time implementation, and whether or not this will inform the debate in associative learning about how stimuli are represented. In Chapter 1 I briefly mentioned that, although a configural model seems the natural model of choice for this paradigm, an elemental model with a distributed stimulus representation, e.g. McLaren, Kay & Mackintosh (1989) or AMAN (Harris & Livesey, 2010) may also be able to account for the data reported here. AMAN has the advantage of being a real-time model. It also has the advantage of having been recently created as a freely available, user-friendly version, designed for other researchers to be able to use (Thorwat, Livesey & Harris, 2012, see http://sydney.edu.au/science/psychology/staff/justinh/downloads/).

AMAN is also of particular relevance as it is an elemental model developed to be able to give an account of how humans and animals learn conditional discriminations

(Harris & Livesey, 2010). The authors of that paper note that simple elemental models have problems solving conditional discriminations such as negative patterning (A+, B+ AB-). This is because all the elements that are present on an AB- trial are also present on A+ or B+ trials, and every element on an A+ or B+ trials is also present on any AB- trial. Hence there seems no way for the model to meaningfully map these elements to whether or not an outcome occurs, as there would always be unhelpful generalisation between the elements and the compound. The same is true of biconditional discriminations (like those which are instantiated by the incongruent stimuli in the task-cueing experiments reported in this thesis), as each stimulus occurs equally often with the different outcomes, i.e. a stimulus on its own is does not predict the correct response, only the compound of two stimuli (in my experiments the cue and the incongruent stimulus) is predictive of the response. Hence, there would always be unhelpful generalisation that would prevent a purely elemental model from learning the correct response. Harris & Livesey (2010) note that this is usually taken as evidence against a purely elemental approach and in favour of a configural approach such as the one taken so far in this thesis.

However, Harris & Livesey (2010) note that the fact negative patterning and biconditional discriminations can be solved by animals and humans does not necessarily support a configural approach as an elemental model with a distributed representation such as McLaren & Mackintosh (2000, 2002) or Harris (2006) can also solve these discriminations. These models solve the discrimination by allowing different elements of the stimuli to be more active in different compounds and so counteract to some extent the unhelpful over generalisation between compounds that require opposite responses. Although both of these models are potentially capable of providing an associative explanation of the CSR condition (indeed initial investigations using a version of McLaren & Mackintosh (2002) were able to capture congruency effects and showed some evidence of a switch cost), here closer consideration to will given to the more recent development of Harris's (2006) model AMAN by Harris & Livesey (2010) and Thorwat, Livesey & Harris (2012).

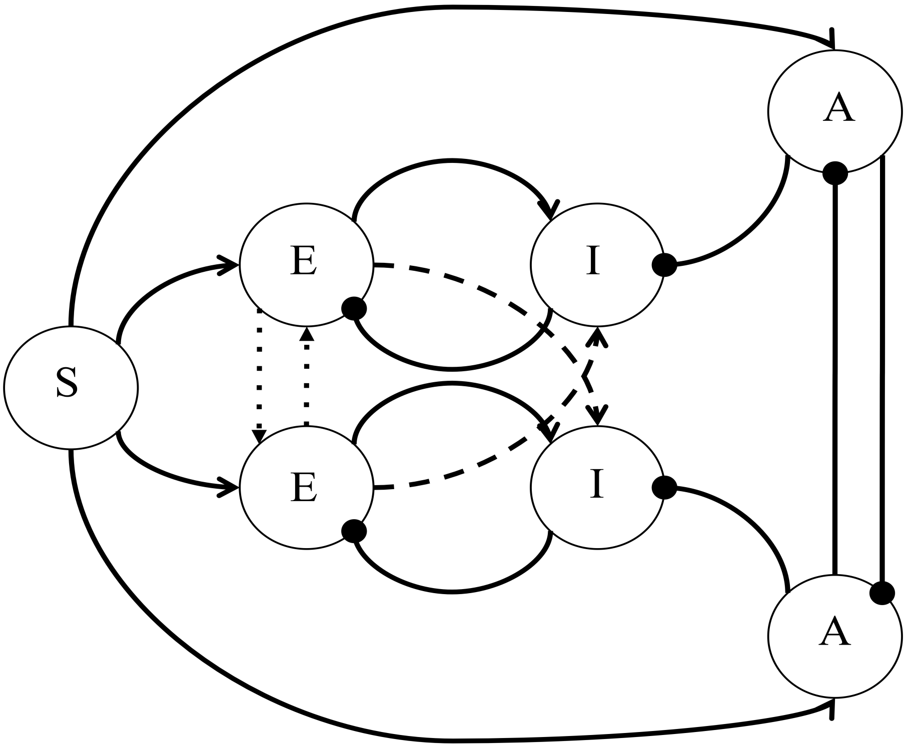


Figure 6.3 A simplified diagram of the structure of AMAN, draw as Figure 1 in Harris & Livesey, 2010. S is the stimulus presented to the network. This is then linked to various excitatory (E) units. Each excitatory unit has a corresponding inhibitory (I) and attention (A) unit. The excitatory unit excites its corresponding inhibitory unit fully but also partially activates other related inhibitory units. The attention units (A) are excited by the stimulus. These inhibit corresponding inhibitory units which releases the excitatory units (E) from inhibition.

The AMAN network consists of three types of units: excitatory, inhibitory and attention units (Figure 6.3). A given stimulus is linked to a selection of excitatory units, said to encode its spatial and featural properties. It is the links between the excitatory units that represent the learning in the network. The strength of the connections between E units depends on the temporal correlations between elements' rate of change of activation. These units each have a corresponding inhibitory unit and attention unit. The inhibitory units are used to inhibit activation of the excitatory units to normalize activity within the network. The attention units, like the excitatory units, receive input from the stimulus itself. The attention units are used to inhibit the inhibitory units, releasing the excitatory units from their inhibition. An attention unit is also normalized by activity in other attention units, so that the ability of a given attention unit to excite a given excitatory unit is diminished with more competing sensory input. This allows some of the excitatory units to be effectively selected and for the learning to take place between these selected units.

In this network, a unit's potential activation (R_{pot}) is determined by the following equation:

$$R_{pot} = \frac{Input^{p}}{Input^{p} + N^{p} + D}$$

Where $Input^{P}$ is the sum of the input to a given unit (from S and E elements). N^{P} is the sum of all normalising inputs (i.e. for E elements the I elements) and D is a constant. In order to create real-time activation the actual activation of a given unit was also determined by its previous activation so that its change in activation at a given tick is determined by:

$$\frac{dR}{dt} = \delta(R_{pot} - R)$$

Where R is the current activation of a given unit and δ is the rate parameter. The change in the weights between excitatory units is determined not simply by the concurrent activation of E units but by the rate of change of the activation of the excitatory unit and its corresponding inhibitory unit. So the change in the weight between unit x and y is related to:

$$\Delta_{x,y} = \alpha_x(\beta_E \times \frac{dE_y}{dt} - \beta_I \times \frac{dI_y}{dt})$$

where E and I denote activations of the subscripted excitatory and inhibitory units, β_E is 0.02 if dE_y/dt is positive and 0 otherwise, and β_I is 0.1 if dI_y/dt is positive and 0 otherwise. The associability of an element (α_x) is determined by:

$$\frac{d\alpha_x}{dt} = k_\alpha(E_x - \alpha_x)$$

where k_a is a constant with a value between 0 and 1; in the simulations below it was set to 0.33. Finally the actual rate of change in the weights ($V_{x,y}$) is determined by:

$$\frac{d^2 V_{x,y}}{dt^2} = k_v(\Delta_{x,y} - \frac{dV_{x,y}}{dt})$$

where k_v is a constant between 0 and 1; in the simulations below it was set to 0.05.

It is easy to see how this network would solve a component discrimination (even when the irrelevant part of the component discrimination (the cue) is also part of a biconditional discrimination). Units from the stimulus would be selected and the connections between these units and the response would be learnt. In a way this is a simple matter of conditioning. Thus the network should find this type of discrimination relatively easy to learn.

In contrast a biconditional discrimination should be much harder for the network to learn. In order to solve a biconditional discrimination Harris & Livesey (2010) suggest that the network takes advantage of the random variation in the links between the elements when the network initialises. They say that there will be some elements which are naturally more active in a given compound than other compounds containing the same elements. This means that the network effectively solves a biconditional discrimination by picking out elements which are more active for a compound and learning the mapping to the correct response from these elements. Obviously, this is more difficult than learning a component discrimination as those units which specify a compound uniquely have to first emerge from the overall stimulus representation. In other words the network has to learn to treat the compounds differently. As there are likely to be very few of these elements, this process is much slower than the component discrimination where there are many more predictive elements. Even in the case of sharing the cues between the two types of the discrimination it is still the case that fewer elements associated with the opposing response will be activated (as these can only be activated through the cue and not the stimulus for the component discrimination). This also suggests that the learning rate for congruent trials will be higher than the rate for the incongruent trials, so that there may be a development of the congruency effect across blocks in a similar manner to that shown in the CSR conditions across the experiments presented so far.

AMAN has been shown to be capable of solving these types of discriminations (Harris & Livesey, 2010) and I've already argued that it will be capable of at least showing the strong congruency effect found in the experiments presented in this thesis. However, the big question for its ability to explain the data presented in this thesis is not its ability to capture the congruency effect (as this was the reason for its selection as a possible associative model of the CSR condition) but whether or not it would also produce a "switch cost". It is obvious that the mechanisms outlined in the Chapter 5 are not available to AMAN in the same form, as AMAN does not have a hidden layer to be able to create the same configural representations that APECS does and so cannot acquire equivalence between compounds in the way that APECS can. So will it still produce a switch cost? In order to ascertain this AMAN was used to model Experiment 2 (Chapter 3).

6.2.1 AMAN Simulation of Experiment 2.

In Harris & Livesey (2010) and Thorwat, Livesey & Harris (2012) they note that their model takes into account the similarity of the stimuli used. Two stimuli, if considered to be similar, will contribute to each other's normalization. In their papers they consider two stimuli to be similar if they come from the same modality. In the experiments presented in this thesis, both the cues and stimuli used were visual and so are from the same modality. Hence one could argue that they should all contribute to each other's normalization. The responses were encoded as if they were USs, that is, they were given no specific modality.

In order to simulate Experiment 2 using AMAN, the same sequencing was used as in Experiment 2. The cues were coded discretely – that is they were represented as W, X, Y & Z which in the model meant that each cue was composed of unique elements. The stimuli were similarly coded discretely. As noted above, the cues and the stimuli were coded as being from the same modality. The ratios of the timings were also kept true to the experiment such that one "tick" represented 100ms. The parameters of the model were altered from the default settings of the freely available version outlined in Thorwat, Livesey & Harris (2012) to the parameters used in Harris & Livesey (2010). This was done because with the newer parameters the model struggled to improve its learning of the biconditional discrimination. This was probably due to both the smaller number of units used to represent a stimulus (10 in the 2012 paper and 20 in the 2010 paper) and the change in the probability of two E units being connected (1 in the 2012 paper and 0.5 in the 2010 paper) – both changes would reduce the likelihood of the model being able easily to find unique elements to carry the halves of the biconditional discrimination. In order to evaluate how well the model had learnt the responses, the extent to which the model activated the correct response minus its activation of the incorrect response in the last 5 ticks of the cue and stimulus presentation was evaluated. As with the results of the static APECS model this was subtracted from the maximum possible difference score (1), in order to make the graphs easily visually comparable with the behavioural data. The model was run 64 times.

6.2.3 Results of AMAN Simulation of Experiment 2

In order to evaluate which effects were significant in the AMAN modelling data, a repeated measures ANOVA with the following factors: task switch (switch, repeat), congruency (congruent, incongruent), CSI (long, short) and block pair (2,3,4, or 5) was performed on the difference between the maximum possible difference score and the difference between the level of activation of the correct response minus the incorrect response.

6.2.3.1 Block Pairs 2-5

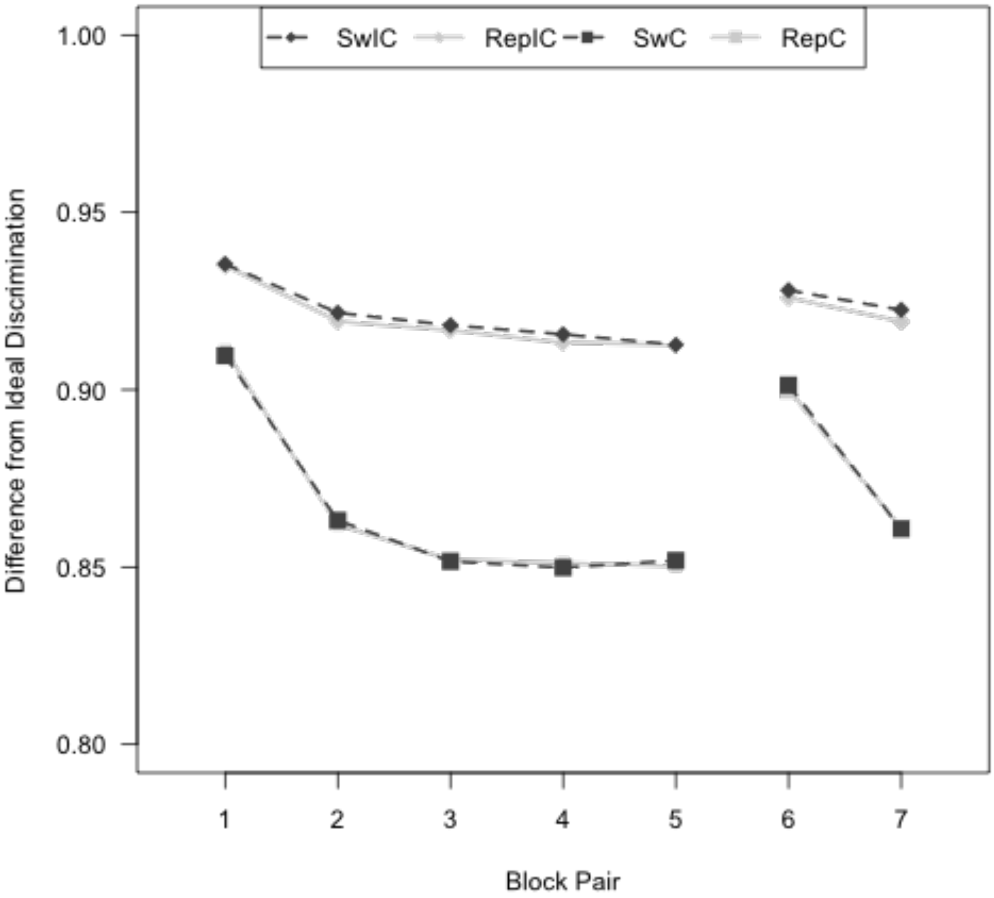


Figure 6.4 shows results of modeling Experiment 2 using AMAN. Diamonds represent incongruent stimuli, whilst squares represent congruent stimuli. Dashed dark lines represent switch trials, whilst solid lighter lines represent repeat trials. The break represents when novel stimuli were introduced.

Figure 6.4 shows that the AMAN simulation of Experiment 2 demonstrated an improvement in performance from block pairs 1 to 5. It then shows a notable perturbation in block pair 6 when the stimuli used are altered, for both the congruent and incongruent trials. The model also shows a very clear congruency effect; it is much faster to learn the correct response for the congruent than the incongruent trials. The graph also is suggestive of a very small switch cost for the incongruent trials, where the performance of the model is very slightly worse on switch than repeat trials.

Congruency. There was a large congruency effect (0.062) this was reliable, F(1,63)= 108, $p<0.001$. This effect is consistent with the large and highly significant congruency effect found in the behavioural data.

Task switch. In agreement with the behavioural data for the CSR condition in Experiment 2, switch trials (0.8856) had worse performance than repeat trials (0.8847), F(1,31) = 7.47, $p<0.001$. The switch cost was larger in the incongruent stimuli (0.00153) than the congruent stimuli (0.000245) but this interaction did not reach significance, F(1,63)=2.52, p=0.118.

Preparation. As this model can encode time we can examine the effect of CSI. Preparation increased the switch cost very slightly, long CSI switch (0.0011); short CSI (0.0007), but this effect was not significant, F<1. The model performed overall better in a long CSI (0.881) than a short CSI (0.890), F(1,63)=305, $p<0.001$.

Acquisition effects. Overall performance evidently improved with block pairs, and this was reliable F(3,189)= 237, $p<0.001$. The size of the congruency effect changed over time, F(3,189) = 19.1, $p<0.001$, see Figure 6.4. In agreement with the behavioural data and APECS, the congruent stimuli were learnt more quickly than the incongruent stimuli.

There was also a significant interaction between congruency, switch and block pair, F(3,189)=2.94, $p<0.05$. This is driven by the switch cost being larger for incongruent stimuli (block pair 2: 0.00245, block pair 3: 0.00140, block pair 4: 0.00238, block pair 5: -7.73×10^{-5}) than congruent stimuli (block pair 2: 0.00122, block pair 3: -0.0005048, block pair 4: -0.00132, block pair 5: 0.00158) in all block pairs except block pair 5.

6.2.3.2 Transfer Analysis

The model shows a marked perturbation in block pair 6 when the new stimuli are introduced (Figure 6.4). The effect of transferring participants from one set of stimuli to another was tested using a mixed ANOVA with the following factors: task switch (repeat or switch), congruency, CSI, block pair (5 or 6).

Transfer. As expected there was a main effect of transfer, in which the model was worse after the stimuli had been changed (0.914) than before (0.882), F(1,63)=4737, $p<0.001$.

The congruency effect changed when the stimuli were changed, it was much smaller after the stimuli had been changed (0.026) than before (0.062), F(1,63)= 103, $p<0.05$.

6.2.4 Discussion of the results of the AMAN model of Experiment 2

All in all AMAN is broadly able to account for the findings of Experiment 2. It produces a large congruency effect that changes over time due to the congruent stimuli being much easier for AMAN to learn. It also produces a small but reliable switch cost that, although not significantly so, is limited to the incongruent trials. Indeed although the interaction is not significant, t-tests revealed that the switch cost was not significant for the congruent stimuli, t(63)=0.356, p=0.553, but was for the incongruent stimuli, t(63)= 6.30, $p<0.05$. This suggests that AMAN, like APECS, is able to generate a switch cost that is largely due to the incongruent stimuli.

AMAN produces a large effect of CSI, in which the performance of the model is better at a long CSI than at a short CSI. This is agreement with the results of Experiment 2, which also showed an advantage for a long CSI. This results is also in agreement with the results of Experiment 3 which showed a main effect of CSI. The ability to capture this effect may suggest the way in which AMAN encodes time is better than the real-time APECS model presented above. Although as mentioned above it may be that the main effect of CSI also represents processes which are beyond the scope of an associative model, such as phasic alertness effects (Niemi & Näätanen, 1981; Nobre, Correa & Coull, 2007). The relative contribution of the potential for associative processes to be driving these effects and other processes would require further research.

Figure 6.5 shows the weights in the excitatory layer of AMAN at the end of an example simulation. It shows that, as expected, the congruent stimuli (A,B,E,F) are positively connected to their correct outcomes. There are also more elements carrying these mappings to a given response than is the case for the incongruent stimuli (C,D,G,H). The incongruent stimuli have connections to both responses, using some different elements but also some of the same elements. The cues have only very weak connections to the outcomes compared to the stimuli. This is probably because the cues may have a more occasion setting role, rather than a direct link to the output due to the long CSI blocks. This may explain the main effect of CSI described above, as if the cue is able to pre-activate the elements in the stimulus which are associated with the correct response before the stimulus is presented, this process would help performance in that trial.

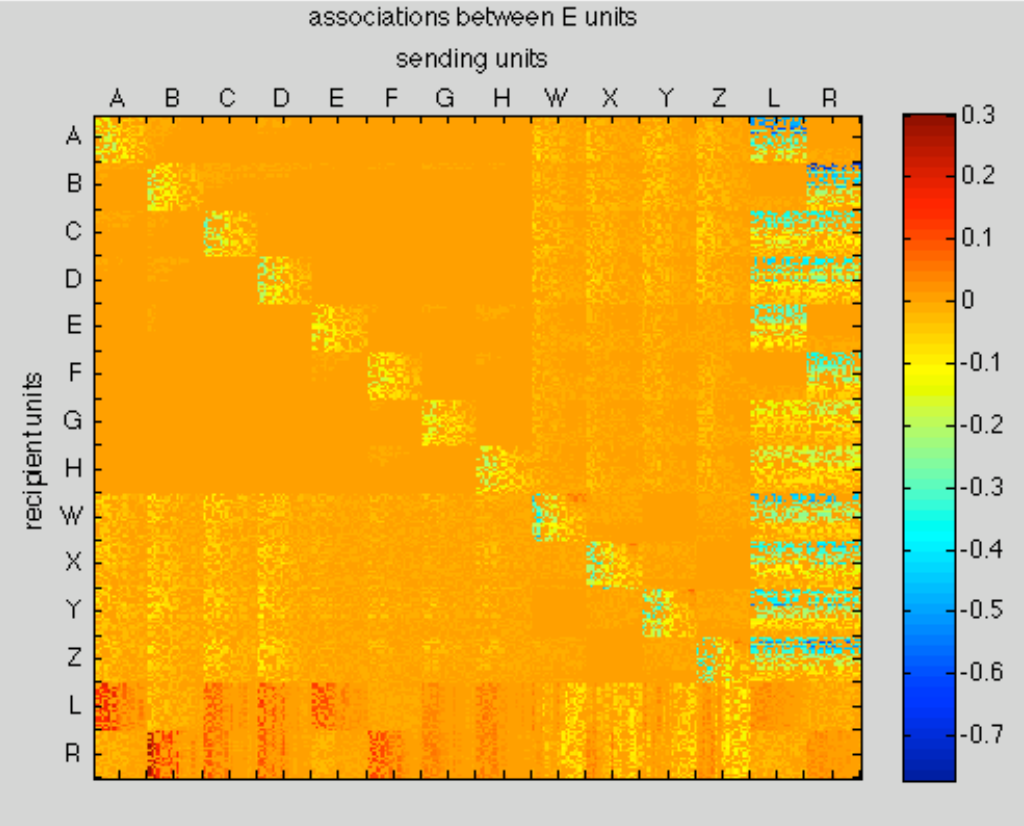


Figure 6.5: An example of the strength of the links between the E units at the end of training. This model was trained with A, B, C, D and transferred to E, F, G, H. Redder colours represent excitatory links and bluer colours represent inhibitory links. A, B, E and F are congruent stimuli, whilst C, D, G and H are incongruent stimuli. W and X are equivalent cues as are Y and Z. L and R are the response units. See Tables 3.1 & 3.2 for a table of all the mappings.

Given that AMAN is reasonably successful at modelling the results of Experiment 2, it is worth considering whether the switch cost produced by AMAN is driven in the same way as the switch cost produced by APECS. In order to do this we can look at the results of modelling the experiments presented in Chapter 5, and ask whether AMAN will still predict a switch cost when different cues are used for the component and biconditional discriminations? Or indeed when there is no longer a CSI?

6.2.5 AMAN model of Experiment 5

AMAN was run 64 times in each condition using the same sequencing as used for APECS and described in more detail in Chapter 5.

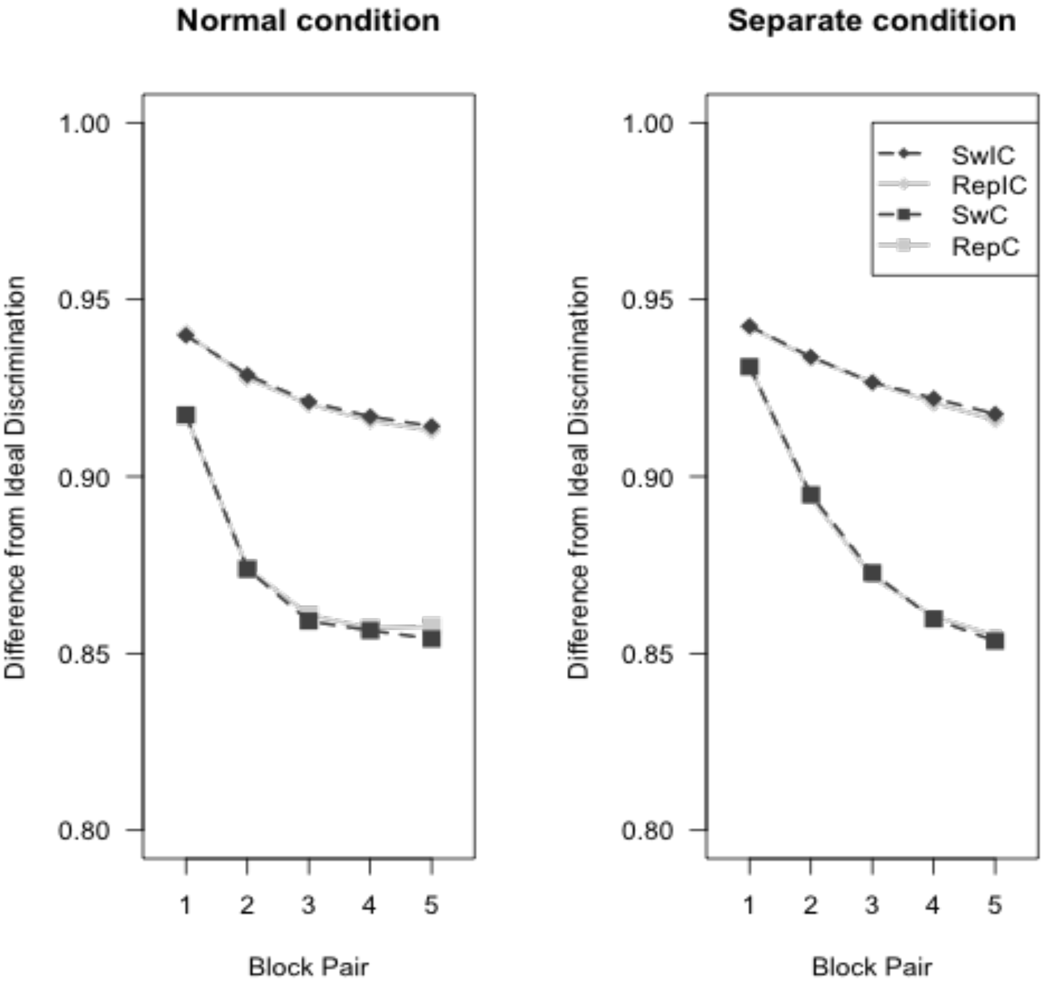


Figure 6.6 shows the results of modelling Experiment 5 using AMAN including direct cue-stimulus repetitions. Diamonds represent incongruent stimuli, whilst squares represent congruent stimuli. Dashed, dark lines show switch trials whilst solid lighter lines represent repeat trials.

Figure 6.6 shows the modelling results for the two conditions. Unlike the AMAN model of Experiment 2 there now seems to be no switch cost evident in either the Normal or Separate conditions. The main notable difference between the conditions seems to be the rate at which the mappings are learnt. Learning both the congruent and incongruent stimuli seems to be slower (at least initially) in the Separate condition (where cues are not shared across biconditional and component discriminations) than the Normal condition.

In order to show if AMAN's performance was affected by the a mixed ANOVA was run on the data with the factors, "task switch" ("switch" or "repeat"), congruency (congruent, incongruent), block pair (2,3,4,5)[13] and condition (Normal or Separate). This ANOVA was followed up with ANOVAs on each condition where appropriate.

"Task switches" and condition There was no overall effect of task switch, F(1,126)<1 nor did the differences in the switch cost between the Normal condition (-0.000287) and the Separate condition (0.000345) result in a significant difference between the groups, F(1,126)=2.33, p=0.13. Planned comparisons showed that there was no reliable effect in either the Normal condition, F(1,63)<1 or the Separate condition, F(1,63)=1.77, p=0.188. Note that these results for the Normal condition are rather different to those for Experiment 2. However, there are differences between the two experiments, unlike Experiment 2, Experiment 5 had no CSI, cue repeats were allowed to occur and the "task" only "switched" a third of the time rather than half.

"Switch" by congruency and condition. The "switch cost" in the Normal condition was larger for incongruent trials (0.000979) than the congruent trials (-0.00155) and this difference was larger than the difference between the "switch cost" seen in the Separate incongruent trials (0.000798) and congruent trials (-0.00108), unlike for the APECS simulation, this interaction did not quite reach significance, F(1,126)= 3.71, p=0.056. Separate analyses revealed that the "switch cost" was significantly larger for the incongruent stimuli in the Normal F(1,63)=16.3, p<0.001, but not the Separate condition, F(1,63)=2.57, p=0.114. There were nearly significant "switch costs" in the incongruent trials for both the Normal, F(1,63)=3.65, p=0.06, and Separate conditions, F(1,63)=3.61, p=0.06 and a significant "switch benefit" for the congruent trials in the Normal condition, F(1,63)=16.8, p<0.001 but not in the Separate condition, F<1.

Congruency and condition There was a significantly larger effect of congruency in the Normal condition (0.0581) than the Separate condition (0.0545), F(1,126)=28.1,

---

[13] Note that block pair 1 was not included in this analysis as for some of the models they seemed to need this block pair to settle in their performance.

p<0.001. The congruency effect was reliable in both the Normal, F(1,63)=9742, p<0.001, and Separate conditions, F(1,63)=20427, p<0.001.

Acquisition effects and condition. Overall performance reliably improved from block pairs 2 to 5, with the improvement being faster in the Normal condition than the Separate condition (Figure 6.6), F(3, 378)= 552, p<0.001. There were significant effects of block pair in both the Normal condition, F(3,189)=2067 p<0.001 and the Separate condition, F(3,189)=6069, p<0.001.

The three-way interaction between condition, block pair and congruency was significant F(3,378)= 94.5, p<0.001. There were significant block pair and congruency interactions in both the Normal condition, F(3,189)=14.1, p<0.001 and the Separate condition, F(3,189)=389, p<0.001. This interaction can be seen in Figure 6.6, which shows the congruency effect initially being larger in the Normal condition. This initial difference seems to be largely due to the congruent trials taking longer to learn in the Separate condition than the Normal condition.

Condition There was also a main effect of condition, whereby performance was better in the Normal condition (0.109) than the Separate condition (0.103), F(1,126)=432, p<0.001.

6.2.6 Discussion of AMAN’s simulation of Experiment 5

The foregoing suggests that AMAN gets different results from APECS and from the behavioural data (including direct cue-stimulus repetitions). It suggests that in cases without a CSI and allowing cue repeats to occur, AMAN no longer predicts a switch cost, even for the Normal condition. It can be seen in Figure 6.7 (and also in Figure 6.8) that the nature of the cue’s (W,X,Y,Z) relationship to the outcome differs from that seen in Figure 6.5. The cues and the incongruent stimuli are now nearly equally strongly connected to the responses, suggesting that instead of the cue in some way mediating the relationship between the stimulus and the response units, it is now playing a more direct, joint role in activating the response, i.e. the cue is no longer an “occasion setter”.

Why might the cue becoming an “occasion setter” in the simulation of Experiment 2 lead to a switch cost? As with APECS, one might consider the trial transition XC-L, YC-R to be a particularly disadvantageous transition for AMAN, as presenting XC will interfere with (extinguish) YC leading to weaker responding to YC (Livesey, personal communication). This could produce a switch cost as this trial transition can only occur on a switch trial. Indeed removing this trial type significantly decreases the size of the switch cost found in the model of Experiment 2 (including XC-L, YC-R: 0.00258; excluding XC-L, YC-R: 0.000859), $F(1,63) = 6.00$, $p<0.05$. However, this is not the sole driver of the switch cost, as even after this transition is removed the remaining switch cost is still significant, $F(1,63)= 6.79$, $p<0.05$. When the cues are being treated as occasion setters the majority of the mapping is carried by the stimulus, and so stimulus-based interference is larger than cue-based interference. However, when the cue is not an occasion setter, cue-based interference is just as damaging as stimulus-based interference. Thus, consecutive trials in which the same cue requires opposing responses are just as challenging as consecutive trials in which the same stimulus requires opposing responses. This suggests that part of the reason AMAN may have produced a switch cost in the simulation of Experiment 2 may be its ability to treat the cues as occasion setters. In support of this the switch cost seen in Experiment 2 in the incongruent trials is only separately significant in the long CSI (a parameter value that should favour occasion setting), $t(63)=2.10$, $p<0.01$, but not the short CSI, $t(63)=1.60$, $p=0.12$.

In the previous chapter it was noted that the transition XC-L, XD-R is in many ways associatively equivalent to the transition picked out above but can only occur on repeat trials. AMAN would consider these two transition types to be equivalently disadvantageous only if the cues were not being treated as occasion setters. Indeed in the model of Experiment 5 there was no difference between these two transition types in either the Normal condition (XC-L, XC-R: 0.0565; XC-L, XD-R: 0.0562), $F(1,63)<0$ or the Separate condition (XC-L, XC-R: 0.0561; XC-L, XD-R: 0.0571), $F(1,63)=2.87$, $p=0.095$[14]. So more simply the lack of switch cost in the model of Experiment 5 but the finding of one in the model of Experiment 2 may be also due to subtle differences in the sequences, which mean that in Experiment 2 the trial transition XC-L, XD-R was not present, as there were no cue repeats.

---

[14] There was no interaction between the Normal and the Separate condition, $F(1,126)=2.33$, $p=0.130$, for the differences between these two transition types.

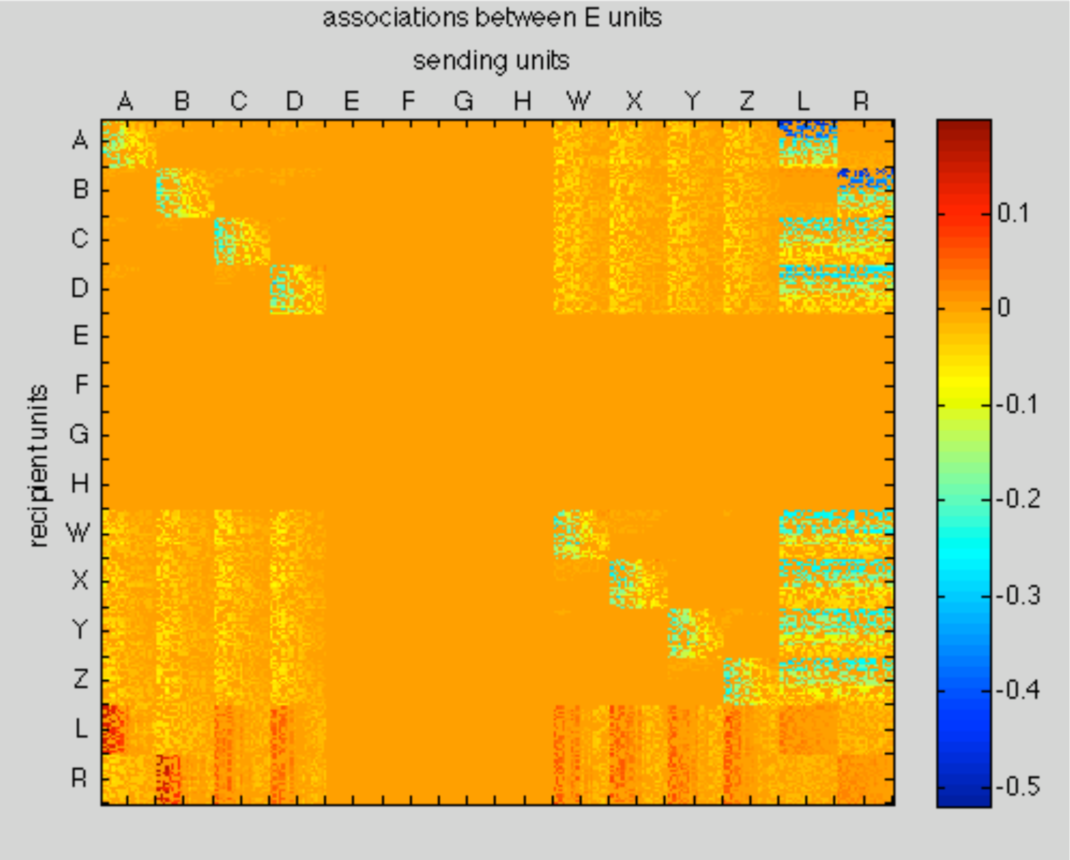


Figure 6.7 shows an example of the final weights between the excitatory units for the Normal condition. For the model only the stimuli A, B, C & D were used. These could appear with any of W, X, Y & Z. A & B were the congruent stimuli, where A was always associated with a left response (L) and B was always associated with a right response (R). C & D were incongruent stimuli. See Table 6.1 for a table of the mappings.

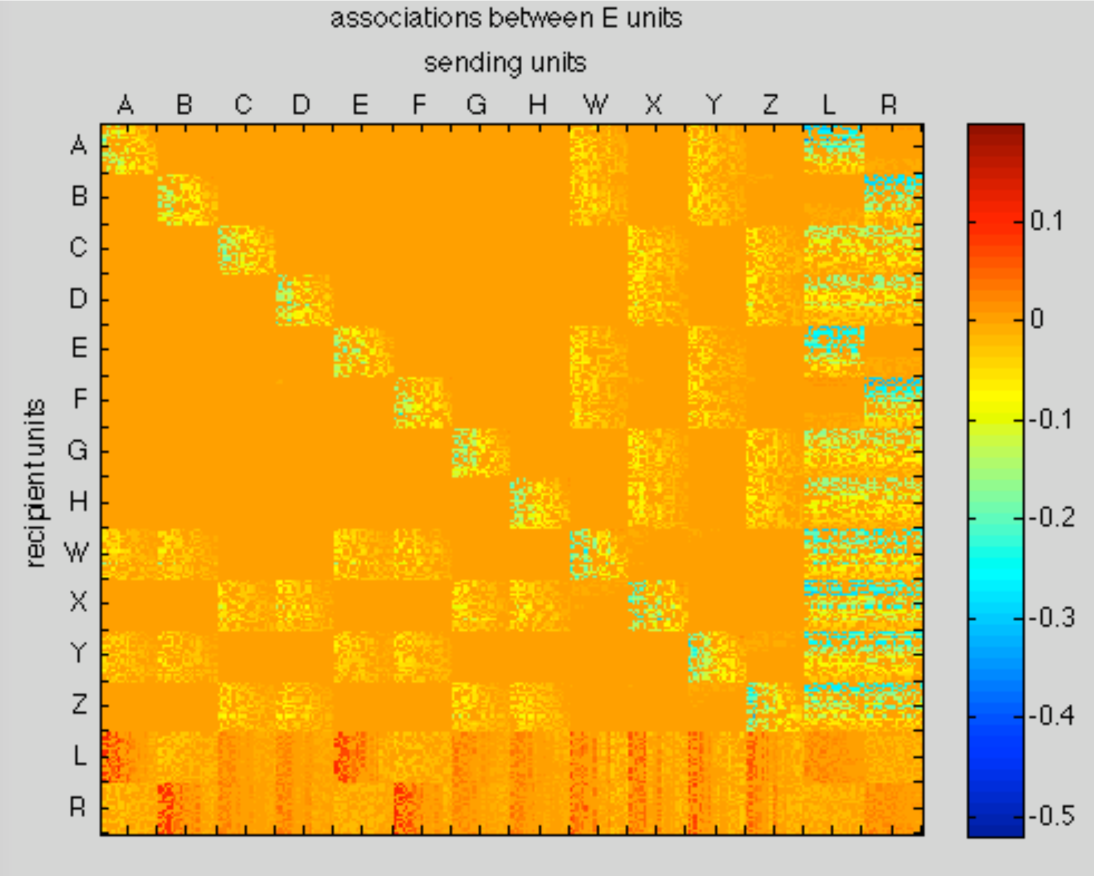


Figure 6.8 shows the final weights between the E units for an example from the Separate condition. Here A,B,E & F was congruent stimuli which only appear with W & Y, where A & E are always associated with a left response (L) and B & F are always associated with a right response (R). C,D,G & H are incongruent stimuli which only appear with X & Z. See Table 6.2 for a table of the mappings.

Given that a major change, apart from the lack of CSI, from Experiment 2 to Experiment 5, is that the sequencing allowed direct cue-stimulus repetitions to occur, the data was analysed without the direct cue-stimulus repetitions included.

6.2.7 Modelling results without direct cue-stimulus repetitions

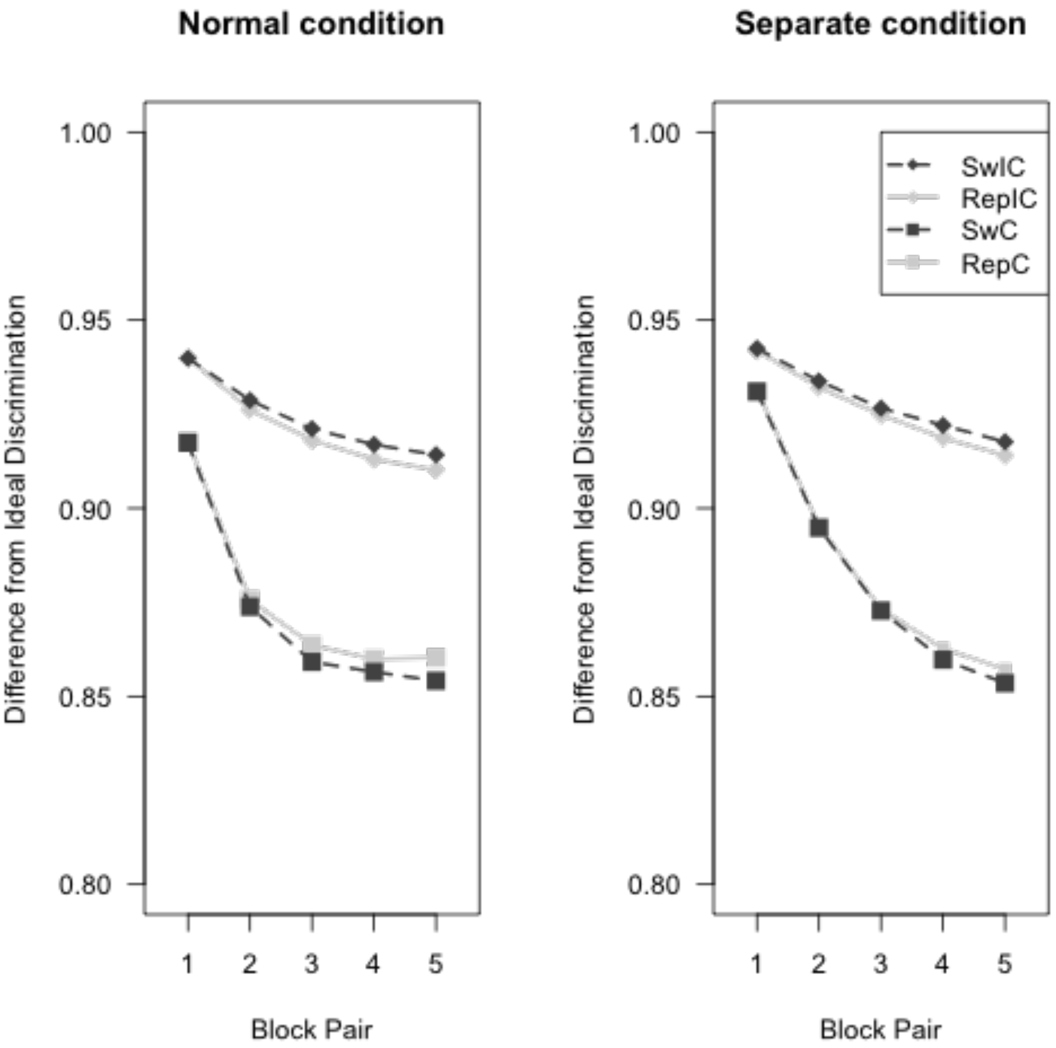


Figure 6.9 shows the results of modelling the Normal condition and the Separate condition when direct cue-stimulus repetitions are removed from the analysis. Diamonds represent incongruent stimuli, whilst squares represent congruent stimuli. Dashed, dark lines show switch trials whilst solid lighter lines represent repeat trials.

By contrasting Figure 6.6 with Figure 6.9 one can see that AMAN seems to show the emergence of a switch cost for the incongruent stimuli and the emergence of a switch benefit for the congruent stimuli when direct cue-stimulus repetitions are removed in both conditions. The same ANOVA technique was used to analyse the data.

“Task switches” and condition There was a larger “switch cost” in the Separate condition (0.000358) than in the Normal condition (-0.000328) but this was not

significant, F(1,126)= 2.81, p=0.096. These costs were not reliable in the Normal condition, F(1,63)=1.09, p=0.301, or the in the Separate condition, F(1,63)=1.88, p=0.176.

"Switch" by congruency and condition. The "switch cost" in the Normal condition was significantly larger for incongruent trials (0.00326) than the congruent trials (-0.00391) and this difference was larger than the difference between the "switch cost" seen in the Separate incongruent trials (0.000255) and congruent trials (-0.00184) and was reliable, F(1,126)= 10.3, p<0.01. Separate analyses revealed that the "switch cost" was significantly larger for the incongruent stimuli than the congruent stimuli both for the Normal condition F(1,63)=129, p<0.001, and the Separate condition, F(1,63)=54.6, p<0.001. The "switch cost" was only reliable in the Normal condition for the incongruent trials, F(1,63)=42.4, p<0.001, and in the Separate condition's incongruent trials, F(1,63)=34.7, p=0.201. The "switch benefit" was reliable for the congruent stimuli in the Normal condition, F(1,63)=105, p<0.001, and the Separate condition, F(1,63)=27.0, p<0.001.

Congruency and condition There was a significantly larger effect of congruency in the Normal condition (0.0556) than the Separate condition (0.0525), F(1,126)=18.4, p<0.01. The congruency effect was reliable in both the Normal, F(1,63)=8452, p<0.001, and Separate conditions, F(1,63)=18814, p<0.001.

Acquisition effects and condition. Overall performance reliably improved from block pair 1 to 5, the improvement was faster initially in the Normal condition than the Separate condition (Figure 6.4), F(3,378)= 455, p<0.001. There were significant effects of block pair in both the Normal condition, F(3,189)=1769, p<0.001 and the Separate condition, F(3,189)=5061, p<0.001.

The three-way interaction between condition, block pair and congruency was significant F(3,378)= 93.0, p<0.001. There were significant block pair and congruency interactions in both the Normal condition, F(3,189)=9.63, p<0.001 and the Separate condition, F(3,189)=353, p<0.001. This interaction can be seen in Figure 6.9 which again shows the congruency effect initially being much larger in the Normal case as AMAN learns the congruent trials faster.

Condition There was also a main effect of condition, whereby performance was better in the Normal condition (0.109) than the Separate condition (0.103), F(1,126)=1407, p<0.001.

6.2.8 Discussion of AMAN's model of Experiment 5 without direct cue-stimulus repetitions

Removing direct cue-stimulus repetitions seems have had opposing effects for incongruent and congruent stimuli. For the congruent stimuli, removing direct cue-stimulus repetitions made performance on task repeat trials worse, suggesting that these trial types were particularly beneficial examples of a task repetition trials. Whereas for the incongruent stimuli, removing direct cue-stimulus repetitions made performance on repeat trials better, suggesting that these trials types were particularly disadvantageous examples of task repetitions trials.

In general AMAN suggested a different pattern of results for Experiment 5 than APECS. It predicted that the main difference between the Normal and Separate conditions would be that learning was slower in the Separate condition for both the congruent and incongruent stimuli. This is to be expected as the Separate condition uses more stimuli than the Normal condition and so even if there are the same number of pairs in both conditions, there are fewer repetitions of a given stimulus in the Separate condition. This can be seen by comparing Figures 6.7 & 6.8 which clearly show weaker connections to the outcomes in the Separate condition (Figure 6.8) than the Normal condition (Figure 6.7). However, no real difference was found between the conditions in terms of overall performance.

AMAN also showed that for the incongruent stimuli there would be no "switch cost" in either the Normal or Separate condition when direct cue-stimulus repetitions were included, but that one would emerge in both conditions when they were excluded. This is contrary to the results of Experiment 5 which showed a larger "switch cost" when direct cue-stimulus repetitions were included than when they were excluded. AMAN also showed that for congruent stimuli in both the Normal and Separate condition there would be no switch cost when direct cue-stimulus repetitions were included, but when they were excluded a "switch benefit" emerged. Again this was not observed in the behavioural data.

In total these results suggest that, although AMAN was able to account for the data of Experiment 2, it seems, using the parameters from Harris & Livesey (2010), to be unable to account at present for the findings of Experiment 5.

6.3 Conclusions

In this chapter, I have attempted to tackle one of the major limitations of the modelling previously presented in this thesis. The version of APECS used up to this stage has been a static one, unable to represent temporal relations (see Sutton & Barto, 1981, 1987 for a discussion of these issues). The static version of APECS is unable to model reaction times and error rates, make predictions relating to the CSI, or realistically model the way that human participants learn from experience.

The real-time version of APECS presented in this chapter models reaction times and error rates, and when carefully parameterised it performed at least as well as the static version of APECS in accounting for the data of Experiment 2. It achieved these results by successfully learning in a more realistic way from feedback, and in that regard provides a potentially closer model of human behaviour.

This chapter also considered an alternative real-time associative learning model, AMAN. As with APECS, this model was able to account for most of the data of Experiment 2, and even produced a main effect of CSI– an effect produced by human participants but not by any version of APECS presented in this thesis. AMAN and real-time APECS, then, were both able to present an account of how humans might be learning under CSR conditions in a task-cueing paradigm, in particular modelling the initially surprising “switch cost”. But these accounts differ in their underlying logic, and the differences become apparent when the models are applied to Experiment 5. The mechanisms by which APECS achieves the switch cost have already been discussed in Chapter 5. For AMAN, the investigation presented here does not constitute a full analysis, but one mechanism for the switch cost comes from the network’s ability to occasion set. This leads to the cue having a different status to the incongruent stimulus with regards to its relation to the response. When the cue is occasion setting, the incongruent stimulus has a stronger link to the response, and this means that there is more stimulus-based than cue-based interference. This leads to more unlearning of the response for the opposing task for the incongruent stimulus than for the other stimuli that can appear with the cue. This process means that the network is disadvantaged when the "task" changes, but the incongruent stimulus repeats, as the link from the stimulus to the opposing response has been weakened. Although the evidence presented earlier in this chapter suggests a role for a process such as this producing a switch cost in AMAN, there may be other mechanisms in play.

However, both of the models in this chapter have limitations, and neither was able to account for all of the data presented in this thesis. The real-time APECS model failed to predict a “switch cost” when the probability of a “task switch” was changed to a third. AMAN was unable to account for the findings of Experiment 5. An examination of the limitations of both models, however, indicates some potential ways forward

More consideration of how a real-time APECS could encode time may be one avenue, as although simply activating the units within the existing static model in real-time had some success, it was far from perfect. One possibility would be to explore the recurrent version of this type of network (c.f. the SRN, Elman, 1990), this would mean that the network would not just have to learn from the decaying representations of the cue and stimulus as it would also have their recurrent representation. Another advantage introducing recurrence might give is the ability to learn about sequences. Given that switch cost is potentially partially a sequential effect it seems reasonable consider what happens when the model is allowed to be able to learn about sequential effects.

Similarly, although AMAN was unable to account for all the data, it was able to account for a good deal of it, and this suggests alternative associative mechanisms that may be driving some of the effects found in the CSR condition. For example, it might be of interest to consider whether the mechanisms enabling AMAN to produce a switch cost in its simulation of Experiment 2 are part of the reason human participants produce a “switch cost”. I’ve already noted that one reason it may be producing a switch cost is the network’s ability to occasion set. This is not an unreasonable hypothesis, but the evidence presented in this thesis does not seem to support it as the full explanation (given that there are still clear switch costs present when there was no CSI in Experiment 5). One could argue for a very slight contribution of occasion setting to the size of the “switch cost”, as in Experiment 5 the “switch cost” after direct cue-stimulus repetitions was removed was 25ms in the Normal condition but was 32ms in Experiment 3, its closest comparator, with a long CSI. But this difference is also in line with the static APECS model which also suggested a slightly smaller “switch cost” for the model of the Normal condition without stimulus repeats (0.018) than for the model of Experiment 3 (0.02). However, there is some additional evidence to suggest a role for an occasion setting mechanism, as there was a slightly larger switch cost (but not significantly so) in a long CSI than a short CSI in the Monsell Experiment and

Experiments 2 and 3 but not 1, this is consistent with the distribution of the switch cost shown by AMAN. Given this mixed picture it would be worthwhile to investigate the potential contribution of an occasion setting mechanism to an associative "switch cost" in the future. It would also seem worthwhile to investigate whether the failure of AMAN to produce a "switch cost" in Experiment 5 represents a failure of the principles of AMAN or a failure of the specific parameters and details of its particular instantiation.

The results of this chapter have suggested that, despite some successes, neither the real-time version of APECS nor AMAN are fully satisfactory real-time models of the associative processes involved in a task-cueing paradigm run under CSR instructions. However, both models were able to account for the data to some extent, and an examination of the mechanics of their successes and failures has provided some insight into ways in which these models may be adapted in future to more fully account for the data.

In addition, considering the different classes of associative model in this way has lead to a greater understanding of what role the different parameters of the experiment, such as the CSI and the presence of specific trial transitions, may play in generating associatively based accounts of effects. Understanding which effects can be accounted for by associative models can help illuminate whether we need to posit controlled processes to explain certain aspects of behaviour, most notably it worth considering that neither associative model generated a RISC effect.

Chapter 7 General Discussion

The experiments reported in this thesis have found clear evidence of a difference in both learning and performance when participants approach a task-cueing paradigm using CSR mappings compared to when they approach it using task-sets. The thesis has also explored in detail the processes underlying performance in the CSR condition. This chapter summarises and draws together the main experimental findings and discusses them with reference to the key concepts introduced in Chapter 1: the compound retrieval model; what a task-cueing paradigm measures; the advantages of each approach; how participants in the CSR condition learn; and whether it is useful to posit both associative and controlled processes. Finally the chapter suggests directions for further research, first testing the conclusions made in this thesis and then generalizing them to a wider context.

7.1 Central findings

In all of the experiments participants saw coloured shapes which served as cues, in the Tasks condition (Monsell Experiment, Experiments 1-3) these specified the categorisation task to be performed on the stimulus (a digit in the range 1-9) that subsequently appeared superimposed on the coloured shape cue. In the Tasks condition, the tasks used were always to classify whether a digit as higher/lower than 5 or as odd/even. In all the experiments the core phenomena considered were: the effect of changing the task-set versus repeating the same task-set, i.e. the switch cost; the effect on the switch cost of having a longer time to prepare (a greater time between the presentation of the cue and stimulus; the difference in performance between response-congruent and response-incongruent stimuli. The *key* manipulation used throughout this thesis was to compare the standard Tasks condition with one where participants were not informed of the tasks, but instead told to directly learn (either by trial and error or via explicit instructions) the cue-stimulus->response (CSR) mappings present in the experiment.

The central finding from the Monsell Experiment and Experiments 1-3 was that when participants approach a task-cueing paradigm using CSR mappings and without reference to the available tasks (as was the case for the CSR condition in my experiments), then their performance differs from the performance of those participants who have knowledge of the tasks and appear to be using the cue to select a task and then categorise the stimulus accordingly. Specifically, the CSR participants show a smaller switch cost, which does not reduce with time to prepare, and a much larger congruency effect which changes over time. This can be contrasted with those participants in the Tasks condition who show a larger switch cost that reduces with preparation and a stable, smaller congruency effect. These findings were relatively insensitive to the manner in which the CSR group was trained as the results were broadly the same when: participants were merely given a hint not to use the tasks (Monsell Experiment), given an instruction sheet containing all the mappings (Experiments 1 & 3) or just told to learn by trial and error (Experiments 2, 4 & 5). Experiment 3 suggested that it was possible for participants to change their approach from not using task-sets to using task-sets and to some extent vice versa. It also suggested that different processes were involved under Tasks and CSR instructions, and the asymmetry in the benefits of prior practice in the other condition suggested that the CSR condition involved more automatic learning which could proceed in the background during the performance of the Tasks strategy whilst the Tasks condition involved more controlled processes which required explicit practice to perfect.

Consideration of why the associative structure of the task-cueing experiment would lead to an apparent "task switch" cost led to Experiments 4 and 5, which investigated the consequences of changing the associative structure so that the biconditional and component discrimination no longer shared cues and so that there was only one cue per discrimination per task. The static instantiation of APECS predicted that with these structural changes, after removing direct cue-stimulus repetitions from the analysis the switch cost would be abolished in the CSR condition. But Experiment 4 suggested that when these changes were made to the associative structure they encouraged task induction as it was now possible to solve the problem by inducing just one task rule with an exception as corollary. When participants were prevented from inducing task rules, in this way, by replacing the numbers with symbols, Experiment 5 found that participants' behaviour was in line with the results of the modelling. This suggests that the nature of the experimental design influences the associative structures formed, and that the use of cues common to both biconditional and component discriminations may be causing a "switch cost" to exist in the CSR condition.

7.2 Implications for the compound retrieval model

Taken together, the results of these experiments lead to the conclusion that the standard pattern of performance in the task-cueing paradigm involves more than just learning and using CSR mappings, modulated by priming of cue encoding, as proposed by Logan et al's compound retrieval model. The large discrepancy between the pattern of and relation between task-switch, congruence, preparation and transfer effects in the Tasks and CSR condition clearly demonstrates that very different processing strategies are involved, as does the state-trace analysis of Chapter 4.

In addition as the compound retrieval model account of the switch cost relies on the differential amount of time taken to encode a cue on a switch and repeat trial, the model may struggle to account for the finding in Chapter 5 that with the same cue sequence there are differences in the sizes of the switch costs between the Normal and Separate conditions. With the same cue sequence the cue encoding aspect of the model should be the same in both conditions, and it this which is largely responsible for the switch cost in Logan and colleagues model. In addition their model may have problems explaining why in the Separate condition without including direct cue-stimulus repetitions in the analysis there is no switch cost, even though cue repeats can still occur on repeat trials, but not on switch trials.

One could perhaps save the compound retrieval/cue priming account as a model of the standard Tasks condition by allowing the cue to not just act as part of the retrieval for the cue+stimulus mapping but to bias attention to the features of the stimulus that are task relevant. However, this seems no more than a variant of a generalized task reconfiguration account.

7.3 Implications for the Allport and colleagues theories

As well as considering the impact of this research for Logan and colleagues' theories it is also worth considering how they impact on the theories of Allport and colleagues (Allport et al, 1994; Waszak, Hommel & Allport, 2003). Allport et al (1994) originally argued that "the switch cost was driven by a kind of pro-active interference (PI) from competing CSR mappings with the same stimuli, persisting from the instruction set on preceding trials" (p. 436). In this explanation of the switch cost the PI from competing S-R mappings is also available to the CSR condition. However, the fact that this interference is related to the persistence of the instruction set is unavailable in the CSR condition. The arguments presented in Chapter 5 suggested that there may be some part of the switch cost in the CSR condition which is caused by interference between competing representations of stimulus response mappings, but only for the incongruent trials. The difference in size of the switch cost (and indeed the residual switch cost) between the Tasks conditions and CSR conditions suggests that if this theory does explain part of the switch cost, then the persistence of the task-set is necessary for it to explain more than just the small switch cost seen for just incongruent stimuli.

A more recent version of this type of theory was proposed by Waszak et al (2003). They suggest that bindings (associations) occur between all the encoded constituents of the action event: not only between the 'immediate stimulus and its response ("S-R bindings"), but also with the distal goal of the action, the task and task-specific processing operations, and possibly other relevant contextual features' (p 367). Again, in this theory some elements of the associations are available in the CSR conditions: stimulus and its response, the distal goal of action and other relevant contextual features. As the performance differs in the Tasks condition, it would seem that the connections to a form of mental representation of the tasks may mediate these links, suggesting that in the CSR condition that they could be merely learning a sub-set of all the associations formed in the Tasks condition.

However, Experiment 3 might suggest otherwise. The CSR-Tasks group's performance under Tasks conditions was equivalent to those participants who started under Tasks conditions. One would expect that if the same S-R links mediated performance in both conditions then there should be some advantage of having set up these links under CSR conditions. So it could be that the links set up under Tasks conditions are different to those under CSR conditions, instead of a 7->Left link, perhaps the links are more of the form 7->odd->Left and hence the existing SR links do not necessarily help with performance. This could be part of the difference between the two conditions: the task-sets bias the route between the stimulus and the response. Indeed Waszak et al's (2003) description of SR episodes is that they should be viewed as stimulus-task-action links. The results from the Tasks-CSR condition suggest that the 7-> Left link is being formed in the same manner while participants are using task-sets, although this link is not at that time in control of the participant's performance.

7.4 What do standard task-cueing experiments measure?

The results presented in this thesis support the popular interpretation of performance in a task-cueing paradigm under Tasks instructions. That is, that task-cueing paradigms for the most part measure participants' ability to use and flexibly change between task-sets. The switch cost seen under Tasks instructions is substantially larger than the switch cost seen when participants are not relying on task-sets (i.e. under CSR instructions), suggesting that the switch cost under these conditions may genuinely reflect processes involved in top-down control or reconfiguration of task-set (Monsell, 2003). It seems that this hierarchical structuring of action control (select task-set, then select response under this task set) is an efficient strategy for performance, even when a small stimulus set makes a "flat" single task-set CSR strategy readily available. Among the advantages of a task-set strategy are the ability to prepare in advance, and the ability to cope easily with novel or unfamiliar stimuli. Finally, task-sets seem to reduce the congruency effect, suggesting perhaps that one of the roles of using tasks, as Dreisbach and Haider (2008, 2009) suggested, is to reduce interference. This suggests that task-cueing paradigms under Tasks instructions may, among other things, be measuring the task-sets' ability to reduce interference when contrasted to performance under CSR instructions.

However, it should be noted that although the majority of participants in the Tasks groups presented in this thesis did use the task-sets, a minority did not (Experiment 1: 8/24 participants failed to use the tasks under Tasks instructions, Experiment 2: 6/22, Experiment 3: 11/73). This means that it would be worthwhile for researchers using the task-cueing paradigm, at least with small stimulus sets, to de-brief their participants thoroughly to ascertain whether they were actually relying on the task-sets when performing in the experiment. Failure to exclude these participants would add noise to the data, as in general these participants showed effects more similar to those seen in the CSR conditions, that is a smaller switch cost that doesn't reduce with time to prepare, and a larger congruency effect.

That such a substantial minority did not use the task-sets in each experiment may be due to the relative efficacy of a CSR strategy when small number of cues and stimuli are used. In the majority of these experiments overall performance was roughly matched between the CSR and the Tasks groups (except in Experiment 2 where the Tasks group had a significant advantage, as the CSR group were learning by trial and error). Where either approach is on average equally advantageous to use, for some participants one may appeal more than the other. Experimenters wanting to use the task-cueing paradigm to explore task-set control would be wise to make the CSR strategy an inefficient option, for example by using larger stimulus sets.

7.5 Why use task-sets if you can perform without them?

Using task-sets does convey some advantages over not using task-sets. As suggested above, the smaller congruency effect when using task-sets suggests that their use may help ameliorate interference from the stimulus' response in the other task. This key advantage of using task-sets has been examined by Dreisbach and Haider (2008, 2009) who showed that when using task-sets, participants were less vulnerable to interference.

Additionally the transfer tests in the Monsell Experiment and Experiments 1&2 show that task-sets also allow easier transfer to novel stimuli. Intuitively this makes sense, as when using the task-sets one knows how to respond when the novel stimuli appear, whereas when using CSR mappings one has no choice but to learn the new responses by trial and error. Hence, it seems sensible to assume that using task-sets allows for easier generalization to novel stimuli. However, this result was in contrast to the results of Dreisbach, Goshke & Haider (2006) who found that using SR mappings allowed easier transfer. These two results can be reconciled if we consider two factors. Firstly, there was indeed a disadvantage for the Tasks conditions, but this was nevertheless smaller than that observed here for the CSR conditions and so Dreisbach et al's result of a disadvantage under task-set conditions is not in disagreement with my findings. Secondly, Dreisbach et al used univalent stimuli, i.e. stimuli which only permit a response under one task-set, and so it is just a case of learning the stimulus response associations. Hence, it was much easier to learn the correct response to these stimuli than for the bivalent stimuli (and particularly for the incongruent stimuli) used in this thesis. So their experiment may not have picked up the cost of learning the responses afresh as these responses were so quick to learn.

Another key advantage is that using task-sets gives is the ability to communicate information about how to perform efficiently and effectively. This was shown in several ways. Firstly it is worth considering that it takes far fewer words to outline how to respond to every stimulus presented when specifying a task-set stratgey. This makes communication much easier and more reliable as it is easier to impart two task-sets rather than run through a list of mappings.

It is also clearly a more efficient and effective strategy than just allowing the participant to learn by trial and error. This is evident in the difference in the initial error rates in Experiment 2. One of the main advantages of being able to communicate via language is that it removes the need to directly experience errors to know how to respond. This has obvious advantages.

7.6 How do participants in the CSR condition learn?

I've suggested in this thesis that the pattern of performance under CSR conditions can be captured by associative models. Chapter 3 showed that a static version of APECS could produce the same qualitative pattern of data as that observed in the CSR conditions across experiments, whilst Chapter 6 suggested AMAN and a real-time version of APECS would also be able to account for the main effects seen in Experiment 2. That is, both associative models, when trained on the task-cueing paradigm used in Experiment 2, showed small but reliable switch costs and large congruency effects which varied in size with training due to the differential ease of learning about congruent and incongruent stimuli. These findings mean that is possible for an association formation model to explain the performance under CSR instructions. However, it is worth noting that in their current incarnations (and variants examined in Chapter 6) neither APECS nor AMAN was able to fully account for all the data presented in this thesis.

Experiment 3 provided probably the most clear-cut evidence that the processes involved under CSR conditions were different in kind to those used under Tasks conditions. It provided this evidence in several ways. Firstly, the state-trace analysis, showed that when participants were using task-sets the function describing the relationship between switch and repeat RTs, as well as the function describing the relationship between incongruent and congruent RTs, differed significantly from the functions obtained when participants were using a CSR approach[15]. This suggests that the processes involved under these two instructional sets are different. It also indicated that participants were able to change from one type of process to another, or to emphasize one type of process rather than the other. Secondly, the data showed that performance altered when participants changed from a CSR approach to using task-sets, in a manner that was consistent with the notion that the participant had just begun the experiment under these instructions. By contrast, when participants changed from using Tasks to using the CSR approach, their performance was immediately at about the same level as participants who had consistently used the CSR approach up to that point. This suggests that participants had been able to learn the CSR mappings whilst using task-sets, but had not been able to learn about using task-sets whilst learning the CSR mappings. The result is also consistent with the idea that the processes involved in

---

[15] Although not presented in this thesis for reasons of space and to avoid undue repetition, the other experiments reported here also gave distinct functions on a state-trace plot for the Tasks and CSR conditions.

learning the CSR mappings may be automatic in nature, as associative processes are often considered to be.

Although these findings all suggest that automatic, associative processes could be behind performance in the CSR condition, they do not necessarily prove it. It remains possible that these are coincidences, and that there is no need to postulate associative processes, and that an entirely rule-based explanation can account for the data (Mitchell et al, 2009). But how viable is an entirely rule-based approach? In order to evaluate this possibility I examined what participants in the CSR condition said they were doing.

7.6.1 Were the participants in the CSR group using rules?

At the end of all of Experiments 1-3 participants in both the CSR and Tasks group were asked how they were "remembering which response to make to a given number". Participants in the Tasks groups were only included if their answer to this question was that they were using the task-sets as instructed (they were also asked the more explicit question of 'were they using the tasks' if this answer was ambiguous). However, participants in the CSR groups were included if they did *not* say they had used/induced the task-sets (after this question they were also asked if they had noticed the task-sets in order to check this). This means that, although participants' in the Tasks group's answers to this question were constrained to be relatively homogeneous, those of the CSR group were permitted to vary somewhat. A content analysis was done on the answers given by the CSR groups to ascertain whether or not there were any consistencies.

7.6.1.1 Method

The answers from Experiments 1-3 were coded. The answers were coded on treatment of congruent trials, linkage of cues and treatment of incongruent trials.

7.6.1.2 Results

For the congruent trials 76.5% (62/81) of participants reported the mappings for these trials separately, for example, they explicitly mentioned 1 was always left and 8 was always right. In the experiments which used more than one cue per task (excluding Experiment 1) 47.8% (31/65) reported noticing that two of the cues went together for the incongruent trials.

For the incongruent trials the responses were categorized by whether the participant named all four mappings, named two mappings and then said the others were opposites (or equivalent language) or named one mapping and then explained how the others were opposites. In total 44.4% (36/81) of participants mentioned one or more mapping: 3.7% (3/81) mentioned all 4 mappings, 29.6% (24/81) mentioned two mappings and 11.1% (9/81) mentioned just one mapping.

7.6.1.3 Discussion

The modal response for the CSR group to the question "how were you remembering which responses to make to a given number?" was for a participant to name the congruent trials separately, link the two cues for a given "task" together (if there was more than one verbal label needed for the cues) and then name two of the incongruent trials and mention how the others had the opposite responses, e.g. on blue/green 7 was left and 4 was right and on red/yellow the opposite responses are required. This finding may present problems for an associative account of performance, as it suggests that instead of simply learning the associations participants have instead induced and used their own set of "mini rules", which are nevertheless distinct from the task-set rules given in the Tasks condition. The high degree of consistency may lead one to wonder if this is indeed a better explanation of the effects found.

How good an explanation of the patterns of behaviour found in the CSR conditions is the assumption that participants are using these mini rules? The use of such rules could easily account for the large congruency effects seen. The rules for the congruent stimuli are easier to induce and to use, whilst the rules for the incongruent stimuli takes time to spot (they require more hypothesis testing) and then are much more cumbersome to use as they can involve working out the response by means of the rule of opposites. The extra inferences required would mean that it would take more time to respond on an incongruent trial than a congruent trial. It would also mean that this response may not be as accurate as the responses seen with the congruent trials (using a rules of opposite would be much more prone to errors than using a simple rule of response) leading to the same effect in the errors.

This account may also be able explain the small but reliable switch cost that is present when participants perform the experiment in this way. In the case of the incongruent trials the most recently used belief may be the easiest to bring to mind. If the two colours assigned to each task are collapsed into one belief (i.e. if the colour is blue or red and the number is 7 press left), this belief will be easier to access and use. It is the case that a mini-rule repeat can only occur on a repeat trial and not a switch trial, and this could be driving a small switch cost if, as task-switching suggests, there is a cost in changing rules. However, on average the difference here would be much smaller than the difference seen between the congruent and incongruent trials, as rule switches also occur on task repeats.

Additionally, this account can accommodate the lack of reduction in switch cost with preparation almost as well as an associative account. In the majority of experiments identification and application of the rule must await the stimulus. However, one could potentially use rules to prepare for the incongruent stimuli, as one can rehearse the appropriate responses for the incongruent stimuli when the colour is presented. This is advantageous as it reduces the number of possible rules in play. However, as preparation is thought to be a voluntary process and here incongruent rehearsal may hinder performance on congruent stimuli, preparation may not universally occur.

Although conscious mini-rules may be a plausible account of the performance in the CSR condition, there remains the issue of whether it is the best explanation. Where do these rules come from? It was noted in the Chapter 1 in Table 1.1 that although the associative system is generally thought to learn unconsciously, the product of the system is sometimes thought to be subject to conscious awareness. That is, when the learning in an associative system has reached a certain level then it can pass to a "conscious" system. In many ways APECS provides a good explanation of where these rules come from, as its hidden representations group the same rules that people tend to group together. It learns that the congruent stimuli always have the same response and for the incongruent stimuli it depends on the colour. So it may be that these rules are just a post hoc verbalization based on monitoring the activations of the input, hidden and output units for the mappings it experiences. So when given cues, stimuli and outputs use the same hidden units then these may be chunked together by a rule based system. It is worth considering that inducing the task-sets themselves would also be another valid rule-based approach but that this is much less common than the rules described above – perhaps because these rules require the use of reasoning using previously learnt knowledge (parity and magnitude) rather than the interpretation of the product of an associative system. So the mini rule hypothesis fails to establish, at present, why these particular rules are induced, whereas the associative account provides a motive behind the rule selection and grouping.

The mini rule hypothesis would also find it difficult to account for effects of changing the structural design of the mappings in Experiments 4 and 5. This is an interesting manipulation from the point of view of a rule-based hypothesis. In order to equate the total possible number of mappings the Separate condition had twice as many stimuli. This means it does not necessarily equate the total number of mini-rules required. In fact the Separate condition potentially requires more rules. This is because there are more congruent stimuli, 4 instead of 2, so instead of just learning "$" is always "right" and "@" is always "left" one would also have to learn that "&" is always "right" and "%" is always "left". It may be possible for these rules to be chunked together, as they are by APECS, as "$" and "&" are always "right" and "@" and "%" are always "left". However, it is worth noting that although this chunking can occur only 6/24 participants in the Separate condition in Experiment 5 named all of the congruent stimuli, but an additional 10 participants recalled just one or more of the congruent mappings (leaving 8 participants who were unable to articulate the congruent mappings)[16]. This is not in disagreement with an associative account of rule induction as in the Separate condition the congruent trials have had fewer pairings (60 trials) than in the Normal condition (120 trials).

It is also worth noting that for the incongruent stimuli only 7 participants in the Separate condition were able to explain how to respond to all the pairings, whereas 13 participants were able to outline a method in the Normal condition. These findings suggest that the rules are harder to articulate/induce in the Separate condition. However, as predicted by an associative model there was no main effect of condition, suggesting that participants were performing equally well in both conditions. So if performance could only be explained in terms of rules (Mitchell et al, 2010) there would have to be an explanation of why there was no difference between the two conditions: one where rules are commonly articulated and one where they were not.

One could argue that conducting a post experiment interview is not the most sensitive way of knowing if participants were using rules, as they may have forgotten the rules or have difficulty explaining them (Shanks & St John, 1994). However, the test is equally insensitive in both conditions, and in one condition they can articulate the rules and in the other they cannot. It may be worth running more sensitive tests on the

---

[16] One might worry that the use of symbols inhibited participant's ability to articulate a rule, however, 18/24 participants in the Normal condition were able to articulate both congruent mappings, with a further 5 naming at least one congruent mapping. The remaining participant named the incongruent mappings but failed to spontaneously mention the congruent mapping. In light of this, it does not seem reasonable to suggest that the use of symbols as opposed to numbers led to the Separate condition's inability to outline the rules.

explicit knowledge of the participants, perhaps by explicitly asking participants if they noticed the mappings as in Harris & Livesey (2008).

7.7 Do we need/ is it a good idea to posit both associative and cognitive control processes?

The above discussion raises the possibility that we do not strictly need to posit both types of processes to be able to account for the data in this thesis. It is possible to come up with a purely cognitive explanation of all the effects gained. However, is this a good explanation? At present, no. This explanation is entirely post-hoc, and based on interviews of participants at the end of experiments. In order to examine whether it is a good explanation one needs to examine its predictions. Firstly, one could examine what would happen if participants were instructed to use the rules participants verbalise at the end of the experiment. This is a tricky manipulation to implement, as the rules are difficult to communicate, and it is hard to verify whether participants are using them, so a result where performance is the same as the CSR condition would be difficult to interpret. Other manipulations that one would expect to disadvantage performance mediated by this mini-rules, such as the ease of verbalization of the rules (by using non-verbalisable symbols as cues and stimuli), concurrent load or a different transfer test where the rule structure no longer applied could also be considered.

The mini-rule account of CSR performance already struggles to explain all the data reported in this thesis, especially the results of Experiment 3 which suggest different processes are involved under Tasks and CSR conditions, and that those involved under CSR conditions are automatic in nature and so different in kind to those deployed under Tasks conditions. Although one could suggest that participants are concurrently inducing the mini rules whilst using a Tasks strategy and that this accounts for the performance of the Tasks-CSR group when they switch to CSR instructions, why would this happen (or why would participants be less honest) in this condition and not in the CSR-Tasks condition?[17]. There would need to be some explanation of why the mini rules are easier to induce whilst using task-sets than the task-sets are whilst

---

[17] Under both conditions participants who explicitly said they did this and so changed their strategy were excluded.

using mini rules, especially as task-sets rules are relatively simple compared to the family of mini-rules that has to be induced to solve the same problem.

In conclusion, the findings reported in this thesis suggest that taking an explanatory stance that involves positing both associative and cognitive processes seems the better explanation. It may not be the only explanation possible, but for the most part the results seem to sit most easily within this framework, and this framework was used to successfully make predictions about how changing the paradigm would affect behaviour.

7.8 How might a cognitive and associative system interact?

Although this thesis did not experimentally address how the two systems interact it is interesting to briefly consider some of the ways in which this interaction might take place. There are three main points in the stream of processing where the two systems could be thought to interact: at the point of stimulus encoding; the way in which the stimulus is interpreted (how the information is mentally represented); and at the point of deciding a response. The control system might interact with the associative system at all, two or just one of these stages. First consider the stimulus encoding stage. The associative system can only make links between the events that are experienced, i.e. what the participant senses, so at a very basic level there is an interaction at this stage, in that the control system has some impact on what is experienced, e.g. whether the participant is paying attention to the screen.

However, at a higher level of the representational encoding of the stimulus it seems more difficult to know how the two systems might split. Is it that the associative system takes in all the information about the external world or is it more selective in the information it accumulates? Research on attention might hold the answer, but attention is an ill-defined concept that might cover multiple mechanisms. Associative learning theorists have long suggested a role for attention in associative learning theories (Mackintosh, 1975, Pearce & Hall, 1980). The attentional mechanisms in these theories have both proposed algorithms to allocate attention based on stimuli's previous associative histories. This may be a different mechanism from being verbally instructed to pay attention to an aspect of a stimulus, and as such may lead to different aspects of a stimulus being attended to by the associative and controlled systems. Alternatively there could just be different drivers of one attentional system. It would therefore be worthwhile considering whether the stimuli are encoded differently under Task and

CSR conditions, as might be suggested by Dreisbach and Haider (2008, 2009)'s experiments.

Secondly, consider the representation stage. One possibility suggested in this thesis has been that at this stage separate processes are involved. In the CSR condition the associative system forms its links from the cue-stimulus to a response, whilst when task-sets are used this requires going from the cue to the task and then from the stimulus to the memory of how the semantic label for that stimulus leads to a response. These separate, relatively independent routes would be in agreement with the results of Experiment 3, where there was evidence that CSR learning took place in the background whilst participants were using the task-sets. However, this is not necessarily in agreement with the ideas presented in other dual theories. Logan & Gordon (2001) in their ECTVA model suggest that this is precisely the point where the two systems interact. Theories of this type hypothesise that the role of the control system is to bias the automatic system. This approach would have difficulty with the results/interpretation given of the Tasks-CSR condition in Experiment 3, as there would need to be some explanation of why performance in the Tasks-CSR condition was as if they had been under CSR conditions throughout, i.e. why there was evidence to suggest the CSR mappings existed in the same form as they existed under CSR conditions.

Finally it could simply be that the two systems are entirely independent, simply providing two separate answers which compete to produce a response. This notion of independence and interaction is frequently given to these systems in popular science books, e.g. Nudge (Thaler & Sunstein, 2009). Both systems could independently encode the stimulus and interpret the stimulus to a response, and then the system with the most 'confidence' in the response could then be the one to govern performance, as in the COVIS model of categorisation (Ashby et al, 2011). The quest to understand how the two systems interact would then be one of understanding what factors drive greater 'confidence' in one system over the other.

In reality the answer is probably much more complex than these hypotheses might suggest. However, considering the ways in which the systems may or may not interact at each of these stages may help to understand the nature of their interaction.

7.9 Further research

7.9.1 Testing assumptions

Throughout this thesis there have been some claims made which would benefit from further experimental testing. For example, the claim was made in Chapters 1 & 2 that animals, specifically pigeons, as they are thought to learn in a purely associative manner, should behave more like the CSR condition than the Tasks condition. This could fairly straightforwardly be verified by training pigeons on the task in the same manner as the CSR condition. However, it may be rather challenging for the pigeon to learn all the discriminations concurrently, and so an incremental training method akin to Dreisbach et al (2006; 2007) method of introducing pairs of stimuli at a time or Burdyn and Thomas's (1984) and Stevens' (2011) method of training one task first and then the other, before intermixing them might be more suitable. Such an experiment could add further weight to the associative analysis of performance in the CSR condition given in this thesis.

The claim was made in Chapter 2 and in this chapter that using additional stimuli would prevent/discourage participants from adopting a CSR approach when they were instructed to use task-sets. One could run the Tasks conditions with more stimuli, and check using a post-experiment interview whether participants did in fact use the task-sets. One would expect that as the number of stimuli increased, fewer participants in the Tasks condition would adopt a CSR strategy, as it would become a less desirable and more demanding strategy. One might also expect that, as the number of stimuli increased, more participants in a CSR condition would spontaneously induce the task-sets, as there would be more information available to serve as a basis for doing so, and inducing the task-sets would become a more desirable strategy as it would reduce memory load. Manipulations of this kind may help clarify the conditions under which a task-set based strategy is advantageous.

In Chapter 5 it was noted that the static APECS model predicts a smaller switch cost with one rather than two cues per task under the CSR instructions used in my experiments. This is a counter-intuitive prediction of the model, but it may be a tricky experiment to run because the switch cost in the CSR condition is generally only just reliable. Nevertheless, it might be worthwhile to try and establish whether or not people do in fact show a smaller a switch cost with only one cue per task, because this manipulation would affect part of the model based explanation of the "switch cost" given in Chapter 5 and so could be used to further unpack whether this explanation has any merit.

The other assumption made in Chapter 3 was that the precise nature of the cues and stimuli used does not matter. Whilst this may be true for stimuli that would be coded discretely, i.e., where it is reasonable to assume that each stimulus and cue is represented by distinct input units, it may not always be the case. In these experiments the task-sets used were semantic classifications made on digits. However, in task-cueing paradigms perceptual classifications are sometimes used e.g. classification of the colour and the shape of the stimulus, with a range of colours and shapes (e.g. Stevens, 2011). In experiments such as this it may not be reasonable to assume that each stimulus is coded uniquely, as one might assume that two stimuli in the same response category would share one or more input units, i.e. the more similar the stimuli the more their representations would overlap. In the case given this might involve changing the input representation of the problem to WAC-L, XAC-L, YAC-L, ZAC-L; WAD-L, XAD-L, YAD-R, ZAD-R; WBC-R, XBC-R, YBC-L, ZBC-L; WBD-R, XBD-R, YBD-R, ZBD-R; where W, X, Y and Z are cues and A/B is one feature of the stimulus, e.g. colour and C/D is the other, e.g. shape. This changes the nature of the discriminations to be learnt: no individual feature is now completely predictive of the outcome, but each stimulus feature, A,B,C, and D is 75% predictive of a response as 25% of the time it appears in the other task where the other response is required. This might be interesting to explore as it changes the nature of the problem being learnt by the associative model. The input representations created by using perceptual tasks rather than semantic tasks may cause the associative system to adopt more task-based representations at a hidden unit level, where the task relevant property and its task-cue enters into an association with the response, e.g. WA-L, XA-L, YC-L, ZC-L, WB-R, XB-R, YD-R, ZD-R. An associative model may pick up on these regularities as these compounds are all completely and uniquely predictive of the outcome. In a sense these mappings, which may be picked out by APECS, would be the similar to the mappings present in Brown et al’s (2007) model of task-set control. This manipulation could be tricky to investigate behaviourally, as there may simply be more task-set induction in the CSR condition with this type of stimulus, which would lead to a high rejection rate for participants in this condition.

7.9.2 Other ideas to consider

It is in the nature of APECS for it to create (relatively) discrete representations in the hidden layer. In Chapters 3 and 5 it was suggested that the model generally used one hidden unit per incongruent stimulus per "task" and one hidden unit per congruent stimulus. However, it was also noted that occasionally the congruent stimuli would latch onto the representations created by the incongruent stimuli for the task appropriate representation. If this occurred for all congruent and incongruent stimuli then the model would use just four hidden units, which would carry the mappings in a way very similar to the models outlined by Gilbert and Shallice (2002) and Brown et al (2007). In other words it would create task style classifications at the hidden layer, e.g. have an odd, even, high and low hidden unit. Therefore, it may be worth considering whether APECS could develop a solution closer to the structure proposed by Gilbert and Shallice (2002) and Brown et al (2007). This could perhaps be achieved by further limiting the number of hidden units (perhaps even as low as 4), changing the training order to perhaps train one task first, or using more stimuli so that the cue has a larger influence over the hidden units as in the model of the Separate condition in Chapter 5. It was noted above in the discussion of perceptual and semantic tasks that another obvious solution - encoding the stimuli perceptually (in other words by their task relevant characteristics) - might encourage more task-based representations. It may then be worth considering whether a more substantial modification of the forms outlined in the model by Gilbert & Shallice (2002) or Brown et al (2007) is necessary to generate performance similar to the Tasks condition, even when some task style representations have been created at the hidden layer. For example, is it necessary to have fixed weights, to mediate the cues through a task demand unit or to have a link between the tasks and the stimuli, which both models have? That is, what are the necessary additions to an APECS model which are needed to change its performance from matching the CSR condition to matching the Tasks condition.

It might be interesting to explore how changing the training order would affect the nature of the associations formed. For example, one could consider training the incongruent stimuli first before introducing the congruent stimuli. One might hypothesize that for APECS in this case the congruent stimuli would be more likely to latch onto the existing incongruent stimuli's hidden units (as they have already developed), and so be treated more like incongruent stimuli. For AMAN, by contrast, this training order would have little or no effect, as the links involving congruent stimuli would be learnt separately in both cases. Hence this manipulation would help illuminate which model was a better model of the CSR condition. In addition it would be of interest to contrast the CSR and Tasks condition. Experiment 3 suggested that the processes involved in the CSR condition run automatically in the background in the Tasks condition. This may mean that part of the congruency effect seen in the Tasks condition is driven by the same underlying conflict as in the CSR condition (i.e. the tasks simply ameliorate the effect seen in the CSR condition). If this is the case then if the congruency effect has been eliminated in the CSR condition, one might expect there to be an impact on the congruency effect in the Tasks condition.

It would also be worthwhile considering the effect of manipulating some of the other factors commonly implicated in the operation of associative and controlled processes in this paradigm. For example, cognitive control is thought to be resource heavy. Hence, adding an additional task is thought to disrupt controlled processing more than associative/automatic processing (e.g. Waldron and Ashby, 2001). Performance in the Tasks condition should thus be disrupted more than in the CSR condition when an extra task is added. In addition, as noted above, the findings of Experiment 3 suggested that CSR learning occurs in the background whilst participants are using task-sets, and so adding in a dual task may change the performance in the Tasks condition to be more like that in the CSR condition. This would have important implications for our theoretical understanding of the ways in which associative and cognitive systems interact, as it would support the findings of Experiment 3, which implied a degree of independence in the way in which the two systems function.

Many of the manipulations used within the task-cueing literature may be of interest when run under CSR conditions. For example, it might be worth running a manipulation similar to Kiesel, Wendt & Peters (2007), who swapped one of the tasks for an orthogonal task part way through, i.e. swap one of the tasks for an inner/outer discrimination, where 1,2,8,9 (outer) have one response and 3,4,6,7 (inner) have the other. Changing the task would alter the congruency level of some of the stimuli, whilst keeping others the same, i.e. parts of the component discrimination would become part of the biconditional discrimination and vice versa. The interest would be in seeing how elemental and configural models handle learning of the additional mappings for the same set of stimuli, whether this would present more of a challenge than introducing new stimuli, and whether the results would correspond to the behavioural data. Would it be the case that it is easier to learn a new response to a previously incongruent stimulus or a previously congruent stimulus? Another task-cueing manipulation that might be of interest would be to examine the effect of reversing the cue and stimulus order (as per Ruge & Braver, 2007). With this manipulation the associative problem posed by the incongruent stimuli would, in some sense, be the same. But for the congruent stimuli performance should be improved as participants could now fully prepare their response and there might well be an effect of CSI reflecting this preparation. It may also be worth not maintaining the cue throughout the CSI but presenting it and then leaving a gap, as is the case in many task-cueing paradigms (e.g. Stevens, 2011). This is known as trace conditioning, as opposed to delay conditioning in the Pavlovian literature. It is the trace conditioning procedure that is thought to encourage the cue to become an occasion setter, and so this might help establish whether occasion setting is an important component of task switching under CSR conditions, as AMAN might suggest.

In the course of some final, post-hoc analyses, directed towards establishing why an associative model might generate a “switch cost”, I discovered that there is a very slight difference in the average recency of re-occurrence of a given cue+stimulus pair (i.e. the average number of trials before that cue+stimulus pairing occurs again) between a switch and a repeat trial. On average, a given cue+stimulus pair will have occurred more recently on a repeat trial than a switch trial. This was true for all the experiments in this thesis, see Table 7.1. If a given cue+stimulus pair has occurred more recently then there may be a response advantage as there has been fewer trials in between and so less trials on which new learning can interfere with the given cue+stimulus pair’s mapping. Hence if a given cue+stimulus pairing has occurred more recently it may have an advantage. If a given cue+stimulus pair occurs on average more recently on a repeat than a switch trial, this could lead to an apparent “switch cost” as there may be an advantage on repeat trials.

Sequence types	Repeat Recency	Switch Recency	F value
p(switch) = 0.5 (Monsell Experiment, and Experiments 1 &2)	16.1	17.7	F(1,31)=29.2, p<0.001
p(switch)=0.33 (Experiment 3)	19.9	20.9	F(1,31)=14.8, p<0.001
p(switch)=0.33 with cue repeats (Normal condition Experiment 4 &5)	15.1	18.0	F(1,31)=189, p<0.001
p(switch)=0.33 with cue repeats (Normal condition Experiment 4 &5) – no direct cue + stimulus repetitions	16.4	18.0	F(1,31)= 48.7, p<0.001
p(switch)=0.33 with cue repeats (Separate condition Experiment 4 &5)	15.0	18.1	F(1,31)= 109, p<0.001
p(switch)=0.33 with cue repeats (Separate condition Experiment 4 &5) – no direct cue + stimulus repetitions	16.4	18.1	F(1,31)= 35.5, p<0.001

Table 7.1 shows the differences in the recency (mean no. of trials between occurrences of a given cue+stimulus pair) divided up by whether the trial in question is a switch or a repeat trial. For each sequence type a new sample of 32 sequences were generated and analysed, using the same code as was used to generate the sequences experimentally.

This recency difference cannot be the necessary and sufficient driver of the switch cost seen in the CSR condition or the Tasks condition. Evidence for this is that there is a recency difference in the Separate condition in which there is no evidence of a switch cost, and the recency distribution is similar in the Normal and Separate conditions[18] but there is a difference in the size of the switch cost. However, this does not mean that it does not play some sort of role. It would be worth investigating what sort of influence

---

[18] There is no difference because there are the same number of total pairings in each condition

it might have, if any, on the switch cost. In investigations into this difference it would be also worth considering the nature of the interference between repetitions of a given cue+stimulus pair, as it is an established principle that the more similar the interfering event (without being identical) the more it interferes. As the cue and the stimulus can repeat in different conjunctions between occurrences of a given cue+stimulus aspects of the intervening trials may be very similar to the to be remembered mapping in such a way as to interfere. Hence, it may be that any effect of recency is counteracted by the fact that on a repeat trial the preceding trial (at least) is more likely to have been more similar to the current trial, as by the nature of a repeat trial the cue used is more likely to have occurred more recently. This would be especially true with a low probability of switching, where long runs of repeats are more likely to occur. A brief initial investigation suggested that the switch probability may affect whether recency has an impact, as there was some evidence of a trend related to recency in the RT and error data of Experiment 2 (p(switch=0.5)), but not in the data of Experiment 3 (p(switch=0.33)). However, given the speculative nature of this analysis, due to the low numbers of participants and low number of trials for each given value of recency, this analysis can only be taken as an initial investigation and this question would require closer consideration using a larger sample to really establish whether this recency difference contributes to the switch cost. It's also worth noting that Monte Carlo simulations of sequences that are completely randomly generated (rather than constrained as the sequences in my experiments were) still produce this recency effect in the sequences, even with just the constraint of no direct cue+stimulus repetitions. Thus, whether or not this effect contributes to the switch cost, it is not an artefact of the particular methods used for sequence production in the experiments reported in this thesis.

This thesis has somewhat skirted around the issue of attention. However, it might be worthwhile attempting to address this issue. Participants in the Tasks conditions tended to say that they focused more on the cue first, which makes sense, as the cue tells them which task-set to adopt, whereas participants in the CSR condition tended to say they focused more on the stimulus, as from their point of view this was generally more informative. To verify if this was indeed the case an eye tracking experiment (used as a measure of overt attention, Wills, Lavric, Croft & Hodgson, 2007) could be run where the cue and the stimulus were presented in discrete locations (perhaps one above the other) as in Longman, Lavric & Monsell (in press). The test case would be to see whether participants in the different conditions considered the cue or the stimulus first when they were both presented simultaneously. And whether, in the case of the congruent stimuli, those participants in the CSR condition do not use the cue at all whereas those participants in the Tasks condition do.

It might also be worth considering whether participants would naturally switch to a CSR approach after extended training. There may come a point where the performance of the automatic/associative system is faster and more accurate than the controlled system and so comes to dominate responding[19]. This may happen after extended training when the associative system has been able to learn the correct responses for the incongruent stimuli. This could be tested by overtraining a Tasks group, and ascertaining whether their performance eventually became equivalent to that of a similarly treated CSR group. This is an interesting concept to consider given that a task-cueing paradigm is frequently included in brain-training packages, and the results of such an experiment may suggest that with excessive training the paradigm might not be training the same processes as it was initially.

7.10 Concluding Remarks

In conclusion, this thesis has demonstrated that approaching a task-cueing paradigm using CSR instructions leads to a different pattern of responding than using the standard task-set approach. Typically participants in a task-cueing paradigm are found to have large switch costs, which they can reduce if they have time to prepare the task-sets, and relatively small congruency effects. But when participants approach this paradigm using CSR instructions, in contrast to the effects observed in the standard Tasks condition, they show a smaller switch cost, which does not reduce with time to prepare, and a larger congruency effect which changes over time as participants learn the responses to the congruent and incongruent stimuli at different rates. This pattern of results has been found consistently across all the experiments presented in this thesis and is broadly consistent with the simulation results produced by associative models of learning. The differences between these conditions may have important implications for the interpretation of performance in a standard task-cueing paradigm. They suggest

---

[19] This may not happen if the automatisation of the cue-task link occurs at the same rate as the acquisition of the CSR mappings.

that participants in the standard task-cueing paradigm are not just learning the cue+stimulus response combinations, contradicting the idea put forward by Logan and colleagues. My conclusion is that task-cueing paradigms under Tasks instructions are indeed fulfilling their *raison d'être,* people's ability to switch between task-sets. In addition, this thesis has explored the nature of the processes involved in learning under CSR instructions, and suggested that these processes are more automatic/associative in nature than those involved under Tasks instructions, supporting the utility of using this distinction. Finally, this thesis has considered which aspects of the design of the task-cueing paradigm lead to the "switch cost" seen in the CSR condition, and discovered that it was related to the use of two cues per task which were shared by the congruent and incongruent stimuli.

This thesis has attempted to combine ideas from two literatures to consider what can be gained from combining the approaches used in each. The outcome is that the associative learning literature has gained an extra effect (the switch cost) that lay undiscovered in the models in current use, and perhaps also the insight that considering the specifics of trial-by-trial transitions may help to differentiate between modelling accounts. Task-switching researchers will need to take note of the plausible case for associative processes operating within the task-cueing paradigm made in this thesis and the nature of those associations. More importantly, I have provided evidence that the *best* theoretical explanation of my results is to interpret them within a dual system framework, in which there are both automatic/associative and controlled/intentional processes.

References

Adams, D. (1988). *The Long Dark Teatime of the Soul*, London: Pan Books

Allport, D.A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta, & M. Moscovitch (Eds.), *Conscious and Nonconscious Information Processing: Attention and Performance XV* (pp. 421-452). Cambridge, MA: MIT Press.

Altmann, E.M. (2004). The preparation effect in task switching: Carryover of SOA. *Memory and Cognition, 32,* 153-163.

Altmann, E.M., & Gray, W.D. (2008). An integrated model of cognitive control in task switching. *Psychological Review, 115*, 602– 639.

Anderson, J. (1982) Acquisition of cognitive skill. *Psychological Review, 89*, 369-406.

Arrington, C.M. & Logan G.D. (2004) Episodic and semantic components of the compound-stimulus strategy in the explicit task-cuing procedure. *Memory & Cognition*, *32*, 965-978.

Arrington, C.M., Logan, G.D. & Schneider, D.W. (2007). Separating cue encoding from target processing in the explicit task-cuing procedure: Are there “true” task switch effects? *Journal of Experimental Psychology: Learning, Memory and Cognition, 33,* 484-502.

Ashby, F.G., Paul, E.J., & Maddox, W.T. (2011) COVIS. In E. M. Pothos & A.J. Wills (Eds.), *Formal approaches in categorization* (pp. 65-87). New York: Cambridge University Press.

Bamber, D. (1979). State-trace analysis: A method of testing simple theories of causation. *Journal of Mathematical Psychology, 19,* 137-181.

Barnes, J.M., & Underwood, B.J. (1959). “Fate” of first-list associations in transfer theory. *Journal of Experimental Psychology , 58,* 97–105.

Beckers, T., Miller, R.R., De Houwer, J. & Urushihara, K. (2006) Reasoning rats: Forward blocking in Pavlovian animal conditioning is sensitive to constraints of causal inference. *Journal of Experimental Psychology: General, 135*, 92 – 102

Berry, C.J., Shanks, D.R. & Henson, R.N.A. (2008) A single-system account of the relationship between priming, recognition and fluency. *Journal of Experimental Psychology: Learning Memory and Cognition, 34(1),* 97-111.

Bonardi, C. (1996) Transfer of occasion setting: The role of generalization decrement. *Animal Learning & Behavior, 24(3),* 277-289.

Bonardi, C. & Jennings, D. (2009) Learning about associations: Evidence for a hierarchical account of occasion setting. *Journal of Experimental Psychology: Animal Behavior Processes, 35(3),* 440-445.

Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., & Cohen, J.D. (2001). Conflict monitoring and cognitive control. *Psychology Review, 108*, 624 –652.

Botvinick, M.M, Plaut, D.C. (2002) Representing task context: insights from a connectionist model of action. *Psychological Research, 66*, 298–311

Bouton, M.E., and King, D.A. 1983. Contextual control of the extinction of conditioned fear: Tests for the associative value of the context. *Journal of Experimental Psychology Animal Behavioral Processes, 9,* 248-26

Bouton, M.E. & Swartzentruber, D. (1986) Analysis of the associative and occasion-setting properties of contexts participating in a Pavlovian discrimination. *Journal of Experimental Psychology: Animal Behavior Processes, 12(4),* 333-350.

Bowler, P.J. & Morus, I.R. (2005) *Making Modern Science: A Historical Survey*. University of Chicago Press, Chicago.

Brainard, D.H. (1997) The psychophysics toolbox. *Spatial Vision, 10,* 433-436

Brown, J.W., Reynolds, J.R., & Braver, T.S. (2007). A computational model of fractionated conflict– control mechanisms in task-switching. *Cognitive Psychology, 55*, 37– 85

Burdyn, L.E. & Thomas, R.K. (1984) Conditional discrimination with conceptual simultaneous and successive cues in the squirrel monkey. *Journal of Comparative Psychology, 98(4),* 405-413.

Busemeyer, J. R. (1982). Choice behavior in a sequential decision-making task. *Organizational Behavior and Human Performance, 29,* 175–207.

Christie, A. (1926). *The Murder of Roger Ackroyd,* London: William Collins and Sons

Cleeremans, A. & McClelland, J.L. (1991) Learning the structure of event sequences. *Journal of Experimental Psychology: General, 120(3),* 235-253.

Cohen, J.D.M, Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97,* 332-361.

Coutureau, E., Killcross, A.S., Good, M., Marshall, V.J., Ward-Robinson, J. & Honey, R.C. (2002). Acquired equivalence and distinctiveness of cues: II. Neural manipulations and their implications. *Journal of Experimental Psychology: Animal Behavior Processes, 28,* 388–396.

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 Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning and. Behavior, 37,* 1–20

Dickinson, A. (2009). What are association formation models? *Learning and. Behavior, 37,* 21–24.

Dickinson, A., & Burke, J. (1996). Within-compound associations mediate the retrospective revaluation of causality judgements. *Quarterly Journal of Experimental Psychology, 49B,* 60–80.

Dreisbach, G., Goschke ,T., & Haider H. (2006) Implicit task-sets in task switching? *Journal of Experimental Psychology. Learning, Memory, and Cognition, 32*,21-33.

Dreisbach, G., Goschke, T., & Haider, H. (2007) The role of task rules and stimulus-response mappings in the task switching paradigm. *Psychological Research; 71,* 383-92.

Dreisbach, G., & Haider, H. (2008) That's what task-sets are for: shielding against irrelevant information. *Psychological Research, 72,* 355-61.

Dreisbach, G., & Haider, H. (2009) How task representations guide attention: Further evidence for the shielding function of task-sets. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 477-486.

Dunbar, K., & MacLeod, C. (1984). A horse race of a different color: Stroop interference patterns with transformed words. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 622–639.

Dunn, J.C. (2008). The dimensionality of the remember–know task: A state-trace analysis. *Psychological Review, 115,* 426-446.

Elchlepp, H., Lavric, A. & Monsell, S. (in preparation).

Elman, J.L. (1990) Finding structure in time. *Cognitive Science, 14,* 179-211.

Evans, J.StB.T. (2003). In two minds: dual process accounts of reasoning. Trends in Cognitive. Science, 7, 454–59

Fitts, P.M. (1964) Perceptual-motor skill learning. In A. W. Melton (Ed.), *Categories of Human Learning.* New York: Academic Press, 1964.

Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A. & Noll, D.C. (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magnetic Resonance In Medicine, 33(5),* 636-647.

Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious*. New York: Viking

Gilbert, S.J., & Shallice, T. (2002). Task switching: A PDP model. *Cognitive Psychology, 44*, 297–337.

Gopher, D. (1996). Attention control: Explorations of the work of an executive controller. *Cognitive Brain Research, 5,* 23–38.

Goschke, T. (2000). Intentional reconfiguration and involuntary persistence in task-set switching. In S. Monsell & J. Driver (Eds.), *Control of Cognitive Processes: Attention and performance XVIII* (pp. 331-355). Cambridge, MA: MIT Press.

Gratton, G., Coles, M.G., & Dunchin, E. (1992). Optimizing the use of information: strategic control of activation of responses. *Journal of Experimental Psychology: General, 121*, 480-506.

Haddon, J.E. & Killcross, S. (2006) Prefrontal cortex lesions disrupt the contextual control of response conflict. *The Journal of Neuroscience, 26(11),* 2933-2940.

Harris, J.A., (2006). Elemental representations of stimuli in associative learning. *Psychological Review, 113,* 584–605.

Harris, J.A. & Livesey, E.J. (2008). Comparing patterning and biconditional discriminations in humans. *Journal of Experimental Psychology: Animal Behavior Processes, 34*, 144-154.

Harris, J.A. & Livesey, E.J. (2010) An attention-modulated associative network. *Learning & Behavior, 38(1),* 1-26.

Hebb, D.O. (1949). *The organization of behavior: A neuropsychological theory*. New York: Wiley.

Hodder, K.I., George, D.N., Killcross, A.S. & Honey, R.C. (2003) Representational blending in human conditional learning: Implications for associative theory. *The Quarterly Journal of Experimental Psychology, 56B(2),* 223-238.

Holland, P.C. (1986) Temporal determinants of occasion setting in feature-positive discriminations. *Animal Learning & Behavior, 14(2),* 111-120.

Holland, P.C. (1989) Occasion setting with simultaneous compounds in rats. *Journal of Experiment Psychology: Animal Behavior Processes, 15(3),* 183-193.

Hommel, B., Fischer, R., Colzato, L.S., van den Wildenberg, W.P.M. & Cellini, C. (2012) The effect of fMRI (noise) on cognitive control. *Journal of Experimental Psychology: Human Perception and Performance, 38(2),* 290-301.

Honey, R.C., Close, J., & Lin, T.E. (2010). Acquired distinctiveness and equivalence: A synthesis. In C.J. Mitchell & M.E. Le Pelley (Eds.), Attention and associative learning: From brain to behaviour: Oxford: Oxford University Press.

Honey, R.C., & Hall, G. (1989). Acquired equivalence and distinctiveness of cues. *Journal of Experimental Psychology*: *Animal Behavior Processes*, *16*, 178–184.

Honey, R.C., & Ward-Robinson, J. (2001). Transfer between contextual conditional discriminations: An examination of how stimulus conjunctions are represented. *Journal of Experimental Psychology*: *Animal Behavior Processes*, *27*, 196–205.

Honey, R.C., & Ward-Robinson, J. (2002). Acquired equivalence and distinctiveness of cues: I. Exploring a neural network approach. *Journal of Experimental Psychology*: *Animal Behavior Processes, 28,* 378–387.

Honey, R.C., & Watt, A. (1998). Acquired relational equivalence: Implications for the nature of associative structures. *Journal of Experimental Psychology*: *Animal Behavior Processes*, *24*, 325–334.

Honey, R.C., & Watt, A. (1999). Acquired relational equivalence between contexts and features. *Journal of Experimental Psychology*: *Animal Behavior Processes*, *25*, 324–333.

Iordanova, M.D., Killcross, A.S. & Honey, R.C (2007) Role of medial prefrontal cortex in acquired distinctiveness and equivalence of cues. *Behavioral Neuroscience, 121(6),* 1431-1436.

Jacoby, L.L. (1991) A process dissociation framework: separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*, 513-541.

Jamadar, S., Hughes, M., Fulham, W.R., Michie, P.T. & Karayanidis F. (2010), The spatial and temporal dynamics of anticipatory preparation and response inhibition in task-switching. *NeuroImage, 51,* 432-449.

Jennings, J.M., & Jacoby, L.L. (1993). Automatic versus intentional uses of memory: Aging, attention, and control. *Psychology and Aging, 8(2),* 283–293

Jersild, A.T. (1927). Mental set and shift. *Archives of Psychology, 89*, 1–81

Kahneman, D. (2011) *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.

Karbach, J., & Kray, J. (2009). How useful is executive control training? Age differences in near and far transfer of task-switching training. *Developmental Science, 12*, 978 –990.

Kehoe, E.J. & Bosenberg, K.G. (2002) Temporal discrimination using different feature-target intervals in classical conditioning of the rabbit's nictitating membrane response. *Animal Learning & Behavior, 30(3),* 208-216.

Kehoe, E.J., Horne, A.J. & Macrae, M. (1995) Learning to learn: Real-time features and a connectionist model. *Adaptive Behavior, 3,* 235-271.

Kiesel, A., Wendt, M., & Peters, A. (2007) Task switching: On the origin of response congruency effects. *Psychological Research, 71,* 117-125.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A.M. & Koch, I. (2010*)* Control and interference in task switching—A review. *Psychological Bulletin, 136,* 849-874.

Koch, I. (2001) Automatic and intentional activation of task-sets. *Journal of Experimental Psychology: Learning Memory and Cognition, 27,* 1474-1486.

Koch, I., &Allport, A. (2006) Cue-based preparation and stimulus-based priming of tasks in task switching. *Memory and Cognition, 34,* 433-444.

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,* 339–378.

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

Lashley, K.S. (1938) Conditional Reactions in the Rat. *The Journal of Psychology: Interdisciplinary and Applied, 6(2),* 311-324.

Lavric, A., Mizon, G.A. & Monsell, S. (2008) Neurophysiological signature of effective anticipatory task-set control: a task-switching investigation. *European Journal of Neuroscience, 28,* 1016-1029.

Le Pelley, M.E. & McLaren, I.P.L. (2001) Retrospective revaluation in humans: Learning or memory? *The Quarterly Journal of Experimental Psychology, 54B,* 311-352.

Le Pelley, M. E., Suret, M. B., & Beesley, T. (2009). Learned predictiveness effects in humans: A function of learning, performance, or both? *Journal of Experimental Psychology: Animal Behavior Processes, 35,* 312–327.

Lipton, P. (2004) *Inference to the best explanation*. London: Routledge.

Livesey, E.J., Thorwart, A. & Harris, J.A. (2011) Comparing positive and negative patterning in human learning. *The Quarterly Journal of Experimental Psychology, 64(12),*

Livesey, E.J., Thorwart, A., De Fina, N.L. & Harris, J.A. (2011) Comparing learned predictiveness effects within and across compound discriminations. *Journal of Experimental Psychology: Animal Behavior Processes, 37(4),* 446-465

Loftus, G.R., Oberg, M.A. & Dillon, A.M. (2004). Linear theory, dimensional theory, and the face-inversion effect. *Psychological Review, 111,* 835-865.

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,* 575-599.

Logan, G.D., &Bundesen, C (2004) Very clever homunculus: Compound stimulus strategies for the explicit task-cuing procedure. *Psychonomic Bulletin & Review, 11,* 832-840.

Logan, G.D., & Gordon, R.D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review, 108*, 393-434.

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, 978-994.

Longman, C.A., Lavric A. & Monsell S. (in press) More attention to attention? An eye-tracking investigation of selection of perceptual attributes during a task switch. *Journal of Experimental Psychology: Learning Memory & Behavior*

Mackintosh, N.J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcements. *Psychological Review, 82,* 276-298.

Mackintosh, N.J. (1988). Approaches to the study of animal intelligence. *British Journal of Psychology, 79,* 509-526.

Mackintosh, N.J. & Little, L. (1969) Intradimensional and extradimensional shift learning in pigeons. *Psychonomic Science, 14(1),* 5-6.

Marquis, J., Killcross, S. & Haddon, J.E. (2007) Inactivation of the prelimbic, but not infralimbic, prefrontal cortex impairs the contextual control of response conflict in rats. *European Journal of Neuroscience 25(2),* 559-566.

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,* 362-372.

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

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

Mayr, U. and Kliegl, R. (2000) Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology: Learning Memory and Cognition, 26,* 1124–1140

McClelland, D.E. & Rumelhart, J.L. (1986) *Parallel distributed processing: Exploration in the microstructure of cognition volume 1: Foundations* The Massachusetts Institute of Technology

McCloskey, M., & Cohen, N.J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In Bower, G.H. (Eds.) *The Psychology of Learning and Motivation Vol 24,* (pp 109–166).San Diego, CA: Academic Press, Inc.

McLaren, I.P.L. (1993). APECS: A solution to the sequential learning problem. Proceedings of the XVth Annual Convention of the Cognitive Science Society, 717–722.

McLaren, I.P.L. (1994). Representation development in associative systems. In J.A. Hogan & J.J. Bolhuis (Eds.), *Causal Mechanisms of Behavioural Development* (pp. 377–402). Cambridge, UK: Cambridge University Press.

McLaren, I.P.L. (2011) APECS An adaptively parameterized model of associative learning and memory. In E. Alonso & E. Mondragon (Eds). *Computational Neuroscience for Advancing Artificial Intelligence: Models, Methods and Applications.* (pp. 145-164). Hershey, PA: Medical Information Science Reference.

McLaren, I.P.L., Forrest, C.L. & McLaren R.P. (2012) Elemental Representations and Configural Mappings. *Learning and Behavior,*

McLaren, I.P.L., Green, R.E.A. & Mackintosh, N.J. (1994) Animal learning and the explicit/implicit distinction. In N.C. Ellis (Ed.), *Implicit and Explicit Learning of Languages*. London: Academic Press.

McLaren, I.P.L., Kaye, H., & Mackintosh, N.J. (1989). An associative theory of the representation of stimuli: applications to perceptual learning and latent inhibition. RGM Morris (Eds) In *Parallel Distributed Processing: Implications for Psychology and Neurobiology,* pp. 102–30. Oxford: Clarendon

McLaren, I.P.L., & Mackintosh, N.J. (2000). An elemental model of associative learning: I. Latent Inhibition and perceptual learning. *Animal Learning & Behavior, 28,* 211–246.

McLaren, I.P.L., & Mackintosh, N.J. (2002). Associative learning and elemental representation: II. Generalization and discrimination. *Animal Learning & Behavior, 30,* 177–200.

McClelland, D.E. & Rumelhart, J.L. (1985) Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General, 114,* 159-188.

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1423-1442.

Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. *Cognitive Psychology, 41,* 211-253.

Meiran, N., Levine, J., Meiran, N., & Henik, A. (2000). Task set switching in schizophrenia. Neuropsychology, 14, 471– 482.

Mitchell C.J., De Houwer J., Lovibond P.F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences, 32*, 183–246.

Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49–100.

Monsell, S. (2003) Task switching. *Trends in Cognitive Science, 7*, 134–140.

Monsell, S., Sumner, P., & Waters, H. (2003) Task-set reconfiguration with predictable and unpredictable task switches. *Memory & Cognition, 31,* 327-342.

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,* 493–516.

Monsell S & Mizon, G.A. (submitted) Stimulus-task associations and the task-switch cost.

Morell, J.R. & Holland, P.C. (1993) Summation and transfer of negative occasion setting. *Animal Learning & Behavior, 21(2),* 145-153.

Murphy, R.A., Mondragón, E. and Murphy, V.A. (2008) Rule learning by rats. *Science, 319,* 1849-1851.

Newell, B.R., & Dunn, J.C. (2008). Dimensions in data: Testing psychological models using state-trace analysis. *Trends in Cognitive Sciences, 12,* 285-290.

Newell, B.R., Dunn, J.C., & Kalish, M. (2010). The dimensionality of perceptual category learning: A state-trace analysis. *Memory & Cognition, 38,* 563-581.

Newell, B.R., Dunn, J.C., & Kalish, M. (2011). Systems of category learning: Fact or fantasy? In B.H. Ross (Ed.) *The Psychology of Learning & Motivation, 54*, 167-215.

Niemi, P., & Naatanen, R. (1981). Foreperiod and simple reaction time. *Psychological Bulletin, 89,* 133–162.

Nieuwenhuis, S. &Monsell, S. (2002) Residual costs in task switching: Testing the failure-to-engage hypothesis. *Psychonomic Bulletin & Review, 9,* 86-92.

Nisbett, R.E. & Wilson, T.D. (1977) Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84(3),* 231-259.

Nobre, A., Correa, A., &Coull, J.T. (2007). The hazards of time. *Current Opinion in Neurobiology, 17,* 465–470.

North, A.J., Maller, O. & Hughes, C. (1958) Conditional discrimination and stimulus patterning. *Journal of Comparative and Physiological Psychology, 51(6),* 711-715.

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

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

Pearce, J.M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review, 94,* 61–73.

Pearce, J.M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review, 101*, 587–607.

Pearce, J.M. & Hall, G. (1980). A model for Pavlovian conditioning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review, 87*, 532–552.

Pelli, D.G. (1997) The videotoolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision, 10(4),* 437-442.

Pratchett, T. (2001). *Thief of Time,* New York: Doubleday

Rescorla, R.A. (1985). Inhibition and facilitation. In R. R. Miller & N. E. Spear (Eds.), *Information processing in animals: Conditioned inhibition* (pp. 299-326). Hillsdale, NJ: Erlbaum.

Rescorla, R.A., Wagner, A.R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Black, A.H., Prokasy, W.F. (Eds.), *Classical Conditioning II: Current Theory and Research. Appleton-Century-Crofts*, New York, pp. 64–99.

Riopelle, A.J. & Copelan, E.L. (1954) Discrimination reversal to a sign. *Journal of Experimental Psychology, 48(2),* 143-145.

Roberts, A.C., Robbins, T.W. & Everitt, B.J. (1988) The effects of intradimensional and extradimensional shifts on visual discrimination learning in humans and non-human primates. *The Quarterly Journal of Experimental Psychology Section B: Comparative and Physiological Psychology, 40(4),* 321-341.

Rogers, R.D. & Monsell, S. (1995) Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124,* 297-231.

Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review, 110*, 88 –125.

Ross, R.Z., & Holland, P.C. (1981) Conditioning of simultaneous and serial feature positive discriminations. *Animal Learning & Behavior, 9,* 293-303.

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,* 1477-1491.

Ruge, H. & Braver, T.S. (2007) Neural mechanisms of cognitive control in cued task-switching: rules, representations, and preparation. In: Bunge SA, Wallis JD, editors. *Neuroscience of rule-guided behavior*. Oxford: Oxford University Press. pp. 255–282

Rumelhart, D.E., Hinton, G.E. & Williams, R.J. (1986). Learning representations by back-propagating errors. *Nature, 323*, 533 – 536.

Saavedra, M.A. (1975). Pavlovian compound conditioning in the rabbit. *Learning and Motivation, 6,* 314–326.

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,* 343-367.

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,* 370-378.

Schuch, S. & Koch, I. (2003) The role of response selection for inhibition of task-sets in task shifting. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 92-105.

Slagter, H.A., Weissman, D.H., Giesbrecht, B., Kenemans, J.L., Mangun, G.R., Kok, A., & Woldorff, M.G. (2006) Brain regions activated by endogenous preparatory set shifting as revealed by fMRI. *Cognitive, Affective & Behavioral Neuroscience, 6,* 175-189.

Shanks, D.R. (2010) Learning: From association to cognition. *Annual Review of Psychology, 61*, 273–301.

Shanks, D.R. & St John, M.F. (1994) Characteristics of dissociable human learning systems. *Behavioural and Brain Sciences, 17,* 367-447.

Sohn, M., Ursu, S., Anderson, J.R., Stenger, V.A., & Carter, C.S. (2000) The role of prefrontal cortex and posterior parietal cortex in task switching. *PNAS, 97(24),* 13448-13453.

Spector, A., & Biederman, I. (1976). Mental set and mental shift revisited. *American Journal of Psychology, 89,* 669-679.

Spence, K.W. (1952) The nature of the response in discrimination learning. *Psychological Review, 59*, 89-93.

Stanovich, K.E. and West, R.F. (2000) Individual differences in reasoning: Implications for the rationality debate. *Behavioral and Brain Sciences, 23*, 645–726.

Stevens, T. (2011) *Cortical regions involved in proactive control of task-set.* Unpublished doctoral thesis. University of Exeter, Exeter, UK.

Stoet, G., & Snyder, L.H. (2003a). Executive control and task-switching in monkeys. *Neuropsychologia, 41,* 1357–1364.

Stoet, G., & Snyder, L.H. (2003b). Task preparation in macaque monkeys (macaca mulatta). *Animal Cognition, 6,* 121–130.

Stoet, G. & Snyder, L.H. (2007) Correlates of stimulus-response congruence in the posterior parietal cortex. *Journal of Cognitive Neuroscience, 19*, 194–203.

Sudevan, P & Taylor, D.A. (1987) The cuing and priming of cognitive operations. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 89-103.

Sunstein, C., & Thaler, R. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness.* New Haven, CT: Yale University Press.

Sutton, R.S. & Barto, A.G. (1981) Toward a modern theory of adaptive networks: expectation and prediction. *Psychological Review, 88(2),* 135-170.

Sutton, R.S. & Barto, A.G. (1987) A temporal-difference model of classical conditioning. *Proceedings of the Ninth Annual Conference of the Cognitive Science Society,* (pp 355-378). Seattle, WA. Lawrence Erlbaum.

Talairach, J., & Tournoux, P (1988) Co-planar stereotactic atlas of the human brain: 3-dimensional proportional system: an approach to cerebral imaging. Stuttgart: Thieme.

Thorwart, A., Livesey, E.J. & Harris, J.A. (2012) Normalisation between stimulus elements in a model of Pavlovian conditioning: Showjumping on an elemental horse. *Learning and Behavior*

Vandierendonck, A., Liefooghe, B. &Verbruggen, F. (2010) Task switching: Interplay of reconfiguration and interference control. *Psychological Bulletin,136,* 601-626.

Wagner, A.R. (1981). SOP: A model of automatic memory processing in animal behaviour. In N.E. Spear & R.R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 5–47). Hillsdale , NJ: Lawrence Erlbaum Associates, Inc.

Waldron, E.M., & Ashby, F.G. (2001). The effects of concurrent task interference on category learning: Evidence for multiple category learning systems. *Psychonomic Bulletin & Review, 8,* 168-176

Ward-Robinson, J., & Honey, R.C. (2000). A novel contextual dimension for use with an operant chamber: From simple to hierarchical forms of learning. *Journal of Experimental Psychology*: *Animal Behavior Processes*, *26*, 358–363.

Warren, J.M. (1964) Additivity of cues in conditional discrimination learning by rhesus monkeys. *Journal of Comparative and Physiological Psychology, 58(1),* 124-126.

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*, 361-413.

Wills, A.J., & McLaren, I.P.L. (1997). Generalization in human category learning: A connectionist explanation of differences in gradient after discriminative and non-discriminative training. *Quarterly Journal of Experimental Psychology, 50A*, 607-630.

Wills, A.J., Reimers, S., Stewart, N., Suret, M., & McLaren, I.P.L. (2000). Tests of the ratio rule in categorization, and an alternative mechanism. *Quarterly Journal of Experimental Psychology, 53A,* 983-1011.

Yeung, N. & Monsell, S. (2003) The effects of recent practice on task switching. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 919-936.

Yeung, N., Nystrom, L.E., Aronson, J.A. & Cohen, J.D. (2006) Between-task competition and cognitive control in task switching. *Journal of Neuroscience, 26,* 1429-1438.

Young , M.E., Johnson, J.L. & Wasserman, E.A. (2000) Serial causation: Occasion setting in a causal induction task. *Memory & Cognition, 28(7),* 1213-1230.