Journal of Experimental Psychology: © 2014 American Psychological Association Learning, Memory, and Cognition 0278-7393/14/$12.00 DOI: 10.1037/a0035981 2014, Vol. 40, No. 4, 1002–1024.

Please cite as: Forrest, C.L.D., Monsell, S., and McLaren I.P.L. (2014). Is performance in task-cuing experiments mediated by task-set selection or associative compound retrieval? *Journal of Experimental Psychology: Learning, Memory and Cognition*, 40, 1002-1024.

Is performance in task-cuing experiments mediated by task-set selection or associative compound retrieval?

C.L.D. Forrest

S. Monsell

I.P.L. McLaren

School of Psychology, University of Exeter, UK

Revision date: 9 July, 2013

Running head: Associative learning and task-cuing

Length: Abstract 250 words, Text 17,341 words, 77 references, 8 figures, 1 table.

Correspondence Address

C.L.D. Forrest

Psychology

University of Exeter

Washington Singer Labs

Exeter EX4 4QG, UK

E-mail: clf206@exeter.ac.uk

Abstract

Task-cuing experiments are usually intended to explore control of task-set. But when small stimulus sets are used, they plausibly afford learning of the response associated with a combination of cue and stimulus, without reference to tasks. In three experiments we presented the typical trials of a task-cuing experiment: a cue (colored shape) followed, after a short or long interval, by a digit to which one of two responses was required. In a "Tasks" condition, participants were (as usual) directed to interpret the cue as an instruction to perform either an odd/even or a high/low classification task. In a “CSR” condition, to induce learning of mappings between cue-stimulus compound and response, participants were: in Experiment 1, given standard task instructions and additionally encouraged to learn the CSR mappings; in Experiment 2, informed of all the CSR mappings and asked to learn them, without standard task instructions; in Experiment 3, required to learn the mappings by trial and error. The effects of a task switch, response congruence, preparation, and transfer to a new set of stimuli, differed substantially between the conditions in ways indicative of classification according to task rules in the Tasks condition, and retrieval of responses specific to stimulus-cue combinations in the CSR conditions. Qualitative features of the latter could be captured by an associative learning network. Hence associatively-based compound retrieval can serve as the basis for performance with a small stimulus set. But, when organisation by tasks is apparent, control via task-set selection is the natural and efficient strategy.

Keywords: task-switching, task-cuing, associative learning, conditional discrimination, connectionist modeling

Human behaviour is often attributed to two types of processing: a set of controlled, resource-limited, and effortful processes, and a complementary set of involuntary, resource-unlimited and effortless processes. These have been referred to as "cognitive" and "associative" (McLaren, Green & Mackintosh, 1994) or as "intentional" and "automatic" (Jacoby, 1991). Typically, processes at both levels are thought to operate simultaneously, with a degree of independence (see McLaren, Green and Mackintosh, 1994 for a statement of this position and Mitchell, De Houwer & Lovibond 2009; Shanks, 2010 for critical reviews of it). In this paper, we examine aspects of performance in a task-switching paradigm often assumed to index “executive” or “endogenous” control processes, and ask to what extent performance might instead be accounted for by associative learning.

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 experiments intended to exercise and measure task-set control have attracted considerable interest over the last two decades (see Kiesel et al 2010; Monsell, 2003; Vandierendonck, Liefooghe, & Verbruggen, 2010; for reviews). In such experiments, the participant must respond on each trial to a stimulus according to one of several task rules The assumption is that 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, other things being equal. We focus here on the task-cuing paradigm (Meiran, 1996, Monsell & Mizon, 2006, Sudevan & Taylor, 1987).

The Task-Cuing Paradigm

In the task-cuing paradigm, participants are instructed and trained in two or more tasks afforded by a set of stimuli. Then, on each trial, a task cue is presented indicating which task the participant is to perform, followed by a stimulus from the available set. The participant is required to make an appropriate response as rapidly and accurately as possible. They may then receive corrective feedback, there is a pause, the next cue is presented, and so on. The experiments we report here required participants to classify a digit as odd/even or high/low following a task cue in the form of a colored shape. Some other examples of task pairs are listed in Table 1, to give a flavour of the wide range of realizations of the task-cuing paradigm.

-- Table 1 about here --

Among the phenomena seen in such experiments, we focus on three: the task switch cost, the response congruence effect and effects of preparation on the switch cost. We first review these phenomena, and theoretical interpretations of them in terms of task-set control processes (and their limitations). We then consider an alternative class of explanation, retrieval from memory of the learned response associated with the combination of cue and stimulus.

The switch cost. This is the often substantial increase in reaction time (and sometimes error rate) on trials when the task changes, compared to trials on which the task repeats in the same mixed task blocks (see Kiesel et al 2010; Monsell, 2003, Vandierendonck et al, 2010, for reviews).

The reduction in switch cost with preparation. The task-cuing 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: a long CSI 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 remove from the RT the 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 reconfiguration processes and stimulus processing after the stimulus; it may do any combination of these possibilities.

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 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. A third possibility is that task-set preparation simply "fails to engage" on a proportion of trials (De Jong, 2000; Nieuwenhuis & Monsell, 2002 but see Brown, Lehmann & Poboka, 2006). 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.

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 are “bivalent”: they afford both tasks. For example, if the two tasks are odd/even and high/low classifications of a digit, the left key might be used for "odd" and "low" and the right for "even" and "high". *Congruent* stimuli (e.g. 1) would receive the same response irrespective of task, while for *incongruent* stimuli (e.g. 7) the appropriate response depends on the task in play. Congruent stimuli are usually responded to faster than incongruent stimuli, (e.g. Kiesel, Wendt & Peters, 2007; Monsell & Mizon, 2006; Rogers & Monsell, 1995), 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 combines (congruent) or competes (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).

A compound-retrieval interpretation. Our purpose in this article is not to decide among these various task-set reconfiguration (TSR) accounts of how task-set control processes and their limitations might generate the phenomena in task-switching experiments, but to contrast them, as a class, with an account of performance in this situation as based on retrieval from memory of the learned response associated with the combination of cue and stimulus. Most task-cuing experiments use a small pool of stimuli that repeatedly recur throughout the experiment, and just one or two cues per task (see Table 1 for examples). In such circumstances, it is a logical option for the participant simply to learn a combination of simple and conditional discriminations (as they are sometimes described in animal learning experiments, see e.g. Mackintosh, 1974). For example, if the task cue is a background color, 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, only simple S-R associations need logically be learned: e.g. “3 → left”. Hence, although the experimenter may construe the experiment as one that requires switches between task-sets, the participant might (after learning the contingencies through 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 generic strategy as CSR learning (where C stands simultaneously for "cue", "contextual" and "conditional").

There are numerous examples of associative learning models that could, in principle, learn the contingencies for the task-cuing paradigm in the way described above (e.g. McLaren, 1993, 1994; Pearce, 1987), i.e. by forming mappings from the cue+stimulus to the response, though they have not been applied to the task-cuing paradigm. The behaviour of one such model will be discussed later and illustrated with a simulation in the General Discussion section.

Another example of a compound retrieval theory was explicitly proposed for the task-cuing paradigm by Logan & Bundesen (2003) and developed in a number of papers (Arrington & Logan, 2004, Logan, Schneider & Bundesen, 2007, Schneider & Logan, 2005, 2007). This account conceptualises performance in the task-cuing paradigm as based on retrieval from memory of the response maximally activated by the combination of encoded cue and encoded stimulus: the cue and the stimulus act as joint retrieval cues. Arrington and Logan (2004) distinguish two such compound retrieval strategies. The first they describe as

”episodic”: cue-stimulus pairs may be associated with responses in a relatively un-interpreted manner; this requires several repetitions of cue and target pairs before reliable associations develop, but this is plausible for the typical task-cuing experiment with a small set of stimuli (such as 8 digits) and one or two cues per task. But when large stimulus sets are used, there is no opportunity for such learning, yet performance is good and the characteristic switch costs and preparation phenomena are nevertheless observed (e.g. Arrington & Logan, 2004; Monsell & Mizon, 2006; Waszak et al, 2003). In later versions of the model therefore (e.g. Schneider and Logan, 2007) the cue is held to be encoded not as a specific stimulus, but as a task-name mediator (such as “parity”, in a paradigm where there is a single stimulus object such as a digit, or “word”, in a paradigm where one of two objects – e.g. a word and a face – must be responded to[1]). It is the compound of mediator and stimulus that jointly retrieve the relevant response from memory. If the attribute to be discriminated is familiar (such as odd-even) no new learning, except to translate the cue to the mediator, may be needed. In one sense this admits a task-representation to the model. But the critical difference remains that it assumes no change in task-set during a block; on each trial the participant “simply” retrieves the response associated with the encoded cue and encoded stimulus

Where then do “task-switch” costs come from? Logan and Bundesen’s (2003) starting point was that using only one cue for each task confounded the effect of changing or repeating a task with the possible effect of changing or repeating the cue on successive trials. Logan and Bundesen (2003) used two cues per task, which allows three categories of trial sequence: the cue repeats; the cue changes but the task remains the same; the cue and the task change. Contrasting the first two estimates the effect of a cue change, which they found to

---

[1] The quotation marks should not be taken to imply that the mediator is verbal: it may be a more abstract representation.

be large. Contrasting the last two estimated the effect of a task change, which they found to be small. They therefore proposed that what had previously been construed as task-switch costs arise largely because cue encoding is primed when the same cue was encoded on the previous trial. The distribution of the time required to encode the cue (which they model as exponential) is assumed to extend well into the post-stimulus interval when the CSI is short, and Logan et al assume that compound retrieval 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. Thus the driver of the RISC effect is the priming of cue-encoding.

Although Logan and Bundesen's (2003) original findings suggested that almost all the apparent cost of a "task-switch" could be attributed to a change of cue rather than task, other studies using two cues per task have shown that there can be a substantial cost when the task changes, with cue change controlled (e.g. Mayr & Kliegl, 2003; Monsell & Mizon, 2006; Scheider and Logan, 2007), and also clear reductions in those task-change costs with preparation (e.g. Monsell & Mizon, 2006). Schneider and Logan (2005, 2007) elaborated their priming account, proposing that, even when there are two cues per task, the two cues assigned to a task prime each others' encoding. This is either because signalling the same task associates them together, and there is associative priming (Schneider & Logan, 2005), or because the effective encoding of the cue is the mediator referred to above, and it is easier to generate the mediator if the same mediator was generated on the previous trial (Logan & Bundesen, 2004; Schneider & Logan, 2007).

Objections to the compound retrieval model.

Logan and colleagues' *compound-retrieval* model has excited controversy because it rejects what is for most users of the task-cuing paradigm its raison d'être: the hope that we can use it to measure the process and/or limitations of endogenous cognitive control as applied to the selection of task-sets. A number of objections have been made to the model. Monsell and Mizon (2006) discuss the plausibility of some of the assumptions, and noted that the marked variation in relative sizes of task and cue change effects across experiments seemed to be driven more by the probability of a task switch than the degree of association between the cues assigned to a task, or the likelihood of a task mediator being used. Mayr (2006) also found that participants were sensitive to the task-switch probability but not to learning the specific cue-stimulus pairs. Altmann (2006) was able to show that there was still a switch cost present on the second trial after a switch, even when a cue was not presented. Brass & von Crammon (2004) showed differential activation of the left IFJ, the right IFG and the right intra parietal sulcus related to task changes but not cue changes, again suggesting that performance is mediated by task-sets. The evidence most problematic for the account of the switch cost in terms of compound retrieval and priming of cue encoding comes from Logan's lab. Arrington, Logan and Schneider (2007) required subjects to make an overt response to the cue as soon as they had identified the stimulus dimension to be classified. This response then triggered presentation of the stimulus. RT to the cue should now include encoding time, and show priming effects, and cue encoding time should be excluded from RT to the stimulus. In fact there were substantial task-switch costs in the second RT, not in the first RT.

Regardless of such difficulties, the Logan/Bundesen/Arrington/Schneider compound-retrieval account – or at least the "episodic" version of it that applies to small stimulus sets – provides a well-specified example of a theory according to which the phenomena of cued task-switching can in principle be accounted for by CSR learning and retrieval, rather than processes of task-set selection and control. It is worth noting, however, that although this compound retrieval theory requires that CSR learning of some kind must have taken place, its authors say little about the learning process, nor have their experiments assessed it. Yet the amount of practice given before data are collected in such experiments has typically been quite small. We analyse below in more detail what has to be learned.

Rationale for the current experiments.

In the present studies, we took seriously the general proposition that CSR learning and retrieval is a plausible strategy for performance in the typical task-cuing paradigm. We maximised the viability of this strategy by keeping the number of stimuli very small (four to start with) to minimise the number of CSR associations to be learned. The experiment used a standard task-cuing paradigm, with a task cue (a coloured shape) preceding the digit stimulus and specifying an odd-even or a high-low classification rule. In the *Tasks* condition, participants were given standard instructions to interpret the cue as indicating which of two classification tasks to perform in response to the stimulus. Participants in the *CSR* condition received identical trial sequences, but were encouraged by instructional or training manipulations (which differed across the three experiments) to adopt a CSR strategy. The logic is straightforward. If participants in a standard task-cuing experiment (as represented by the Tasks condition) naturally use a CSR strategy, then encouraging them to do so in the CSR condition should not produce any qualitative difference between their performance in Task and CSR conditions (though possibly participants might learn CSR associations faster if we push them explicitly to do so). If, however, we get a qualitatively different pattern of results in the Tasks and CSR conditions, and evidence consistent with CSR learning in the CSR condition, this would suggest that typical performance in the standard task-cuing experiment cannot be reduced to compound retrieval. It would also suggest properties diagnostic of these two strategies, and allow us to probe their relative efficiency.

The Dreisbach et al. experiments.

A related instructional manipulation has been used by Dreisbach and colleagues (Dreisbach, Goschke & Haider, 2006; Dreisbach, Goschke & Haider, 2007; Dreisbach &

Haider, 2008; Dreisbach & Haider, 2009, see Dreisbach, 2012 for a review) to contrast a classification performed according to a task-rule with the same classification accomplished by means of individual S-R rules. Participants are given a set of individual S-R rules for binary classification of a set of stimuli. The stimulus set is designed so that an additional task-rule-based classification strategy is available, but it is sufficiently non-obvious that participants do not notice it until it is drawn to their attention. For example, participants were taught to classify a set of 4 red and 4 green words with what appeared to be eight arbitrary S-R mappings (half of each color left, half right). In fact, hard-to-spot task-based rules, cued by the color, were available: 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. Participants told these task-rules at the outset showed costs when the color switched, 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. Why then did the participants tipped off to the task-rule elect to use it if it leads to worse performance? Dreisbach and Haider (2008, 2009) report further experiments which 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. For example, Dreisbach and Haider's (2008) participants had to classify a word describing an item of clothing by whether it was worn on the top or the bottom of the body. Behind the word was a picture of an animal facing in a direction compatible or incompatible with the response. Using a task strategy reduced the effects of (irrelevant) directional compatibility.

The associative structure of the standard task-cuing experiment.

In Dreisbach et al.'s experiments, each stimulus was associated with only one response and presented only in one task, so that "simple" S-R associative learning would suffice. Hence, for their participants, using task-rules can be said to add unnecessary complexity. In spite of this, the evidence suggests they are used if offered at the outset, and confer some benefits in particular circumstances (e.g. in the presence of irrelevant information). Our experiments instead address the more common task-cuing paradigm in which all or many of the stimuli are bivalent: they afford responses in both tasks. Not only is this the situation which has revealed standard phenomena of task-switching, such as response congruence effects, and the one for which theoretical accounts, including the compound retrieval idea, have mainly been generated. It is also a situation in which switching between classification tasks, contingent on the cue, is plausibly a simpler strategy than CSR learning for solving the problem the experiment poses to the participant.

Considered from an associative learning point of view, bivalent stimuli add complexity because congruent and incongruent trials have different associative structures. Congruent stimuli can be classified by attending to a single feature (e.g. X**A**→ R, Y**A**→ R, X**B**→L, Y**B**→L; where X and Y are cues for two tasks, A and B are congruent stimuli, R is a right response and L is a left response); the cue can be ignored. In contrast, incongruent trials require what the associative learning literature calls a "biconditional discrimination" (e.g. XC→R, YC→L, XD→L, YD→R; where C and D are the incongruent stimuli). There are likely to be differences in the rate of acquisition of these sub-problems. Biconditional discriminations are particularly hard for human participants to learn (Harris & Livesey, 2008). For example, Livesey, Thorwart, De Fina & Harris (2011) in two experiments demonstrated that for human participants, biconditional discriminations were harder to learn than component discriminations (i.e. A → R and B → L). We therefore expect performance mediated by CSR learning to show a substantial effect of congruence, but for reasons different to those leading us to expect a congruence effect when people switch between task-sets.

Does an associative analysis of CSR learning predict any "switch cost"? We will argue that it does, particularly for the incongruent trials, due to facilitation of retrieval by learning on the previous trial (rather than by facilitation of cue encoding as in the Schneider and Logan, 2005, model). This analysis is backed up by the results of simulations of which we will report an illustrative example after the experiments. For now we offer an intuitive rationale. An advantage for "task-repeat" trials[2] can derive from the fact that the mappings between cue, stimulus and response are, to some extent, primed by learning on the previous trial, and this priming will be more facilitatory for "task-repeat" trials than for "task-switch" trials. Even though we take the precaution of using two cues per task, changing one dimension of the cue (Experiments 1 & 2) or the entire cue (Experiment 3) used to signal the task in play on every trial, our simulations suggest the intermediate representations formed to carry the mappings from cue and stimulus to response in a multi-layer network tend to combine the cues that signal the same response for a given stimulus so that they activate one intermediate representation — i.e. the network creates cue equivalence (Honey & Ward-Robinson, 2002). For incongruent stimuli this means that the cues for the same task will be treated equivalently. For incongruent stimuli, there is no benefit from this recent learning when the task switches, as now different intermediate representations and different pathways will be involved. The priming that results on repeat trials is thus due not to persistence of

---

[2] We use the terms "task-switch" and "task-repeat" to define trial sequences even when considering the performance of a hypothetical participant unaware of the task-cuing construal of the trial sequence; the quotation marks signal this usage.

activation over trials, but to the increments in the connections between input, intermediate and output representations that took place on the previous trial.

Given that we expect some switch cost in the CSR condition, what about the reduction in switch cost (RISC) effect? According to the Logan and Bundesen (2003) account, this arises from the priming of cue encoding, and we should expect to see it in the CSR, just as in the Tasks condition. But the basic idea of associative CSR learning does not by itself predict a RISC effect, indeed, it predicts none: the cue alone cannot usefully pre-activate any S-R or CSR associations because each cue is equally associated with both responses. Thus a basic associative account predicts no benefit of increasing the preparation interval on any "switch cost" that might be seen. Any advantage from a long CSI should be limited to generic preparation to process the stimulus and respond, i.e. a main effect of foreperiod. Moreover, while it could be argued that a long CSI removes the need simultaneously to process cue and stimulus, it is also possible that simultaneous presentation of the cue and the stimulus in the CSR condition might benefit CSR learning by facilitating configural processing of the cue and stimulus.

Task-switching in monkeys.

Another basis for predictions in our 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. 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. 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; this is in line with the predictions about the switch cost made above. 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 might well benefit the monkey less on a repeat than it would on a switch trial (as Logan & Bundesen, 2003, suggested). This does not bear upon our prediction that if cue change and task change are un-confounded, then little or no RISC effect will be observed when using a CSR strategy. Stoet and Snyder (2007) also found a large congruence effect in the monkeys they tested, supporting the prediction that approaching the experiment without a task representation should indeed yield a large congruence effect.

Transfer to a new stimulus set.

In addition to examining switch costs, RISC effects and congruence effects, our experiments also included a manipulation especially relevant to assessing learning. Towards the end of each experiment, after practice on just four of the 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 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 as 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. We would thus 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 tasks condition.

The experiments.

The three experiments all compared a group receiving standard task-cuing instructions (the *Tasks* group) to a *CSR* group who received the same trials, but were encouraged to learn and use CSR associations. In Experiment 1, 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. In Experiment 2, 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 for learning as a list at the beginning of the experiment. Finally, Experiment 3 required the CSR participants to learn responses by trial and error alone, much like Stoet and Snyder's monkeys.

Of course, in both Experiments 2 and 3, it was perfectly possible for participants in the CSR group to induce and use a task-set rule, even though they were not told to use such a rule or to look for one. Likewise, it was possible for participants in the Tasks condition to notice and then strive to learn and use CSR associations instead of or as well as the tasks. Hence in Experiments 2 and 3 we interviewed the participants in a post-experiment debriefing, to enable us to select a sufficient group of participants in each condition who appeared to be unaware of the alternative strategy. We leave for future research the idea that, given common conceptions of the automaticity of associative learning, we might expect participants who do not explicitly notice that they can learn S-R and CSR rules to learn them nonetheless.

Experiment 1

On each trial the participant saw a task cue (a square or diamond filled with pink or blue) in the centre of the screen. After a 100 or 1200 ms CSI, one of four digits appeared superimposed on the colored shape. The task, odd/even or low/high classification, was specified by the color of the cue for some participants, the shape for others. Either the shape or color 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 also encouraged to learn the CSR mappings.

Method

Participants

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

Apparatus and stimuli.

The experiment was programmed in SuperLab (v1.5) and presented on a Macintosh Performa computer with a 13″ screen. Participants were tested individually. 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; for half the subjects in each group the color signalled the task, for the other half the shape. The side of the square was ~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 centre of the shape in Geneva font, and was ~1.7° high.

Design.

The experiment had a between-subjects factor of instruction (CSR versus task-set), and within-subjects factors of trial type (switch versus repeat), response congruence (congruent versus incongruent), and CSI (*long*, i.e. 1200 ms versus *short* i.e. 100 ms). After practice and instruction, a session consisted of 14 blocks half of 61, half of 63 trials[3]. 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 test of transfer of training. 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.

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.

There were 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 (see Appendix A) designed to encourage them to learn and use “pairings” such as "Square + 2 → RIGHT". Half the participants were assigned to the CSR

---

[3] The small variation in block length is a by-product of generating trial sequences with an algorithm that ensures the distribution of run lengths per task approximates overall to binomial expectation.

condition and the other half to the Tasks condition. For the reader's benefit, Figure 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.

-- Figure 1 about here --

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, the response-cue interval was 1600 ms in the short blocks, and 500 ms 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 1500 ms, before the experiment continued; if he/she failed to respond within 4000 ms, “Time Out” appeared (in black). At the end of a block the participant’s 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 digits they were going to see would be different.

Results

In all the experiments, trials following an error were excluded from the analysis, and error trials were excluded from the RT analysis. We characterize the effects of practice in terms of changes over block pairs; each pair contained one short- and one long-CSI block whose order was counterbalanced over participants.

-- Figure 2 about here --

As illustrated in Figure 2*,* participants in both the Tasks (shown right) and CSR (shown left) conditions showed an improvement in performance from block pair 1 to 5. Although the CSR group, who received the extra instructions to learn CSR associations, started off slower and less accurate, 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[4] of 685 ms and 5.21% errors, compared to 676 ms and 4.83% errors for the Tasks group. Figure 3 summarizes three main outcomes of this experiment: a larger switch cost in the Tasks condition than the CSR condition, a RISC effect in the Tasks condition but not in the CSR condition, and a much smaller congruence effect in the Tasks condition than in the CSR condition. Figure 2 also shows 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 congruence effect) discussed in the introduction will now be considered in turn, first for the main experiment (block pairs 1-5), then for the transfer blocks.

Block pairs 1-5

To test the critical interactions of instruction condition with effects of a task switch, congruence and preparation time, a mixed ANOVA with the factors: task switch (repeat, switch), congruence (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. The salient contrasts are shown in Figure 3. Effect sizes are reported as generalised eta squared η_G^2 (Olejnik & Algina, 2003) as recommended for designs with repeated measures (Bakeman, 2005).

-- Figure 3 about here --

---

[4] In all experiments we repeated the same analyses with mean RT: effects were slightly noisier but qualitatively very similar. We note in footnotes two cases where a critical effect that reached significance in the analysis of medians did not reach significance for the means.

Task switches and instruction. The Tasks group had a larger RT switch cost (76 ± 7.57 ms)[5] than the CSR group (40 ± 5.93 ms), $F(1,46) = 7.28$, $p < .05$, $\eta^2_G = 0.00124$, but the switch costs in both the Tasks condition, $F(1,23) = 69.8$, $p < .001$, $\eta^2_G = 0.0329$, and the CSR condition, $F(1,23) = 39.4$, $p < .001$, $\eta^2_G = 0.00898$, were reliable. For errors, the overall main effect of a task switch was reliable, $F(1,46) = 39.7$, p < .001, $\eta^2_G = 0.0349$, (repeat: 4.42%, switch: 7.45%), but there was no interaction with instruction: the switch cost was 3.06 ± 0.58% for the Tasks group and 3.01 ± 0.41% for the CSR group. Hence the switch cost over these blocks was greater in the Tasks than in the CSR condition (see Figure 3).

Preparation and instruction. As Figure 3 shows, preparation reduced the switch cost in the Tasks group RT from 97 ms in the short-CSI blocks to 65 ms in the long-CSI blocks, a reduction of 32 ± 10.3 ms. There was no reduction in the CSR group; their switch cost was 37 ms in the short-CSI blocks and 42 ms in the long-CSI blocks, a difference of -5 ± 8.72 ms. The interaction was marginally reliable $F(1,46) = 2.54$, $p = 0.12$, $\eta^2_G = 0.000428$,. Separate analyses revealed a significant interaction of CSI and task switch in the Tasks condition, $F(1,23) = 5.39$, $p < .05$, $\eta^2_G = 0.00244$, 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%), a reduction of 1.59 ± 0.49% For the CSR group the switch cost was 2.45% for the short-CSI blocks, and 3.56% for the long-CSI, a difference of -1.11 ± 0.65% This interaction was reliable, $F(1,46) = 4.96$, $p < .05$, $\eta^2_G =$

---

[5] mean and standard error of contrast

0.000704. Hence there was a clear RISC effect for the Tasks group, but none for the CSR group, and the RISC effect differed between groups.

However, both groups benefited more generally from the opportunity to prepare, as we may see by examining task-repeat trials alone. RT in the Tasks group was 120 ± 22.7 ms faster after a long-CSI (627 ms) than after a short (747 ms), $F(1,23) = 30.1$, $p < .001$, $\eta_G^2 = 0.105$. A similar but less marked effect was seen in the CSR condition, where repeat trials were 30 ± 11.9 faster in the long-CSI (714 ms) than in the short (744 ms), $F(1,23) = 9.63$, $p < .01$, $\eta_G^2 = 0.00842$. The difference in this effect between the two groups was significant, $F(1,46) = 11.7$, $p < .01$, $\eta_G^2 = 0.0130$. This effect was not significant in the errors for either group.

Congruence and instruction. RT and error rates showed a substantially larger effect of congruence (incongruent minus congruent) in the CSR group (299 ± 26.3 ms, 9.28 ± 0.95%) than in the Tasks group (138 ± 16.9 ms, 7.74 ± 0.82%); the interaction was highly reliable for RTs, $F(1,46) = 12.4$, $p < 0.01$, $\eta_G^2 = 0.0220$, but not error rates. In separate analyses for the two groups, the congruence effect was reliable for the Tasks group, $F(1, 23) = 32.2$, $p < .001$, $\eta_G^2 = 0.0725$, for RT; $F(1,23) = 46.0$, $p < .001$, $\eta_G^2 = 0.208$, for errors; and for the CSR group: $F(1,23) = 61.6$, $p < .001$, $\eta_G^2 = 0.217$, in RTs; $F(1,23) = 60.9$, $p < .001$, $\eta_G^2 = 0.256$, in errors.

Acquisition effects. As may be seen in Figure 2, overall performance evidently improved from block pair 1 to 5, and this was reliable for RTs $F(4,184) = 184$, $p < .001$, $\eta_G^2 = 0.0648$. Although the RT improvement looks greater for the CSR 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, $F(4,184) = 3.38$, $p < .05$, $\eta_G^2 = 0.00556$. Separate analyses revealed significant effects of block pair in both the Tasks, $F(4,92) = 2.73$, $p < .05$, $\eta_G^2 = 0.00817$, and CSR groups, $F(4,92) = 10.0$, $p < .001$, $\eta_G^2 = 0.0341$.

The associative account outlined in the introduction predicted that congruent stimuli should be learned more rapidly than incongruent stimuli. This is supported by the interaction between block pair, congruence and instructions in the error rates, $F(4,184) = 3.90$, $p < .05$, $\eta_G^2 = 0.00586$, driven largely by an interaction between block pair and congruence in the CSR group, $F(4,92) = 7.84$, $p < 0.001$, $\eta_G^2 = 0.0228$, but not in the Tasks group, $F < 1$, and by the non-reliable interaction in the RTs (see Figure 2). For the CSR group, performance on the congruent stimuli appears to asymptote in the error rates early on, and after three block pairs in the RTs, while the much poorer performance on the incongruent stimuli continues to improve steadily. For the Tasks group, there was a 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 is expected on a control account if the irrelevant task-set is more active on switch than repeat trials. The switch cost was larger for incongruent (Tasks: 95 ms, CSR: 75 ms) than for congruent stimuli, (Tasks: 57 ms, CSR: 4 ms), and this effect was much more pronounced in the CSR condition than in the Tasks condition. Separate analyses revealed a significant interaction in the CSR group, $F(1,23) = 59.7$, $p < .001$, $\eta_G^2 = 0.00514$, but not in the Tasks group. In the error data the pattern was similar. Again there were larger switch costs for the incongruent stimuli (Tasks: 4.96%, CSR: 5.68%) than for the congruent stimuli (Tasks: 1.15%. CSR:

0.34%). This did not interact with instructions. Separate analyses revealed a significant effect in both the CSR group, $F(1,23) = 18.3$, $p < .001$, $\eta_G^2 = 0.0240$, and the Tasks group, $F(1,23) = 19.2$, $p < .001$, $\eta_G^2 = 0.0190$.

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: 170 ms; CSR: 315 ms), than with a long CSI (Tasks: 105 ms; CSR: 283 ms), $F(1,46) = 11.3$, $p < .01$, $\eta_G^2 = 0.000643$, but this change did not differ much between the instruction conditions: this effect was not significant in the errors.

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), congruence, CSI, block pair (5 or 6) and instruction condition. This seems the most straightforward of the several possible measures of transfer. 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 137 ± 20.9 ms (2.41 ± 0.59%) as compared to the Tasks group's 55 ± 20.3 ms (1.15 ± 0.47%). The interaction was significant for RT, $F(1,46) = 4.13$, $p < .05$, $\eta_G^2 = 0.00753$, but not error rate, $F(1,46) = 1.36$, $p = 0.25$, $\eta_G^2 = 0.00163$, although there was a main effect of transfer in the errors, $F(1,46) = 10.7$, $p < .01$, $\eta_G^2 = 0.0128$. Under both instruction conditions transfer to new stimuli increased the switch cost, by 48 ± 7.6 ms

(1.17 ± 0.58%) in the Tasks group, and 64 ± 6.0 ms (-0.23 ± 0.41%) in the CSR group. In the RTs there was a significant interaction between block pair and task switch, $F(1,46) = 16.6$, $p < .001$, $\eta^2_G = 0.00348$, but not in the error rate $F < 1$. This interaction was separately significant in RTs in the Tasks, $F(1,23) = 9.70$, $p < .001$, $\eta^2_G = 0.00329$ and CSR groups, $F(1,23) = 7.99$, $p < .05$, $\eta^2_G = 0.00373$, but in the errors for neither condition, $F < 1.5$.

A further analysis showed that performance improved from block pair 6 (Tasks: 731 ms, 5.98%, CSR: 822 ms, 7.62%) to 7 (Tasks: 688 ms, 5.89% CSR: 738 ms, 5.52%) significantly in both the RTs, $F(1,46) = 22.6$, $p < .001$, $\eta^2_G = 0.0393$, and the errors, $F(1,46) = 10.7$, $p < .01$, $\eta^2_G = 0.0128$. This decrease was larger in the CSR group than the Task group in the RT data, $F(1,46) = 4.14$, p < .05, $\eta^2_G = 0.00753$.

Discussion

In the Tasks condition, as is standard in task-switching experiments, we instructed the participants to construe the cue as indicating the appropriate categorical task-rule to apply (odd-even or high-low). In the CSR condition, the only difference was that we augmented the standard instructions with a strong hint to try to learn stimulus-specific S-R and CSR mappings. The hint was effective in changing performance, though it did not improve it. The Tasks group showed the typical large switch cost, a modest response congruence effect, and the reduction in switch cost (and congruence effect) with preparation expected under these conditions. In contrast, the CSR group showed a smaller switch cost, no reduction in switch cost with preparation, and a larger congruence effect, consistent with the predictions of a simple associative account. The initial performance of the CSR group was poorer, perhaps because they were simultaneously grappling with task-level and stimulus-level mappings, but they had largely caught up by the fifth block pair. 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 showed no switch costs after just one block pair. Performance on the incongruent stimuli showed much more gradual improvement with practice, and continued to show a switch cost. This is consistent with CSR learning.

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 were reliable.

We draw two preliminary conclusions. First, it appears that associative learning of S-R and CSR associations can come to mediate performance in a task-cuing paradigm with a small set of cues and stimuli presented many times, and is a reasonably effective strategy. We do not claim that the CSR group were reliant only on this associative strategy: we had, after all, instructed them in a task-set level strategy as well, and the mere addition of a hint to use an associative learning strategy is a relatively weak manipulation. Participants may well have varied considerably in the mixture of strategies they were using. Hence in Experiments 2 and 3 we used training and instruction regimes designed to produce a better segregation of the task-based and association-based strategies.

The hint to learn associations 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 our second conclusion is that it is unlikely that participants in our Tasks group, and by extension other task-cuing experiments, were merely retrieving the response associated with each cue + stimulus combination, contrary to the proposal of Logan and Bundesen (2003). Rather they were using the cue to select a categorical task-rule, and then applying 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, a reduction of the otherwise complex mapping of stimuli to responses to a simpler one. 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 effective and natural.

However, because we are likely seeing a mixture of strategies in the CSR group, this experiment merely suggests features of the data diagnostic of associative versus task-set responding. We cannot rule out, for example, the possibility that the CSR group learned and used only S-R, and none of the CSR, associations — i.e. their performance on the incongruent stimuli may still have been based largely on application of categorical task rules. The fact that the switch cost in the CSR condition was as large for the incongruent trials in the CSR group (75 ms) as the overall switch cost seen in the Tasks group (75 ms) would be in line with this hypothesis. A stronger manipulation of strategy is needed.

Experiment 2

In Experiment 2, the instructions to the CSR group contained no mention of two cue-contingent categorisation task-rules. Participants were simply provided with a piece of paper listing the CSR mappings (in the same form as the right half of Figure 1) and asked to learn them, using block pair 1 as practice. Unlike Experiment 1 there was no practice nor any exposure to all the digits before block pair 1. This also meant that when participants were transferred from one set of stimuli to the other in block pair 6 they had not previously responded to the new digits at all. The Tasks group received exactly the same blocks of trials as the CSR group, but with standard task-cuing instructions prior to block pair 1.

Participants shown a list of stimulus-specific rules may nevertheless work out that categorical odd/even and low/high rules contingent on the cue can be used. Participants given task-cuing instructions may nevertheless adopt a CSR strategy. We therefore debriefed participants on what they thought they were doing, and replaced participants in each group until we had two groups of 16 whose introspections matched the instructions in force for their group. An analysis of the performance of the replaced participants will be discussed after Experiment 3.

Method

Participants

The participants were 44 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 contingent on their performance (average payment £2.11, range £1.75-£3.00).

Apparatus and stimuli

The experiment was programmed in Matlab (2008b) with Psychtoolbox installed, and run on an iMac computer. The participants were seated 50 cm from the 20″ screen and tested individually. Cues were colored 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 centre of the cue in Courier bold font size 60 (~1.3° tall).

Design and Procedure

The design and randomization of the main part of the experiment were almost identical to those of Experiment 1, with the minor difference that which four digits were congruent and which were incongruent was counterbalanced by swapping the response mappings for the high/low task for half the participants. The procedure differed from Experiment 1 only in the following respects. First, the single-task practice and switching practice blocks of Experiment 1 were deleted; what was described to the participant as "practice" was now effectively block pair 1 of Experiment 1. Second, participants in the CSR condition were at the outset given the CSR combinations as a list (as on the right side of Figure 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 1). Third, as the participants in the present experiment were, unlike those in Experiment 1, participating for credit, we added small bonus payments for good performance: 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 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).

Eight Tasks group participants were replaced because when asked “did you use the tasks to help you?”; they said they did not, leaving 16 who said they did. Four CSR group participants were replaced because they mentioned using at least one categorical task rule, leaving 16 who did not. The excluded participants will be discussed after Experiment 3.

--- Insert Figure 4 about here ---

Results

In Experiment 2, block pair 1 reflects the initial encounter with these conditions, and block pair 6 now represents the first encounter with the transfer stimuli. So the analysis strategy was similar to that in Experiment 1, except that the initial block pair was excluded. Both groups showed substantial improvement over the first 5 block pairs, with a marked improvement from the first to the second block pair. As Figure 4 shows, as in Experiment 1, 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. By the fifth block pair, the difference was less marked, with the CSR group having a median correct RT of 740 ms (3.57% errors), and the Tasks group 639 ms (4.76% errors). The critical contrasts summarised in Figure 5 show that, as in Experiment 1, the Tasks group showed a larger switch cost than the CSR group, a RISC effect when the CSR group showed none, and a smaller congruence effect than the CSR group. Figure 4 shows that, as in Experiment 1, 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.

Block Pairs 2-5.

Overall the CSR group were slower (795 ms) and made more errors (5.44%) than the Tasks group (681 ms, 4.78%). This was a significant difference for RTs, $F(1,30) = 4.32$, $p < .05$, $\eta_G^2 = 0.0682$, but not for the errors, $F < 1$.

Task switches and instruction. As Figure 5 shows, the Tasks group had a much larger RT switch cost (139 ms) than the CSR group (29 ms), $F(1,30) = 9.19$, $p < .001$, $\eta^2_G = 0.0168$, a difference of 110 ± 18.2 ms The switch cost in the Tasks condition was highly reliable, $F(1,15) = 17.6$, $p < .001$, $\eta^2_G = 0.117$, but the effect in the CSR group only marginally reliable, $F(1,15) = 3.39$, $p = 0.086$, $\eta^2_G = 0.00372$. The error switch cost was also larger for the Tasks (1.95 ± 0.42%) than the CSR group (0.15 ± 0.56%), but this difference (1.8 ± 0.50%) was only marginally reliable, $F(1,30) = 3.28$, $p = 0.080$, $\eta^2_G = 0.00378$; there was a significant switch cost in the Tasks group, $F(1,15) = 10.6$, $p < .01$, $\eta^2_G = 0.0188$, but not in the CSR group, F<1.

--- Insert Figure 5 about here ---

Preparation and instruction. As Figure 5 shows, preparation reduced the switch cost in the Tasks group from 184 ms in the short-CSI blocks to 94 ms in the long-CSI blocks, a reliable RISC effect of 90 ± 17.6 ms, $F(1,15) = 6.42$, $p < .05$, $\eta^2_G = 0.0136$. The equivalent effect in the CSR group was much smaller and not significant: the switch cost was 40 ms in the short-CSI blocks and 17 ms in the long-CSI blocks, a difference of 23 ± 8.41 ms, $F(1,15)= 1.91$, $p = 0.187$, $\eta^2_G = 0.000634$.. However the three-way interaction was only marginally reliable, $F(1,30) = 2.86$, $p = 0.1$, $\eta^2_G = 0.000152$. For the error rates, a similar pattern was present: for the Tasks group there was a larger switch cost in the short-CSI blocks (3.1%) than the long-CSI blocks (0.8%), a reduction of 2.3 ± 0.58%, $F(1,15) = 4.05$, $p = 0.063$, $\eta^2_G = 0.00661$. For the CSR group the cost was 0.57% for the short-CSI blocks, and -

0.26% for the long-CSI, a difference of .83 ± 0.49%, $F < 1$. The 3-way interaction was not reliable, $F < 1$.

Participants in the Tasks condition also showed a overall preparation effect when just the repeat trials were considered, such that they were faster in a long-CSI (531 ms) than a short-CSI (692 ms), $F(1,15) = 59.0$, $p < .001$, $\eta^2_G = 0.184$. However, they also made more errors at the long-CSI (4.41%) than the short (3.19%) but this was not reliable, $F(1,15) = 1.52$, $p = 0.237$, $\eta^2_G = 0.00423$. The CSR participants were also less accurate but only slightly faster at the long-CSI (775 ms, 5.67%) than at the short (787 ms, 5.06%). The interaction was not reliable, $F < 1$.

Congruence and instruction. RT and error rates showed a much larger effect of congruence (incongruent minus congruent) in the CSR group (299 ms, 6.68%) than in the Tasks group (58 ms, 6.54%); the difference (241 ± 26.0 ms, 0.14 ± 0.90%) was highly reliable for RTs $F(1,30) = 21.5$, $p < .001$, $\eta^2_G = 0.0749$, but not error rates, $F < 1$. In separate analyses of the two groups, the congruence effect was reliable for both the Tasks group, RT: $F(1, 15) = 12.6$, $p < .001$, $\eta^2_G = 0.0223$, errors: $F(1,15) = 26.3$, $p < .001$, $\eta^2_G = 0.176$ and for the CSR group: RT: $F(1,15) = 36.6$, $p < .001$, $\eta^2_G = 0.294$ errors: $F(1,15) = 28.3$, $p < .001$, $\eta^2_G = 0.164$.

Acquisition. As may be seen in Figure 4, overall performance evidently improved from block pair 2 to 5, and this was reliable for RTs, $F(3,90) = 26.2$, p < .001, $\eta^2_G = 0.0495$ and the errors, $F(3,90) = 10.5$, $p < .001$, $\eta^2_G = 0.0389$. But 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. However, the congruence effect in the CSR group changed across block pairs (as with Experiment 1), as indicated by the interaction between block pair, congruence and instructions in the RTs, $F(3,90) = 4.91$, $p < .01$, $\eta_G^2 = 0.00483$, but not in the errors; see Figure 4. In separate analyses of the two groups there was a reliable interaction in the CSR group, RT: $F(3,45) = 5.11$, $p < .05$, $\eta_G^2 = 0.0130$; errors: $F(3,45) = 5.42$, $p < .01$, $\eta_G^2 = 0.0348$, 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 block 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, $F(3,90) = 2.76$, $p < .05$, $\eta_G^2 = 0.00683$. Separate analyses showed this was a marginally significant interaction in the Tasks group, $F(3,45) = 2.59$, $p = 0.064$, $\eta_G^2 = 0.0121$, 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 with a short CSI (7.12%) than a long CSI (5.65%) but then by block pair 4 they were making more errors with a long CSI (6.11%) than a short CSI (3.42%).

Other significant interactions. The present results revealed a slightly larger switch cost for incongruent (Tasks: 147 ms, CSR: 33 ms) than congruent trials (Tasks: 130 ms, CSR: 23 ms), but the interaction was not reliable, and the error data did not show the same trend. There was a difference in error rate between the two instructional conditions, $F(1,30) = 16.1$, $p<0.001$, $\eta_G^2 = 0.00943$. Separate analyses revealed a significant interaction between switch and congruence in the Tasks group, $F(1,15) = 10.8$, $p < .01$, $\eta_p^2 = 0.0190$, and in the CSR group, $F(1,15) = 5.36$, $p < .05$, $\eta_p^2 = 0.00342$. In the Tasks group there were larger error 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 congruence (significant in Experiment1), was not significant in Experiment 2, $F < 2$, for either RTs or errors.

The transfer test.

As in Experiment 1 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), congruence, CSI, block pair (5 or 6) and instruction condition. As expected, the CSR group was more affected by the transfer to new stimuli; their mean RT (error rate) increased by 177 ± 51 ms (1.97 ± 0.88%) as compared to the Tasks group's (48 ± 8.5 ms, -0.94 ± 0.60 %). However, the interaction was only marginally reliable for both RTs, $F(1,30) = 3.15$, $p = 0.086$, $\eta_p^2 = 0.0158$, and the error rate, $F(1,30) = 3.72$, $p = 0.063$, $\eta_p^2 = 0.0104$.

Transfer to new stimuli increased the size of the congruence effect in the CSR group (26 ± 26.9 ms, 4.78 ± 0.89 %) but not in the Tasks group (4 ms ± 10.9, -1.48 ± 0.83 %). This was supported by a significant interaction in the errors, $F(1,30) = 6.57$, $p < .05$, $\eta_p^2 = 0.0120$, but not in the RTs, $F < 1$. Separate analyses of the errors revealed a significant interaction between block pair and congruence only in the CSR group, $F(1,15) = 11.1$, $p < .05$, $\eta_p^2 = 0.132$.

Under both 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 at a short (Tasks: 2.2%, CSR: 1.89%) compared to a long-CSI (Tasks: 0.87%, CSR: -0.71%), to having a larger switch cost at a long-CSI (Tasks: 0.95% CSR: 2.41%) rather 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 $F(1,30) = 4.99$, $p < .05$, $\eta_G^2 = 0.00411$. This was not the case in the RTs, $F(1,30) = 1.83$, p = 0.186, $\eta_G^2 = 0.000522$, 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: 135 ms, CSR: 42 ms) than after transfer (Tasks: 78 ms, CSR: 4 ms).

A further analysis showed that performance improved from block pair 6 (Tasks: 687 ms, 3.83%, CSR: 917 ms, 5.54%) to 7 (Tasks: 626 ms, 4.20% CSR: 765 ms, 3.70%) significantly in the RTs, $F(1,30)=18.4$, $p<0.001$, $\eta_G^2 = 0.0409$, but not in the errors, $F(1,46)=1.16$, $p = 0.291$, $\eta_G^2 = 0.00256$. The improvement was greater for the CSR group, especially for incongruent stimuli, but the corresponding interactions were not reliable; there was a considerable variability over participants within groups in how they coped with the transfer test.

Discussion

The results of Experiment 2 show that participants in a task-cuing paradigm under instructions to perform the task using specified CSR rules (and with no mention of the possibility of using a task-set strategy) produce a pattern of results quite different from participants given standard task-switching instructions. As in Experiment 1, the CSR group showed large congruence effects, small switch costs and no reduction in switch cost with preparation, while the Tasks group showed relatively small congruence effects and large switch costs which 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 then for the Tasks group. The consistent pattern across both experiments encourages us to believe that it reflects genuine processing differences contingent on the participant's approach to the task rather than mere artifacts of procedure. These effects will be discussed further in the General Discussion.

The interaction between switch and congruence in the CSR errors seems inconsistent with Experiment 1. However, the RTs show a non-significant effect in the opposite direction, so we do not dwell on this. The lack of a CSI by congruence interaction in this experiment is another point of disagreement between the first two experiments. However, as Monsell and Mizon (2006) and Kiesel et al (2010) note, this interaction has an inconsistent history in the literature.

The transfer test was somewhat unsatisfactory in this experiment, because the new stimuli in block pair 6 were accompanied by an instruction sheet giving the new CSR mappings for the CSR group, and the task rules applied to the new stimuli for the Tasks group. Thus, although the CSR group showed a greater perturbation in performance, this may be 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 3, in which we adopted a procedure perhaps more suited to the idea of associative learning: we taught the CSR mappings to the CSR group by means of trial and error alone.

Another objection that could be raised against both experiments so far is that the set of task cues used do not truly constitute two distinct cues per task. Though the cue changed on all trials, the dimension that changed on repeat trials was essentially irrelevant to the response, and participants could have learned to ignore that dimension. Hence it is possible that the small switch cost seen in the CSR condition was caused by the cue changes being more salient on switch trials, so that the apparent switch cost is an effect of cue change. Experiment 3 addressed this issue by using 4 shape cues (a circle, triangle, square and pentagon), with two arbitrarily assigned to each task, and a trial sequence such that the cue changed on every trial.

Another issue concerns the way in which knowledge is acquired. In the CSR groups of Experiments 1 and 2 participants either knew the task-sets (Experiment 1) or were explicitly told the correct responses for a given cue and stimulus pair (Experiment 2), so that any associative learning that occurred was probably (initially) supported by another (declarative) representation of the contingencies. In Experiment 3, we removed these props and, while presenting the same sequences of trials, required participants to learn the responses simply by trial and error, as perhaps the most straightforward "invitation" to the system to use basic associative learning to solve the problem.

Experiment 3

There were two differences between this and Experiment 2. First, no instructions at all were given to the CSR group, except that they had to learn by trial and error the correct response to each combination of cue and stimulus. Second, two distinct shape cues were used for each task.

Method

Participants

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

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 centre of the cue in size 60 Courier bold font as before. The two sets of digits used were again 1,4,7,8 and 2,3,6,9. The computer and screen were as in Experiment 2.

Design and procedure

The design and randomization of the trial sequences was as in Experiment 2 except that there were now four distinct shape cues, with two shapes signaling one task and two the other. Thus, the cue changed on every trial, but only some changes signaled a change of task.

Participants in the Tasks condition were given standard task-set instructions verbally, 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. As in 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 2, 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 later.

Results

--- Insert Figure 6 and 7 about here ---

Performance is shown as a function of block pair in Figure 6, and critical contrasts are summarized in Figure 7. The main results are similar to those in the previous experiments. A small but reliable switch cost in the CSR group, but a larger one in the Tasks group: a much larger congruence effect in the CSR than the Tasks group; a large RISC effect in the Tasks group, none in the CSR group; transfer to new stimuli much more disruptive in the CSR group.

Block Pairs 2-5

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 (735 ms) than the Tasks group (651 ms) and they made more errors (5.02%) than the Tasks group (3.09%); the difference stems largely from performance on incongruent stimuli.

As for Experiment 2, a mixed ANOVA with the factors: task switch (repeat, switch), congruence (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 (777 ± 106.2 ms) and made more errors (11.9 ± 6.4 %) than the Tasks group (700 ± 94.6 ms, 4.62 ± 2.97%). This was a significant difference for errors, $F(1,30) = 25.6$, $p < .001$, $\eta^2_p = 0.118$, but not RTs, $F(1,30) = 2.47$, $p = 0.13$, $\eta^2_p = 0.036$.

---- Insert Figure 7 about here ---

Task switches and instruction. As with Experiments 1 and 2 there was a larger switch cost in the Tasks group (73 ± 13.4 ms) than in the CSR group (30 ± 8.8 ms), $F(1,30) = 3.83$, $p < .06$, $\eta^2_p = 0.00307$ (see Figure 7). The costs were reliable for both the Tasks condition,

$F(1,15) = 15.1$, $p < .001$, $\eta^2_G = 0.0452$ and the CSR condition, $F(1,15) = 5.58$, $p < .05$, $\eta^2_G = 0.00420$. For errors, the overall main effect of a task switch was reliable, $F(1,30) = 6.58$, $p < .05$, $\eta^2_G = 0.00597$ (repeat: 7.98%, switch: 9.05%), but there was no reliable interaction with instruction, $F(1,30) = 1.31$, $p = 0.26$, $\eta^2_G = 0.00119$: the switch cost was 0.86 ± 0.49 % for the Tasks group and 2.24 ± 0.70 % for the CSR group

Preparation and instruction. As Figure 7 shows, preparation reduced the RT switch cost in the Tasks group by 50 ± 10.5 ms, from 99 ms in the short-CSI blocks to 49 ms in the long-CSI blocks, $F(1,15) = 5.84$, $p < .05$, $\eta^2_G = 0.00559$. There was no such effect in the CSR group, for whom the switch cost was 27 ms in the short-CSI blocks and 32 ms in the long-CSI blocks, an increase of 5 ± 11.3 ms, $F < 1$. The interaction was marginally reliable $F(1,30) = 3.27$, $p = 0.084$, $\eta^2_G = 0.00122$. 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 (527 ms) than with a short-CSI (849 ms), $F(1,15) = 155$, $p < .001$, $\eta^2_G = 0.481$. For the same contrast the CSR group was also faster with the long- CSI (745 ms) than at the short-CSI (845 ms), $F(1,15) = 16.9$, $p < .001$, $\eta^2_G = 0.0344$, but the effect was considerably smaller. The three-way interaction was significant, $F(1,30) = 64.0$, $p < .001$, $\eta^2_G = 0.107$. These interactions were not significant in the errors ($F < 1$).

Congruence and instruction. RT and error rates showed (Figure 6) a much larger effect of congruence in the CSR group (308 ± 31.4 ms, 17.1 ± 1.96%) than in the Tasks group (60 ± 12.1 ms, 4.26 ± 0.83%); the interaction was highly reliable both for RTs, $F(1,30) =$

27.3, $p < .001$, $\eta^2_G = 0.0881$ and in the error rates, $F(1,30) = 18.2$, $p < .001$, $\eta^2_G = 0.0934$. In separate analyses, the congruence effect was reliable for both the Tasks group, $F(1, 15) = 12.5$, $p < .001$, $\eta^2_G = 0.0306$, for RTs, and $F(1,15) = 13.3$, $p < .01$, $\eta^2_G = 0.0802$, for errors, and the CSR group, $F(1,15) = 48.4$, $p < .001$, $\eta^2_G = 0.318$, for RT, and $F(1,15) = 37.9$, $p < .001$, $\eta^2_G = 0.331$, for errors.

Acquisition effects. Overall performance improved from block pair 2 to 5 (Figure 7), and this was reliable for RTs, $F(3,90) = 8.99$, $p < .001$, $\eta^2_G = 0.0287$. Participants in the CSR group learning by trial and error alone had a much higher initial error rate than in Experiment 2, and a steeper learning curve than the Tasks condition, as indicated by a significant interaction between block pair and instructions, $F(1,30) = 16.7$, $p < .001$, $\eta^2_G = 0.00545$.

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 show continuous and 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, congruence and instructions was significant in the errors only, $F(3,90) = 10.0$, $p < .001$, $\eta^2_G = 0.0344$. Separate analyses revealed a highly significant block pair by congruence interaction in the CSR condition, $F(3,45) = 18.5$, $p < .001$, $\eta^2_G = 0.104$, but not in the Tasks group, $F(3,45) = 2.34$, $p = 0.09$, $\eta^2_G = 0.0110$. Figure 6 shows that in the CSR condition performance on the incongruent trials did not approach asymptote.

Other significant interactions. In agreement with Experiment 1 the switch cost was again larger for incongruent trials for the CSR group (39 ms, 3.85%) than congruent trials (19 ms, 0.63%). For the Tasks group the error switch cost was larger for incongruent trials (1.82%) than for congruent:(-0.14%) but the RT error cost was smaller for incongruent (63 ms) than for the congruent trials (85 ms). Overall there was a significant interaction between task switch and congruence in the errors, $F(1,30) = 7.02$, $p < .05$, $\eta^2_G = 0.00422$, but not in the RTs. However, this interaction did not differ reliably between the two experimental conditions in the errors or RTs. The congruence effect was larger in the short (Tasks: 82 ms, CSR: 342 ms) than the long-CSI (Tasks: 39 ms, CSR: 276 ms), $F(1,30)= 9.79$, $p<0.01$, 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, $F<1$.

The transfer test

As can be seen in Figure 6, the perturbation of performance by the introduction of novel stimuli at block pair 6 was more marked and persistent for the CSR group, especially for incongruent stimuli. The effect of transfer was analyzed as for Experiments 1 and 2, although it is worth noting that in this experiment participants were told to treat block pair 6 as practice. As expected, the CSR group was more affected by the transfer to new stimuli; their RT (error rate) increased by 106 ± 35.7 ms (9.18 ± 1.43%) as compared to the Tasks group's 64 ± 13.9 ms (2.18 ± 0.56%). The interaction was significant for the error rate, $F(1,30) = 10.5$, $p < .01$ $\eta^2_G = 0.0377$, but not the RTs, $F(1,30) < 1$.

Performance improved from block pair 6 (Tasks: 715 ms, 5.28%, CSR: 841 ms, 14.2%) to 7 (Tasks: 632 ms, 4.07% CSR: 756 ms, 5.33%) significantly both in the RTs,

$F(1,30) = 24.5$, $p < .001$, $\eta_G^2 = 0.0349$, and in the errors, $F(1,46) = 33.4$, $p < .001$, $\eta_G^2 = 0.0669$.

The CSR group reduced their error rate significantly more than participants in the Tasks group, $F(1,30) = 19.4$, $p < .001$, $\eta_G^2 = 0.0399$— but had much more room for improvement.

Discussion

The CSR group, who learnt only by trial and error, showed a similar pattern of results to the CSR groups in the previous experiments, i.e. much larger congruence effects and smaller switch costs than the Tasks group. Preparation reliably reduced the switch cost in the Tasks group but not in the CSR group (though the interaction was only marginally reliable). In the CSR group, the learning curves for congruent stimuli reached asymptote while the learning curve for the incongruent stimuli did not, indicating that the correct responses were easier to learn for the congruent stimuli. The consistency of these differences across experiments suggests that they are not greatly dependent on the procedure used to induce CSR learning, or procedural details such as the type of cue.

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 block pairs 2 to 5 and the transfer block (and hence the smaller impact of transfer in the CSR condition in this experiment compared to the previous two: they were starting from a higher baseline).

General Discussion

In three experiments with the trial sequences and response contingencies of a task-cuing experiment, with small sets of four stimuli and four task cues, we have shown that quite different patterns of performance are obtained when participants follow standard task-cuing instructions to apply one of two classification rules as specified by the cue, compared to when they are induced to learn and use cue + stimulus → response (CSR) mappings. In all three 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 perturbation of performance, similar for all combinations of congruence and switch/repeat, consistent with a transient adjustment as the existing rules were applied to new stimuli.

The three experiments induced CSR learning in the other group of participants in different ways. In Experiment 1, we merely supplemented standard task instructions with encouragement to learn CSR mappings. In Experiment 2, the CSR group was given a list of the CSR mappings to learn with no mention of tasks. In Experiment 3, the CSR group learnt the mappings by trial and error alone, again with no mention of tasks. In each case the CSR group exhibited a large congruence effect, and a small but 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. The less efficient transfer in the CSR groups highlights one of the key advantages of using a task-set strategy in this paradigm: its ability to deal with novel stimuli.

It is worth commenting on the difference between the results obtained in our transfer test and in Dreisbach, Goshke & Haider (2006)'s Experiment 1. Our CSR participants who were unaware of the task rules showed worse performance in a transfer test than the Tasks participants. In contrast Dreisbach, Goshke & Haider's (2006) participants showed worse performance in a transfer test in their Tasks condition than in their SR condition. The most obvious difference is that Dreisbach et al's stimuli were univalent, so their participants had to re-learn only SR rules, and may well have succeeded in doing so well within the ten presentations per stimulus in their transfer blocks. In our experiments the conflict created by bivalent incongruent stimuli evidently prompted more extended relearning than could be accomplished in the first few exposures.

The "replaced" participants.

It is of interested to consider the participants we replaced because their introspective reports of their strategies did not match the strategies we intended to induce. In Experiments 2 and 3, in the CSR condition 11 out of 43 participants reported having discovered and used a task-set strategy at some point. That there were more such cases in Experiment 3 (7 versus 4) may be because participants in Experiment 2 were given a more prescriptive demonstration of how to approach the experiment — the instruction sheet provided.

Perhaps of more importance for the interpretation of standard task-switching data is the finding that nearly a third of participants (14 out of 46) in the Tasks condition of Experiments 2 and 3 reported having learned and used at least some CSR mappings instead of, or as well as, task-based classification rules. We describe some of their introspections later. Effects of the critical variables for these participants resembled more those observed for the CSR than for Tasks group: large congruence effects (219 ms for RT, 11.8% for errors), relatively modest "switch cost" (61 ms, 3.57%), and no RISC effect (short-CSI minus long-CSI switch cost: -38.3 ms, -0.66%). In task-cuing experiments where CSR learning is relatively easy (e.g. one cue per task, small numbers of stimuli, half of the stimuli being congruent) 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 such small stimulus sets might best be avoided.

The induction of task rules has recently been considered by Collins and Frank (2013). They suggest that people naturally induce tasks even in situations where using such a strategy is actually disadvantageous for performance, as in the case explored by Dreisbach and colleagues. The experiments of Collins and Frank (2013) explore associative structures different from those examined here. However, it is interesting to consider how our findings, that in a situation where neither strategy confers much benefit relative to the other (and if anything using task-sets affords a slight advantage), impacts on their conclusions about why participants induce and use task-sets. They argue that task-sets allow easier generalizability, prevent conflict (cf. Dreisbach et al) and create a 2-stage decision-making process. Here all these things are certainly true, yet some Tasks participants explicitly given task rules still opted not to use them and the majority of CSR participants did not induce task rules. It may be worth considering what advantages using CSR mappings may have – perhaps they are easier to automatize and so require less cognitive effort to use.

We now discuss (i) the implications of our results for accounts of performance in the task-cuing paradigm, (ii) an associative account of performance in the CSR condition, (iii) the implications for associative versus propositional accounts of learning.

Implications for a compound-retrieval account.

The first conclusion we draw is that our overall findings are problematic for the general class of theory which seeks to explain standard task-cuing performance in terms of memory retrieval; Logan et al's (2003) model is the notable explicit example. According to such a theory, in a task-cuing experiment with standard instructions, participants do not switch between classification rule-sets 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, applying a single CSR task-set throughout. In the CSR conditions of our experiments we used a very small stimulus set to make the CSR contingencies easy to learn, and gave instructions and training designed to encourage a compound retrieval strategy. The results indicate that participants can apply a CSR strategy successfully given a reasonable amount of practice, though it may initially be costly in errors. But, when they do apply this strategy, we see a pattern of performance quite different from that seen when participants are given standard task-set instructions. It would appear to follow that participants who are given standard task-set instructions, and produce the patterns of data usually reported in the task-switching literature, cannot be using a CSR strategy, at least of the "episodic" kind, as defined by Arrington & Logan (2004). If they were already using a CSR strategy, telling them or training them to use it should not change the pattern of performance so dramatically.

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-cuing 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-cuing 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, submitted; Monsell & Mizon, submitted) the phenomenology is qualitatively similar to that observed with small stimulus sets.

We temper this argument against compound-retrieval as a general account of task-cuing performance with two observations. First, the results of the CSR condition clearly indicate that people *can* adopt and efficiently use a compound retrieval strategy 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-cuing 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 a compound-retrieval model to allow participants to retrieve the response to the combination not of specific stimulus and cue, but of the stimulus and a mediating representation of the relevant attribute as signaled by the cue (Arrington & Logan's, 2004 "semantic" strategy): essentially the cue asks "is it odd or even?", the stimulus "7" activates both "odd" and "low" and applying both these cues conjointly to memory maximally activates "odd" more than the other response alternatives. This account essentially posits activation by the cue of a representation of the classification task required (Schneider & Logan, 2007), but denies that there is a change in task-set when the task-representation changes: the constant task-set is simply to perform the compound retrieval operation. Is it possible that, in Arrington and Logan's terms, our participants were applying the "semantic" compound retrieval strategy in the Tasks condition, but the "episodic" compound retrieval strategy in the CSR condition?

There would seem to be two kinds of argument against this. First, the rationale would presumably be that in the CSR condition, unlike the Tasks condition, the instructions do not assist the participants to identify the relevant categorical dimensions, and so they tend to use the episodic strategy and learn specific C-S compounds rather than the mediating task labels. But this rationale works only for Experiments 2 and 3. In Experiment 1, CSR participants were given the same task-set instruction as the Tasks participants, but still – apparently – slugged away at learning specific CSR or at least SR associations in spite of the explicit availability to them of the semantic strategy.

The second argument is more fundamental. Both the "semantic" version of compound retrieval and the TSR account invoke representations of the two classification tasks required. In the compound retrieval account, this is generated from the cue and then applied simultaneously with the encoded stimulus to activate responses from memory. In the TSR account it is applied to reconfigure task-set (e.g. bias attention towards the relevant attribute and/or activate/retrieve/select the relevant S-R rules) which is then applied to classify the stimulus. The critical difference between the accounts is in the timing. In the compound-retrieval account, retrieval waits until both task and stimulus representations are ready. In the TSR model, the task-set is reconfigured and then the R corresponding to the S is retrieved. This two-stage two-level property of the TSR account supplies a natural account of the switch cost and its reduction with preparation (the RISC effect): the further ahead the participant can get with task-set preparation before stimulus onset, the better. Either version of the compound retrieval account (semantic or episodic) must rely on the priming of cue-encoding to explain the switch cost and RISC effect, with the facilitation on "task-repeat" trials having qualitatively if not quantitatively identical effects whether the immediately preceding cue is identical, associated, or retrieves the same mediating label (Schneider and Logan, 2005, 2007). Hence if Tasks participants use the semantic compound-retrieval strategy, and CSR participants the episodic compound-retrieval strategy, we should still see a reduction in switch cost with preparation in both conditions. But we did not. Only the Tasks condition showed a RISC effect, consistent with task-set reconfiguration in that condition, and some other strategy in the CSR condition.

An associative account of performance in the CSR condition.

We now examine in more detail the claim that this “other” strategy reflects associative learning and retrieval. Some features of the performance we observed in our experiments are intuitively explicable in these terms. There is a large congruence effect because of associative interference when one element of the cue + stimulus compound is a stimulus associated with both responses. The congruent stimuli are much easier to learn than the incongruent stimuli, as the former require learning of simple stimulus-response associations whereas the incongruent stimuli require learning of biconditional discriminations. The lack of a preparation effect is unsurprising because the cue is equally associated with both responses, so there is nothing to prepare[6]. Less intuitively obvious, however, is the finding of what appears to be a small “switch cost” even when there are no task-representations to switch between.

---- Insert Figure 8 about here ---

However it turns out that a small "switch cost" is generated by an associative learning network with no representation of the underlying tasks. In order to demonstrate this we modeled the data using the most recent version of the APECS model, a three-layer back-propagation localist connectionsist network — see Figure 8, McLaren, (2011) and McLaren, Forrest & McLaren (2012) for more details of the model. The main differences between

---

[6] In fact some participants in the CSR condition in Experiments 2 and 3 said that that they examined the digit first and then looked to the cue for the digits with inconsistent responses.

APECS and a standard back-propagation network are that APECS has adaptive learning parameters which favor the unit best suited for carrying a given mapping, and adaptive biases to preserve this learning, and that it is trained to replicate the input on the output units as well as generate the right response. This gives APECS a more discrete representation at the hidden layer than a standard back-propagation network. The model is able to learn the correct responses for a given cue and stimulus pair based purely on modifying the weights on links between input units, hidden units, and output from feedback, with no control mechanism. We ran a simulation with input units for each of the four cues and eight digits, two output units for the responses and an intermediate layer of 14 hidden units. The network was trained on trial sequences generated in the same way as in the above experiments; 32 networks initialized with random weights were run as pseudo-subjects. No attempt was made to capture the temporal dynamics of a trial (and hence compare short and long CSI conditions); on each trial the combination of cue and stimulus was applied to the network, it generated activation of the output units, and error-correction was applied to modify the connection weights using the APECS algorithm.

As expected the performance of the model was worse on incongruent than congruent stimuli. This difference between incongruent and congruent trials evolved over blocks in a way similar to the results of Experiment 3, as the congruent trials were much easier for the model to learn. Critically, the model's performance was also worse on "task-switch" than "task-repeat" trials (significantly so if we treat the 32 networks as "subjects", F (1,31) = 4.46, $p < 0.05$, $\eta_p^2 = 0.00630$; switch trials 0.063, repeat trials 0.056[7]). This "switch cost" was

---

[7] The performance of the model was examined by comparing the activation difference between the two output units, which corresponded to the two responses, and the target

significantly larger for the incongruent (0.0130) than the congruent trials (0.000599), $F(1,31) = 4.16$, $p = 0.05$, $\eta^2_G = 0.00525$, and was reliable only for the incongruent stimuli, $F(1,31) = 4.99$, $p < 0.05$, $\eta^2_G = 0.0123$; not for the congruent stimuli, $F < 1$. Unsurprisingly performance also suffered when the stimuli were exchanged for a new set of four. In other words the model showed a pattern of data broadly similar to the CSR groups in our experiments, suggesting that "dumb" low-level associative learning is a viable explanation for the pattern of behaviour seen in the CSR conditions.

Why is there an advantage for "task-repeat" trials in the absence of a task-representation? It is somewhat easier to intuit the basis of this advantage for the case where an incongruent stimulus is repeated on successive trials: on a "task-switch" trial, a different response has just been made to that stimulus and the wrong CSR association reinforced, whereas on a "task-repeat" trial the same response is required, and the appropriate CSR association has been reinforced on the previous trial. However, this does not account for all the switch cost seen in the CSR condition: when immediate stimulus repeats were excluded from the data from Experiment 3, there was still a persistent switch cost in the CSR condition, $F(1,15) = 9.784$, $p = 0.007$, $\eta^2_G = 0.00364$ (repeat: 795 ms, switch: 824 ms). But the APECS simulations also deliver a small switch cost in these circumstances. When stimulus repeats were removed from the simulated data there was a significant switch cost for the simulation of Experiment 3, $F(1,31) = 12.1$, $p < .05$, $\eta^2_G = 0.00691$ (repeat: 0.056, switch: 0.063). Moreover, the interaction between task switch and congruence was inconsistent across the

---

difference (target difference-trained difference). On this measure larger scores mean worse performance. The same results were obtained if MSE was used as the dependent measure.

behavioural experiments. The basis of these effects in both the experiments and the model is at present unclear, and further explorations are needed. We speculate that the internal (hidden unit) representations used by the network overlap for different CSR mappings (this is clearly true of our simulation) and that hence interference between trials can take place even when they do not share the same cue and stimulus. What is not clear is why this should favor task repeat trials over task switch trials. It may be that these internal representations do capture the "equivalence" of the two cues used to signal the same task and thus generate a small switch cost — an account similar to Logan and Schneider's (2005) proposal of priming of a mediating task representation. For now, the moral is that the "task-switch" cost in the CSR condition is compatible with an associative account in which there are no task representations as such.

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. An alternative possibility is that participants were learning and using CSR rules, but in no less a "cognitive" or "symbolic" way than the Tasks group construes the situation as one involving switches between two classification rules. Mitchell, De Houwer and Lovibond (2009), for example, have argued that there is no need to postulate an underlying automatic associative system at all; they claim 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 propositional and associative accounts of the CSR group's performance? One argument for the associative account is that a relatively simple connectionist learning model captures important aspects of performance, even aspects that are not obvious consequences of a propositional CSR strategy such as the “switch cost”. Another argument could be drawn from Stoet and Snyder's (2003a, 2003b, 2007) rhesus monkeys, who, after massive practice, showed only small switch costs but a larger congruence effect than human controls who received standard task-set instructions and relatively little practice. Of course, if one is willing to grant 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 — often 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 Experiments 2 and 3, participants in the CSR condition articulated propositional rules. They often noted that two of the numbers (the congruent stimuli) were each linked to only one response, e.g. "8 was always left". They also outlined mappings for the incongruent stimuli, most often by describing one (or two) of them and then explaining how various changes meant the opposite response. In these outlines, in Experiment 3, they would always refer to the two cues which went with each task as being equivalent, e.g. "square and circle with 7 means right but if it’s a triangle or pentagon it’s the opposite response and 4 was the opposite of 7". Explicit use of such rules would evidently be easier for the congruent stimuli as their rules are simpler and easier to learn. 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 discovered what rule structure is required to solve the problem. One test of a rule, by contrast with a simple association, is whether it can be generalized (Murphy, Mondragón & Murphy, 2008) — but in this case a simple associative model can also be shown to produce a degree of generalization, as the same cues are used in the first part of the experiment and the transfer phase. Hence in this case the transfer test is not discriminative between stimulus-specific rules and associations.

Of course the ability to articulate a propositional rule does not mean that performance is being driven by it. It has 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.

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 a sufficiently 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. Rogers & Monsell, 1995). We are pursuing these questions with further experiments.

For the present, we conclude that inducing a CSR strategy in the task-cuing 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-cuing experiments. Performance mediated by CSR learning is broadly consistent with an associative account. But learning and retrieving responses to compounds of cue and stimulus appears, on the whole, neither a natural nor an effective strategy for the standard task-cuing 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.

References

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

Allport, A., & Wylie, G. (2000). 'Task-switching', stimulus-response bindings, and negative priming. In S. Monsell & J. S. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 35–70). Cambridge, MA: MIT Press.

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

Altman, E.M. (2006). Task switching is not cue switching. *Psychonomic Bulletin & Review, 13,* 1016-1022.

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.

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, *37*, 379-384.

Brass, M. & von Crammon, D.Y. (2004). Decomposing components of task preparation with functional magnetic resonance imaging. *Journal of Cognitive Neuroscience, 16,* 609-620.

Brown, S.D., Lehmann, C., & Poboka, D. (2006). A critical test of the failure-to-engage theory of task-switching. *Psychonomic Bulletin & Review, 13,* 152-159.

Collins, A.G.E., & Frank, M.J. (2013). Cognitive control over learning: Creating, clustering and generalizing task-set structure. *Psychological Review, 120,* 190-229.

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.

Dreisbach, G. (2012). Mechanisms of cognitive control: The functional role of task rules. *Current Directions in Psychological Science, 21,* 227-231.

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.

Elchlepp, H., Lavric, A. & Monsell, S. (submitted). A change of tasks delays and modulates early processes in lexical tasks: evidence from ERPs.

Fitts, P. M. (1964). Perceptual-motor skills learning. In A.W. Melton (Ed) *Categories of human learning* (pp. 243-285) New York: Academic Press.

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.

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

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.

Jacoby LL (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.

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.

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,* 446-465.

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., & 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.

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.

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. (2006). What matters in the cued task-switching paradigm: Tasks or cues? *Psychonomic Bulletin and Review, 5,* 794-799.

McLaren, I. P. L. (1993). APECS: A solution to the sequential learning problem. In *Proceedings of the XVth Annual Convention of the Cognitive Science Society* (pp. 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 parameterised model of associative learning and memory. In Alonso, E. & Mondragón, E. (Eds.). *Computational neuroscience for advancing artificial intelligence: Models, methods and applications (pp.??-??).* Hershey, PA: IGI Global.

McLaren, I.P.L., Forrest, C.L. & McLaren R.P. (2012). Elemental representations and configural mappings. *Learning and Behavior, 40,* 320-333

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* (pp. 313-332). London: Academic Press.

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.

Mitchell C.J, De Houwer J., & Lovibond P.F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences, 32*, 183–246.

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.

Murphy, R.A., Mondragón, E., & Murphy, V.A. (2008). Rule learning by rats. *Science, 319,* 1849-1851.

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.

Nobre, A., Correa, A., & Coull, J. T. (2007). The hazards of time. *Current Opinion in Neurobiology, 17,* 465–470.

Olejnik, S., & Algina, J. (2003) Generalized eta and omega squared statistics: Measures of effect size for some common research designs. *Psychological Methods, 8,* 434-447.

Pearce, J.M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review, 94,* 61–73.

Rogers, R.D. & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124,* 297-231.

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.

Rumelhart, D. E., Hinton, G. E. & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature, 323*, 533 – 536.

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.

Shanks, D.R. (2010). Learning: From association to cognition. *Annual Review of Psychology, 61*, 273–301.

Shi, Y., Zhou, X., Muller, H.J. & Schubert, T. (2010). The neural implementation of task rule activation in the task-cuing paradigm: An event-related fMRI study. *NeuroImage, 51,* 1253-1264.

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.

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.

Vandierendonck, A., Liefooghe, B. &Verbruggen, F. (2010). Task switching: Interplay of reconfiguration and interference control. *Psychological Bulletin*, *136,* 601-626.

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.

Yeung, N. & Monsell, S. (2003). Switching between tasks of unequal familiarity: The role of stimulus-attribute and response-set selection. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 455-469.

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.

Appendix

Experiment 1 Instructions

Participants in both conditions were initially given identical instructions explaining that the cue indicated the task, and specifying the response categories for each task. The only difference between the conditions was that the CSR group was shown the following advice after the practice blocks:

TOP TIP!

Only four digits will be used in the following blocks:

<1, 7, 4 and 8>.

Here's a tip that may help you improve your performance.

At the moment, on each trial, you have to work through something like:

"oh, the background's <square>

that's odd/even,

the digit's a 2, that's even,

and so the response must be LEFT."

But, as you practice, you may find that you can gradually learn to respond on the basis of a small set of "rules" or "pairings", including

Square + 2 —> RIGHT

Diamond + 2 —> LEFT

With only 4 digits, there are not many such pairings.

If you can learn them, you won't need to worry so much about which classification you are supposed to be doing!

Onscreen instructions for Experiment 2 and 3

The onscreen instructions were very similar for Experiments 2 and 3 except that in Experiment 2 participants were informed that they would be given a piece of paper reminding them about how to respond for the practice blocks and the cues varied between the experiments (blue/pink squares/diamonds in Experiment 2 and four different shapes in Experiment 3).

In the task condition they were given the following onscreen instructions:

“Welcome! You should have the index fingers of your left and right hands resting on the marked "z" and "/" keys - we will call these the LEFT and RIGHT keys. On each trial of this experiment a digit is displayed on the screen for you to classify by pressing the left or right key. These numbers will appear on circles, triangles, squares and pentagons. The background will determine which of two tasks you are doing. The first task is to classify the digit as ODD or EVEN if the background is a SQUARE or CIRCLE.

If the digit is ODD (e.g. 3), please press the LEFT key.

If the digit is EVEN (e.g. 4), please press the RIGHT key.

The second task is to classify the digit as LOW (less than 5) or HIGH (greater than 5) if the background is a TRIANGLE or PENTAGON.

If the digit is LOW (e.g. 3), please press the LEFT key.

If the digit is HIGH (e.g. 8) please press the RIGHT key.

You should try to respond to each digit as quickly as you can, while avoiding errors. Shortly after you press a key, the next digit will appear. If you press the wrong key, "WRONG" will appear on the screen. The experiment will consist of 14 blocks. The first two blocks are practice blocks for you to try to learn the tasks. Press any key to continue.”

In the CSR condition participants were given the following instructions:

“Welcome! You should have the index fingers of your left and right hands resting on the marked "z" and "/" keys - we will call these the LEFT and RIGHT keys. On each trial of this experiment a digit is displayed on the screen for you to classify by pressing the left or right key. Only four digits will be used in the following blocks 2,3,6,9. These numbers will appear on circles, triangles, squares and pentagons. As you go through the experiment you can gradually learn to respond on the basis of a small set of "rules" or "pairings" for example:

Hexagon + 0 = RIGHT

With only 4 digits, there are not many such pairings. You should try to respond to each digit as quickly as you can, while avoiding errors. Shortly after you press a key, the next digit will appear. If you press the wrong key, "WRONG" will appear on the screen. The first two blocks are practice blocks for you to try to learn the mappings. Press any key to continue."

Table 1

Examples of task pairs with small stimulus sets used in task-cuing experiments

Stimuli	Tasks	Cues	References
8 digits	Classify odd/even or high/low	Words (“parity”, “magnitude”) or arbitrary letters, shapes or colours.	Monsell, Sumner and Waters (2003) Logan and Bundesen (2003) Schuch and Koch (2003)
16 coloured shapes	Identify colour or shape	Words ("colour", "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, centre.	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 coloured 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, colour or form	A double arrow, yellow patch or “$”	Koch (2001)
6 coloured words	Name the word or the colour.	location –top/bottom of screen.	Allport and Wiley (2000)
4 coloured rectangles	Indicate orientation colour	The digits 3, 4, 5 or 6.	Slagter et al (2006)
8 coloured face pictures.	Classify or colour gender	The words "gender" or "colour" 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” coloured squares.	Jamadar, Hughes, Fulham, Michie & Karayanidis (2010)

Figure Legends

Figure 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 level specifies the response assignments for the selected classification. The right hand panel indicates the "flat" structure of the cue + stimulus-> response strategy, as instructed/trained in the CSR condition, with no mention of tasks. These diagrams were used as the instruction sheets for Experiment 2.

Figure 2. The means of median correct RTs and error% over successive block pairs in Experiment 1 for each combination of congruent/incongruent stimuli and "switch"/ "repeat" trials (as specified for the Tasks condition). Darker dotted lines represent switch trials, lighter solid lines 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.

Figure 3. Contrasts indicating three principal findings of Experiment 1. Reaction time contrasts are shown in the upper panels, error rate contrasts in the lower panels. The three panels show the interaction of instruction condition (Tasks versus CSR) with: (left panel) the effects of a task switch [as specified for the Tasks group]; (middle panel) the reduction in switch cost with increasing CSI; (right panel) the effect of response congruence.

Figure 4. The means of median correct RTs and error % over successive block pairs in Experiment 2, plotted as in Figure 2.

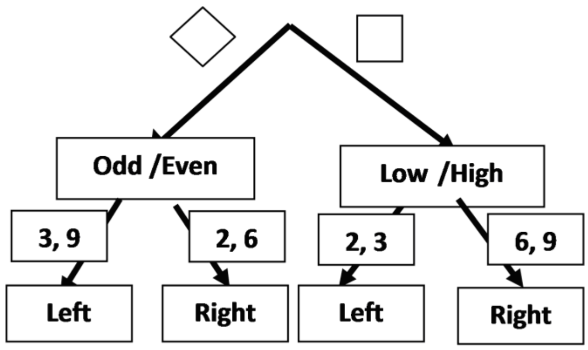
Figure 5. Contrasts indicating three principal findings of Experiment 2, plotted as in Figure 3.

Figure 6. The means of median correct RTs and error % over successive block pairs in Experiment 3, plotted as in Figure 2.

Figure 7. Contrasts indicating three principal findings of Experiment 3, plotted as in Figure 3.

Figure 8. A diagrammatic representation of APECS. Black units represent units activated on an example trial, illustrating the way in which the hidden units could carry the mappings, both to the correct output unit and to the unit replicating the cue on the output layer. The minus sign on the hidden units indicates the negative bias which preserves the learning. Note that more input, hidden and output units were used in the actual simulations.

Figure 1



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

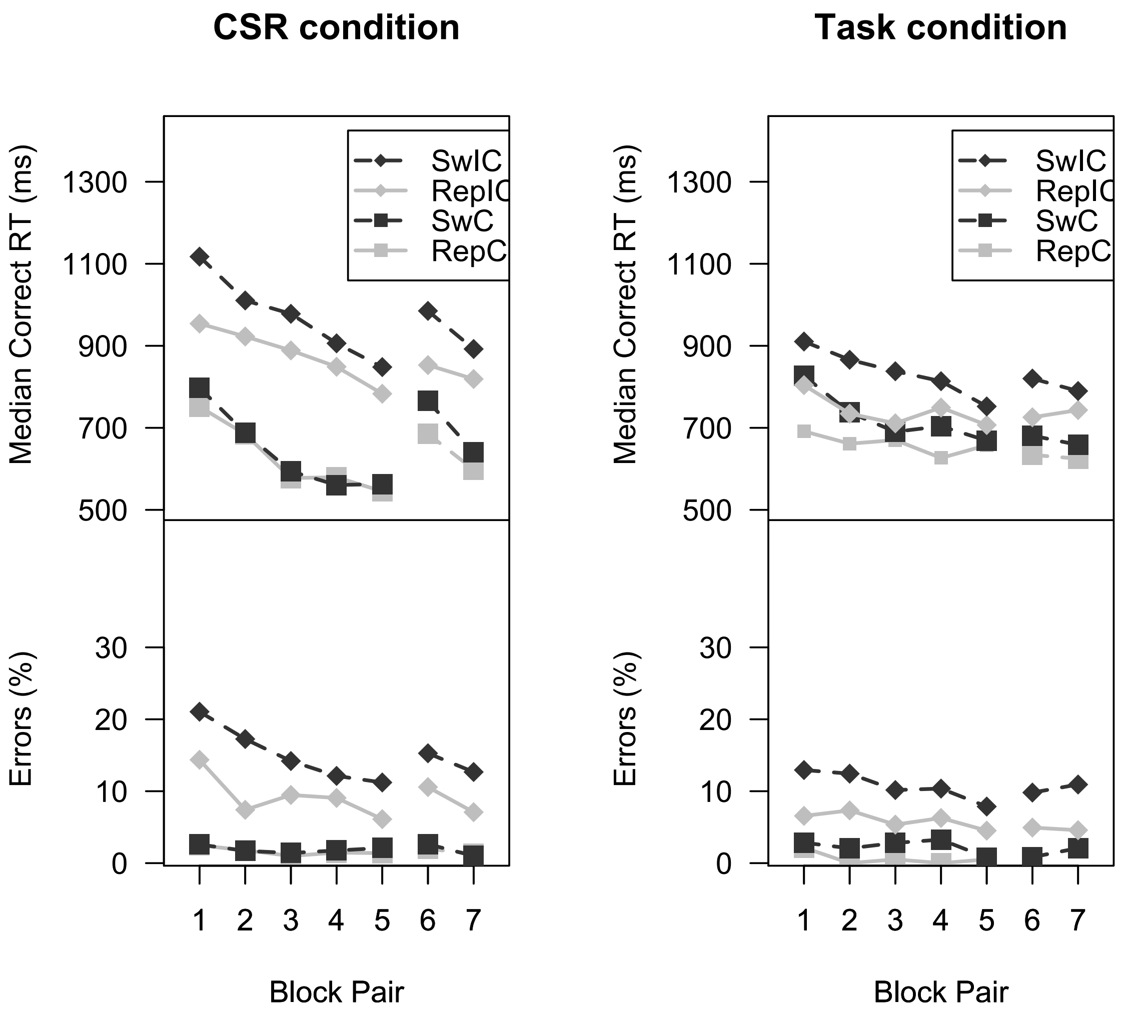

CSR condition
Task condition
Median Correct RT (ms)
1300
1100
900
700
500
SwIC
RepIC
SwC
RepC
Errors (%)
30
20
10
0
1 2 3 4 5 6 7
Block Pair

Figure 3

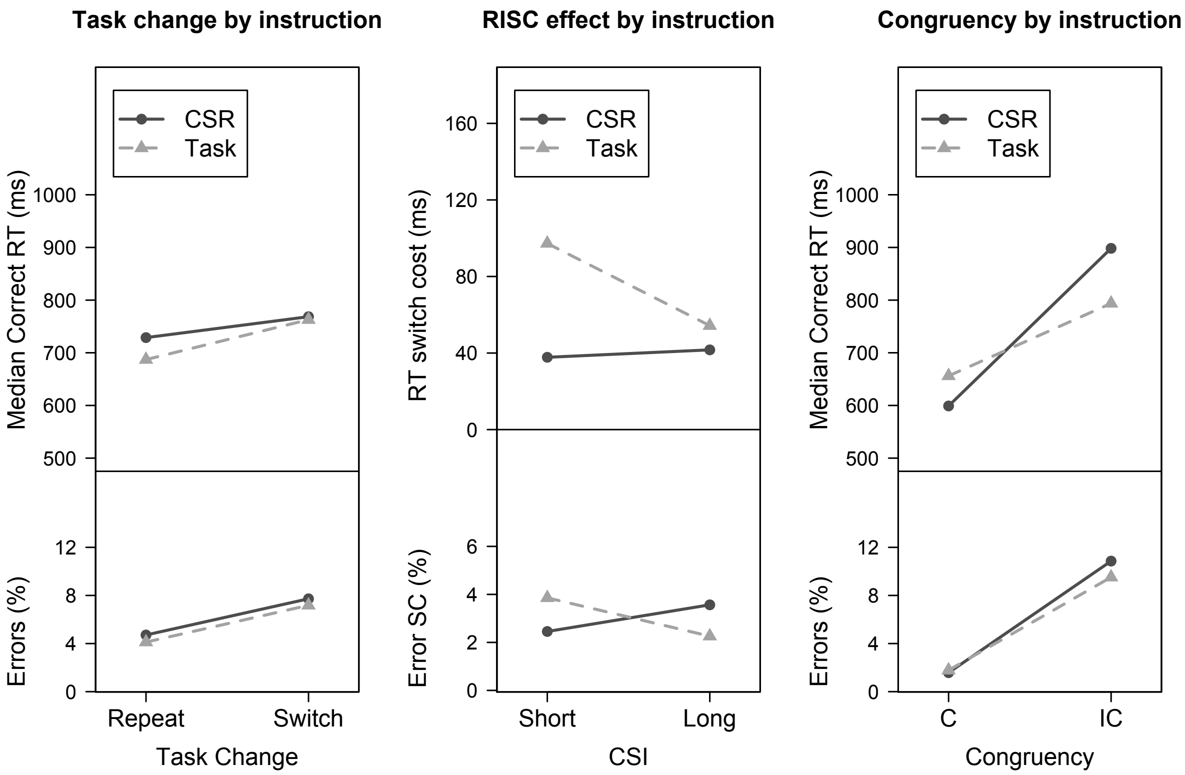

Task change by instruction
RISC effect by instruction
Congruency by instruction
CSR
Task
Median Correct RT (ms)
1000
900
800
700
600
500
Errors (%)
12
8
4
0
Repeat
Switch
Task Change
RT switch cost (ms)
160
120
80
40
0
Error SC (%)
6
4
2
0
Short
Long
CSI
C
IC
Congruency

Figure 4

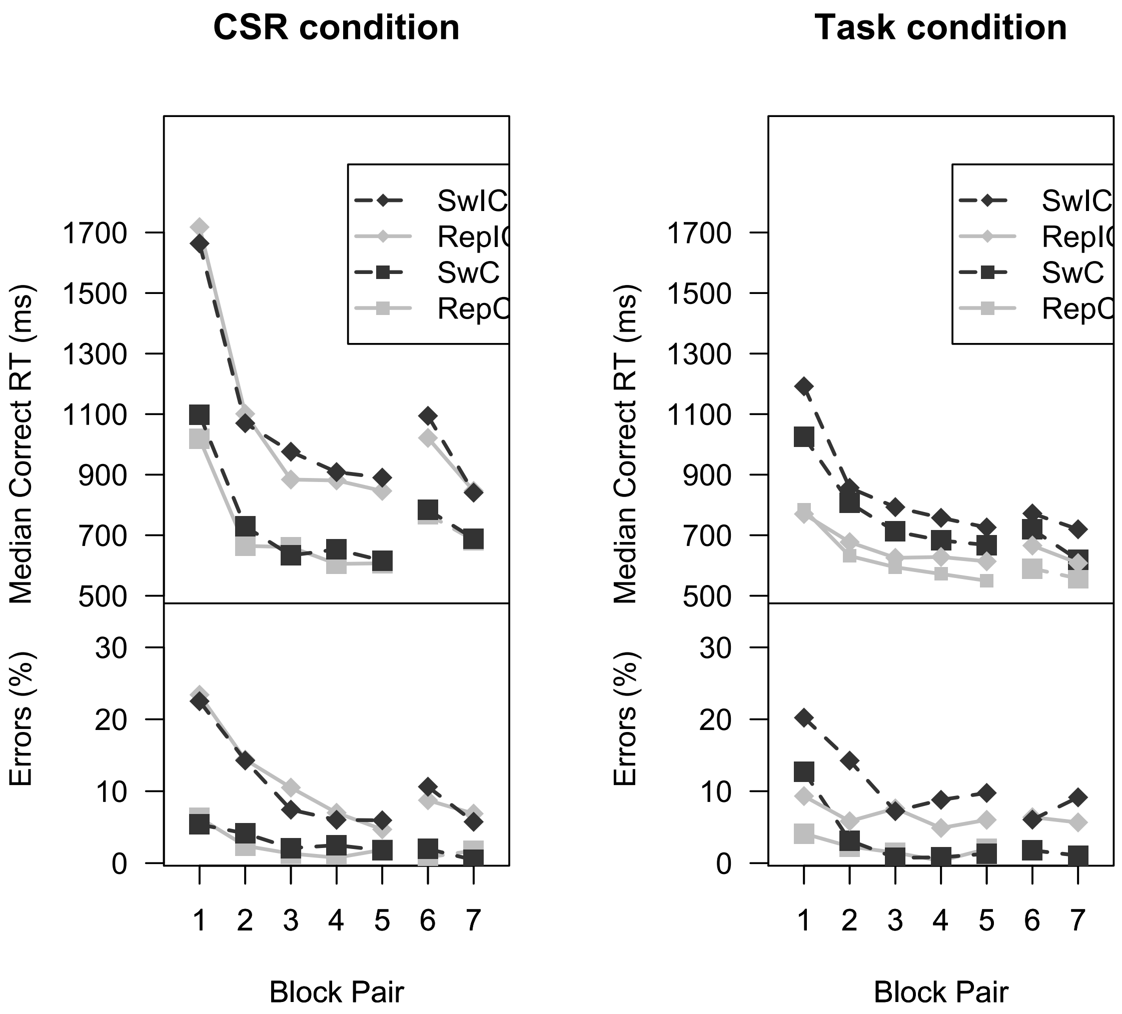

CSR condition
Task condition
SwIC
RepI(
SwC
RepC
Median Correct RT (ms)
1700
1500
1300
1100
900
700
500
Errors (%)
30
20
10
0
1 2 3 4 5 6 7
Block Pair

Figure 5

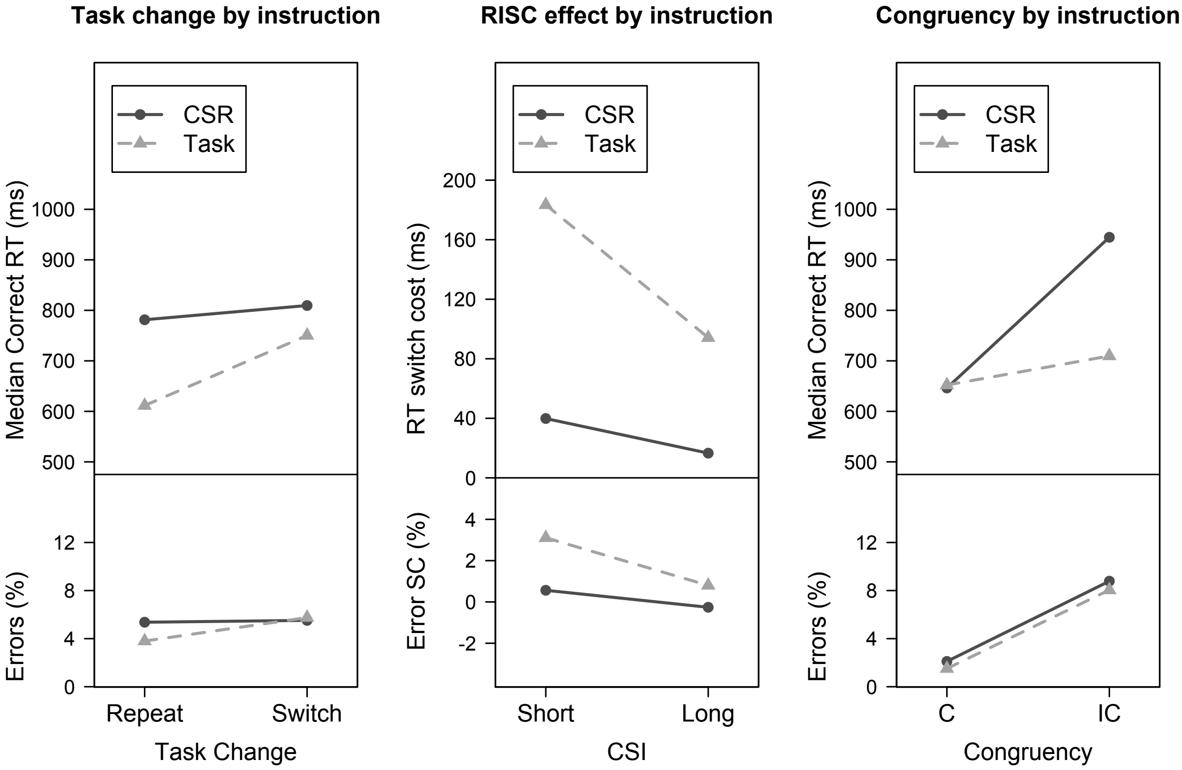

Task change by instruction
RISC effect by instruction
Congruency by instruction
CSR
Task
Median Correct RT (ms)
1000
900
800
700
600
500
Errors (%)
12
8
4
0
Repeat
Switch
Task Change
RT switch cost (ms)
200
160
120
80
40
0
Error SC (%)
4
2
0
-2
Short
Long
CSI
C
IC
Congruency

Figure 6

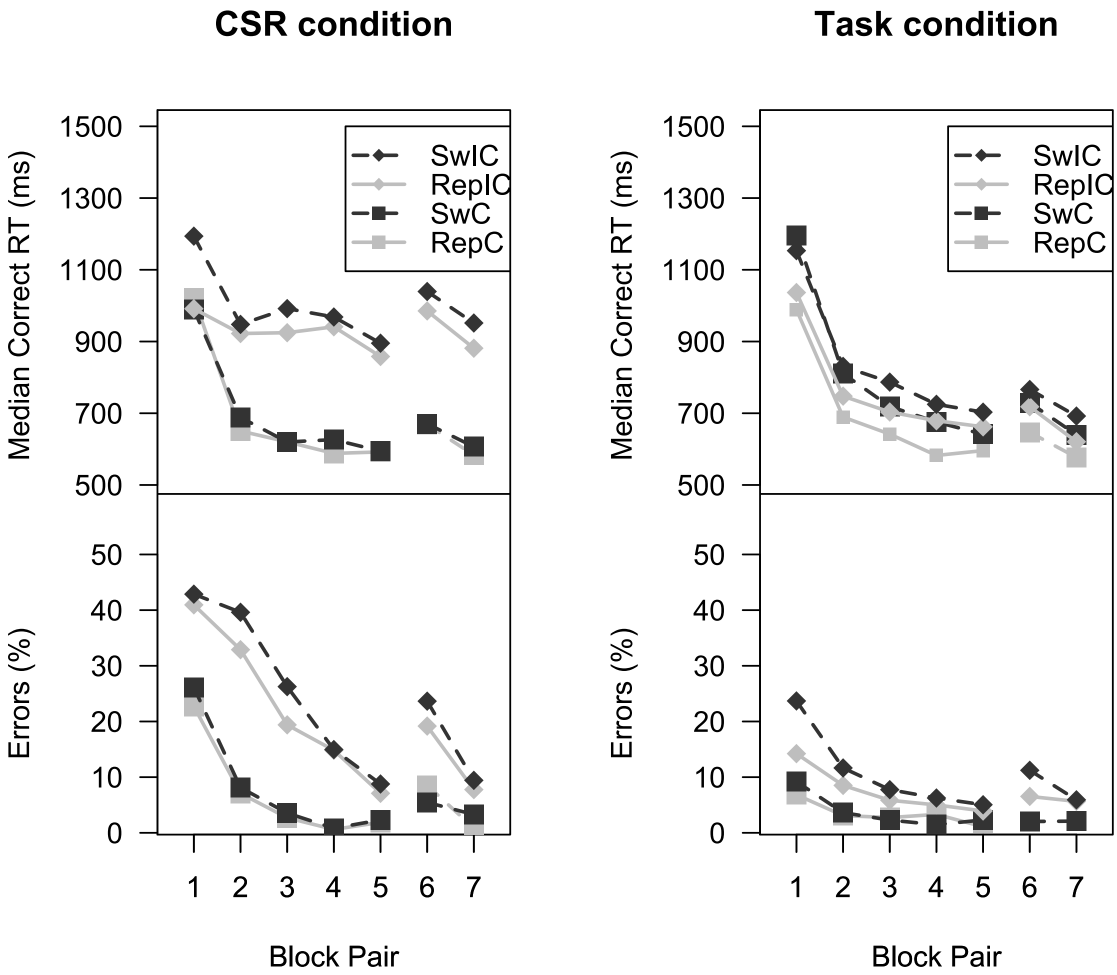

CSR condition
Task condition
Median Correct RT (ms)
1500
1300
1100
900
700
500
SwIC
RepIC
SwC
RepC
Errors (%)
50
40
30
20
10
0
1 2 3 4 5 6 7
Block Pair

Figure 7

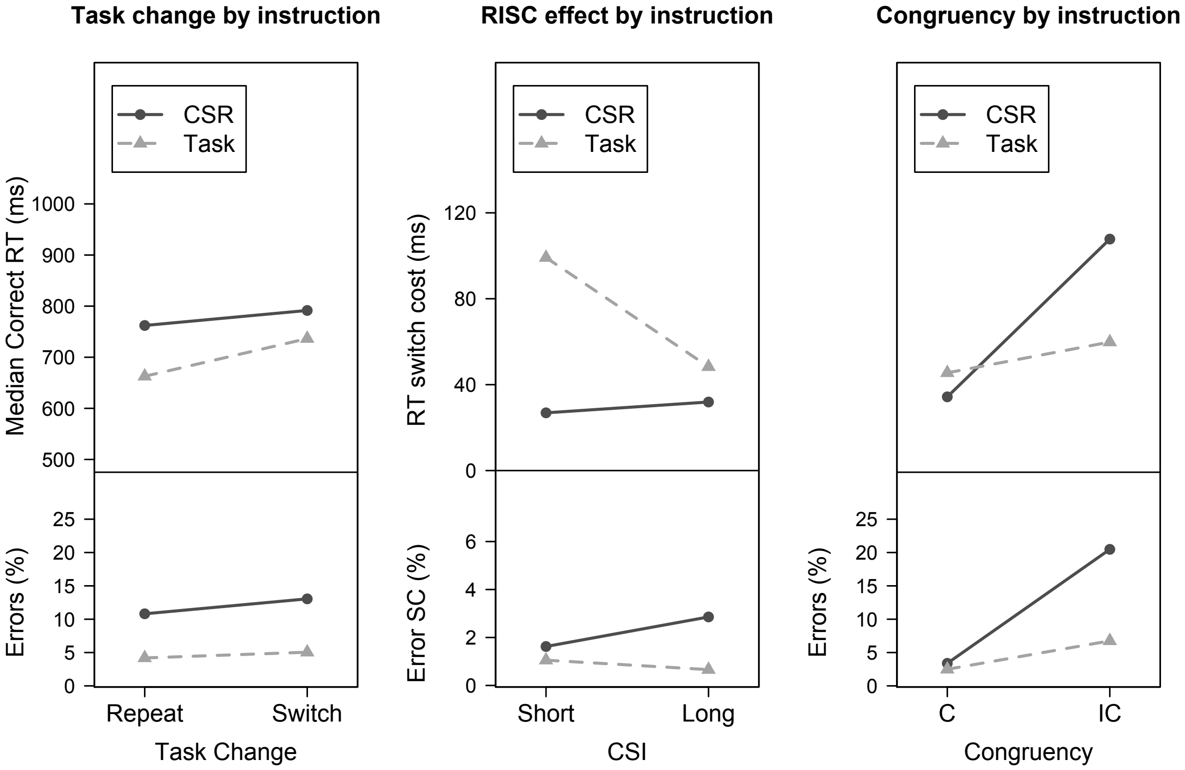
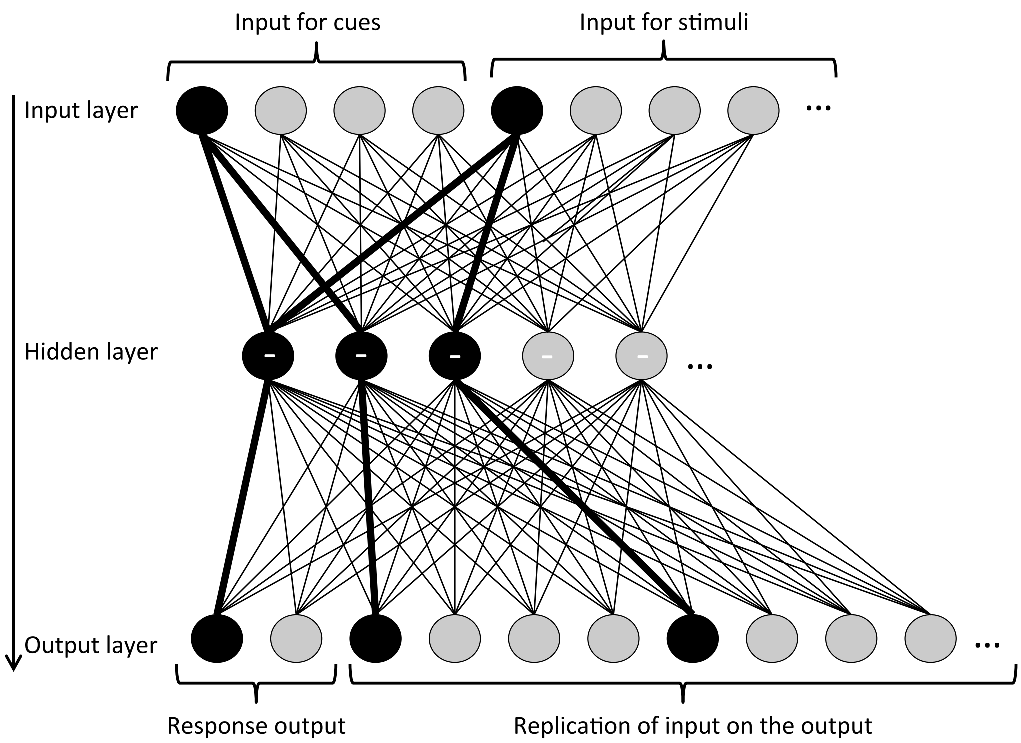

Task change by instruction
CSR
Task
Median Correct RT (ms)
1000
900
800
700
600
500
Errors (%)
25
20
15
10
5
0
Repeat
Switch
Task Change
RISC effect by instruction
CSR
Task
RT switch cost (ms)
120
80
40
0
Error SC (%)
6
4
2
0
Short
Long
CSI
Congruency by instruction
CSR
Task
Errors (%)
25
20
15
10
5
0
C
IC
Congruency

Figure 8


Input for cues
Input for stimuli
Input layer
...
Hidden layer
...
Output layer
...
Response output
Replication of input on the output