Control of stimulus-set and response-set in task switching

1

2

3

Δ

5

James A. Grange¹

¹ School of Psychology, Keele University

Abstract

Successful goal directed behaviour requires not only selecting the correct response to an object in our environment, but also requires selecting the correct *object* in our environment upon which to act. Whilst most task switching studies investigate the selection and maintenance of mental representations of response options (so-called *response-sets*) they often do not investigate the selection and maintenance of mental representations of object selection (so-called *stimulus-sets*). In the current study participants were exposed to a task switching paradigm with multiple stimuli in which the relevant stimulus-set (i.e., which object to respond to) and responseset (i.e., how to respond to that object) independently either repeated or switched on each trial. Of interest was the nature of the task-set representation required, and whether response-set and stimulus-set could be updated independently. Guided by predictions from a computational model of dual-task control (ECTVA; Logan & Gordon, 2001), seven experiments were conducted that evaluated the independence of task-set components. All experiments confirmed ECTVA's predictions of an under-additive interaction between response-set and stimulus-set sequence—diagnostic of independent and parallel reconfiguration of components. However, limitations to this independent updating were observed when participants were encouraged to selectively prioritise response-set or stimulus-set reconfiguration via component-specific preparation manipulations. The results are discussed in terms of various hypotheses on the structure of task-set representation.

Keywords: Task switching; task-sets; ECTVA Word count: 17,025 (not including Appendices)

Correspondence concerning this article should be addressed to James A. Grange, School of Psychology, Dorothy Hodgkin Building, Keele University, Keele, ST5 5BG, United Kingdom. E-mail: grange.jim@gmail. com

6

48

Introduction

Humans live in a rich, multi-task environment where at any one moment we are 7 surrounded by objects upon which we could act: For example, sat in an office there are 8 many objects competing for our attention, such as our computer, our smartphone, papers 9 on our desk, and the coffee machine. The complexity of our environment is compounded by 10 the fact that there are often many tasks one can perform on a single object; that is, stimuli 11 are often multivalent. For example, on a smartphone you can browse the internet, take 12 a photo, play chess, or even sometimes make a phone call. So, at any one moment there 13 are myriad objects and tasks one could act upon, and in order to achieve goal-directed 14 behaviour, cognitive control processes must operate efficiently to prioritise task-relevant 15 objects and actions to those objects. 16

The current study is interested in understanding the mechanisms by which the cog-17 nitive system effectively coordinates the selection of objects and actions to those objects 18 to accomplish goal-directed behavior by utilising a task switching paradigm with multiple 19 stimuli. In many task switching designs, participants are presented with single multivalent 20 stimuli (e.g., digits) and a task cue informs the participant which of two (or more) tasks 21 should be performed. For example, the cue "odd/even" informs the participant to judge 22 whether the stimulus is odd or even, and the cue "low/high" informs the participant to 23 judge whether the stimulus is lower or higher than five. Participants make their response 24 (e.g., press "Z" for "lower" or "odd", and "M" for "higher" or "even") as quickly and as 25 accurately as possible. The main manipulation in these designs is the task sequence: On 26 some trials the task will repeat from that of the previous trial (e.g., odd/even to odd/even) 27 and on other trials the task will switch (e.g., odd/even to low/high). It is a well-replicated 28 finding that task switches lead to slower and more error-prone responses than task repeti-29 tions (Allport, Styles, & Hsieh, 1994; Rogers & Monsell, 1995), a phenomenon referred to as 30 the "switch cost" (see Grange & Houghton, 2014; Kiesel et al., 2010; Koch, Poljac, Müller, 31 & Kiesel, 2018; Logan, 2003; Monsell, 2003; Vandierendonck, Liefooghe, & Verbruggen, 32 2010, for reviews). 33

Successful performance in such paradigms is thought to rely on the formation of 34 a so-called "task-set". Although definitions of task-set differ among researchers (see e.g., 35 Schneider & Logan, 2007, 2014), they can be thought of as "... a set of representations and 36 processes capable of performing a task" (Schneider & Logan, 2014, p. 29). Researchers 37 have explained the contribution of task-sets to the switch cost in a number of ways: as 38 reflecting time-consuming reconfiguration of the relevant task-set that is required on task 39 switch trials but not on task repetition trials (Monsell, 2003; Rogers & Monsell, 1995); as 40 reflecting carry-over effects of the active task-set from previous trials, which can interfere 41 (on switch trials) or facilitate (on repetition trials) establishing the task-set relevant for 42 the current trial (Allport et al., 1994; Altmann & Gray, 2008; Meiran, Chorev, & Sapir, 43 2000); or as a combination of reconfiguration and interference (Meiran, 2000; Meiran et al., 44 2000; Vandierendonck et al., 2010). Others have explained the switch cost as not reflecting 45 influences of the task-set at all, but rather due to priming of cue encoding on task repetition 46 trials (Logan & Bundesen, 2003; Schneider & Logan, 2005; but see Monsell & Mizon, 2006). 47

The paradigm used in the current study (see Figure 1) presents participants with

3

⁴⁹ multiple stimuli (e.g., a digit in orange and a digit in purple) and a task cue (e.g., "odd/even" ⁵⁰ in orange font) that provides two independent pieces of information: (a) which stimulus to ⁵¹ select on the current trial, and (b) which task to perform on that stimulus. As such, the ⁵² established task-set must hold representations of the object and the action required to ⁵³ that object. Borrowing from the classical attention literature, the mental representation of ⁵⁴ which stimulus to prioritise is referred to as the *stimulus-set*, and the mental representation

⁵⁵ of which action or task to perform is referred to as *response-set* (Logan & Gordon, 2001).¹

⁵⁶ From trial-to-trial, the relevant stimulus-set and the relevant response-set can independently

⁵⁷ either repeat or switch from that of the previous trial.



CURRENT TRIAL

Figure 1. Overview of the factorial manipulation of response-set and stimulus-set sequence used as a framework for all experiments reported in the current study. Note that images are not to scale, and the colours used in the individual experiments differ.

The question of interest in the current study is the nature of the task-set representation required for successful performance in such a task, and—more specifically—whether

¹Some authors use the terms stimulus-set and response-set differently. For example, Meiran (2000) presented participants with a single stimulus that could appear in any one of four quadrants. The tasks participants were required to switch between were a vertical judgement (i.e., judging whether the stimulus was in either the upper- or lower-half of the quadrant) and a horizontal judgement (i.e., judging whether the stimulus was in the left- or right-half of the quadrant). Response keys were assigned such that one key reflected an "upper" or "left" response (depending on task), and a second key reflect "lower" or "right". In the terminology of Meiran (2000), *stimulus-set* referred to representing an upper-left stimulus as "mainly upper" (e.g., if the task were "vertical"), and *response-set* referred to representing the "up & left" response key as representing "mainly upper". In the terminology of ECTVA (Logan & Gordon, 2001), representing an upper-left stimulus as "mainly upper" is achieved via the *response-set* (i.e., the representation of which action or task to perform on a chosen stimulus), and representing an upper-left response-key as "mainly upper" is achieved of all response categories assigned to that key (see Equation 4 in the main body of the current paper).

response-set and stimulus-set are unique control parameters that can be updated indepen-60 dently. Some previous work has explored switching of stimulus-set alone (e.g., Gopher, 61 Armony, & Greenshpan, 2000; Rushworth, Passingham, & Nobre, 2005) and switching of 62 response-set alone (e.g., Rushworth, Passingham, & Nobre, 2002), but fewer studies have 63 examined response-set and stimulus-set switching within a single design.² One notable 64 exception is Kieffaber, Kruschke, Cho, Walker, and Hetrick (2013), who presented partici-65 pants with a stimulus display comprising two coloured shapes, and participants were cued 66 on each trial to attend to one of three stimulus dimensions (i.e., stimulus-sets) of shape, size, 67 or colour. Importantly, the response-sets for the three dimensions differed: Two dimensions 68 required a "same/different" response, whereas the third dimension required an identifica-69 tion response. For example, if cued to attend to shape or size, participants were required 70 to judge whether the two presented stimuli were the same or different on the cued dimen-71 sion. However, if cued to attend to "colour", participants were required to judge whether 72 the two stimuli were red or blue. In such a design, the response-sets are shared between 73 the stimulus-sets of "shape" and "size" (i.e., both require a same/different judgement) but 74 are different for the stimulus-set of "colour". Therefore, a switch from "shape" to "size" 75 requires only a switch of stimulus-set (i.e., response-set repeats), but a switch from "shape" 76 to "colour" requires both a switch of stimulus-set and response-set. 77

Across three experiments, Kieffaber et al. (2013) provided behavioural (via switch 78 costs and between-task interference effects in Experiments 1 & 2) and neural (via event-79 related potentials in Experiment 3) evidence of a dissociation between switches of just 80 stimulus-set (i.e., shape to size) and switches of both response-set and stimulus-set (i.e., 81 shape to colour), suggesting stimulus-set and response-set are independent components 82 of task-set that independently contribute to switch costs and between-task interference. 83 However, in this design note that it was not possible to evaluate the cost of a switch of 84 just response-set, as a switch of response-set in this design also necessitated a switch of 85 stimulus-set. 86

The current study overcomes this limitation as response-set and stimulus-set se-87 quences are manipulated factorially. To my knowledge, only one study has explored switch-88 ing of response-set and stimulus-set in a factorial design. Kleinsorge (2004) presented digits 89 on the left or right side of the screen in red or green, with varying quantities up to eight. 90 The relevant response-set was cued by the position of the stimuli (e.g., if on the left, per-91 form odd/even judgement; if on the right, perform a lower/higher than 5 judgement). The 92 relevant stimulus-set was cued by the colour of the digits; if the digits were red, partici-93 pants applied their judgement to the numerical value of the digits, and if the digits were 94 green participants applied their judgement to the number of digits presented. On each trial, 95 either the response-set (odd/even vs. low/high) or stimulus-set (numerical value vs. num-96 ber of digits) either repeated or switched from the previous trial. Results showed a cost 97 of switching both response-set and stimulus-set independently. However, a switch of both 98

²Some paradigms used in previous task switching research requires the control of both response-sets and stimulus-sets, but these are often confounded. For example, in the seminal study of Rogers and Monsell (1995) participants were presented with two stimuli (e.g., "3F") and had to switch between judging whether the number was odd or even, and whether the letter was a vowel or a consonant. In this paradigm, task switches are confounded with object switches, and thus response-set and stimulus-set switching are confounded.

stimulus-set and response-set led to shorter response times than when stimulus-set switched and response-set repeated. Kleinsorge interpreted this pattern as supporting a hierarchical organisation of task-set components (see the hierarchical view of task-set organisation discussed below).

¹⁰³ Component Switching & Task-Set Organisation

¹⁰⁴ Switching of either response-set or stimulus-set requires an updating of that particular ¹⁰⁵ component of the overall task-set. Such component-switching has been studied previously, ¹⁰⁶ but there is currently theoretical disagreement as to the nature of the task-set representation ¹⁰⁷ needed for multi-component tasks and the effect component-switching has on performance.

Some authors propose that a switch in any component leads to an updating of all com-108 ponents of the task-set. For example, Vandierendonck, Christiaens, and Liefooghe (2008) 109 presented participants with Stroop-like stimuli (e.g., four repetitions of the digit 6; Exper-110 iments 1, 2, & 4) or global-local stimuli (e.g., the digit 6 repeated and organised to look 111 like the digit 4; Experiment 3) and they had to perform either an odd/even or lower/higher 112 judgement. Participants were cued on each trial to attend to a single dimension of the stim-113 ulus display (e.g., global vs. local) and to perform one of the two tasks on that dimension 114 (e.g., odd/even). On each trial, the relevant dimension and task could independently either 115 repeat or switch from that of the previous trial. Across four experiments, Vandierendonck 116 et al. (2008) found that a switch in a single component (either the dimension or the task) 117 led to an equivalent performance cost than if both components switched (see Figure 2A). 118 The authors suggest that this pattern of data is diagnostic of a *flat* task-set organisation 119 which requires complete reconfiguration when any component requires updating (for a sim-120 ilar pattern of data, see Allport et al., 1994; and the "linked" condition of Hübner, Futterer, 121 & Steinhauser, 2001). 122

Similar to the flat organisation view is the *integrated* view of Philipp and Koch (2010) 123 in which task-set components become integrated into a single representation, such that a 124 switch in either component leads to a complete updating of the whole task-set. That 125 the whole task-set is reconfigured when a single component changes is similar to the flat 126 organisation view, but the integrated view differs in that it proposes the components are first 127 integrated into a single representation before response selection. Philipp and Koch (2010) 128 argue that such an integration of task-set components should lead to an *under*-additive 129 pattern of switch costs where the cost of switching both components is less than the sum of 130 switching a single component (see Figure 2B). Across three experiments that independently 131 varied the task-set components of task (e.g., odd/even vs. lower/higher judgements) and 132 response modality (vocal vs. manual responses), Philipp and Koch (2010) consistently found 133 such under-additive interactions. 134

Such an under-additive interaction cannot be predicted by the *componential* view of the organisation of the task-set (called the "agglomerated" view by Rangelov, Töllner, Mueller, & Zehetleitner, 2013), where task-set components are represented—and therefore can be updated—independently. This view predicts that a switch of both components is more costly than a switch of either component in isolation, and—in contrast to the integrated view—the costs are additive (see Figure 2C). Such a pattern was observed by



Component 2 - repetition - switch

Figure 2. Schematic predictions from each of the four main hypotheses of the structure of task-sets during component-switching. Each panel shows the factorial combination of the sequence of Component 1 of a task-set (repetition vs. switch) and Component 2 of a task-set (repetition vs. switch). See text for details.

Hübner et al. (2001) using global-local stimuli where participants switched between the components of stimulus dimension (global vs. local) and task (magnitude vs. parity judgements), although in some cases the cost of switching both components was *greater* than the sum of switching either component individually (cf., the prediction of the integrated and flat view) indicative of an additional processing cost due to the coordination of joint updating.

A final view of the organisation of the task-set is the *hierarchical* view (Kleinsorge, 147 2004; e.g., Kleinsorge & Heuer, 1999; Kleinsorge, Heuer, & Schmidtke, 2002, 2004) which 148 suggests components of the task-set are organised such that a switch of higher-level compo-149 nents of the hierarchy leads to complete reconfiguration of the whole task-set, but a switch 150 to lower-level components of the hierarchy only leads to reconfiguration of that level. In 151 several studies, Kleinsorge and colleagues had participants switch between different tasks 152 (e.g., numerical vs. spatial judgements) and different judgement-response mappings (com-153 patible vs. incompatible). Such a paradigm—the authors argue—promotes a hierarchical 154 organisation of the task-set components because the judgement-response mappings (the 155 lower level of the hierarchy) depend on the task (the higher level of the hierarchy). As such, 156 a switch in the judgement-response mapping component would not require updating of the 157 task component, but a switch in the task component would necessitate a change in the 158

¹⁵⁹ mapping component. This organisation makes a striking prediction (see Figure 2D) that ¹⁶⁰ was confirmed in several experiments by Kleinsorge and colleagues whereby a switch in both ¹⁶¹ components is less costly to performance than a switch in just the dominant, higher-level, ¹⁶² component; the hierarchical organisation predicts such a pattern because a switch in the ¹⁶³ higher component leads to a switch also for the lower component, and this switch at the ¹⁶⁴ lower level needs to be reversed in the case of a repetition at the lower level.

165 Executive Control of the Theory of Visual Attention (ECTVA)

The current study approaches the question of the nature of task-set representations 166 in multi-component switching by grounding it within the framework of a theoretical model 167 of executive control in multi-task situations: the Executive Control of the Theory of Vi-168 sual Attention (ECTVA, Logan & Gordon, 2001; which builds upon the Theory of Visual 169 Attention, TVA, Bundesen, 1990). ECTVA provides a rich theoretical framework where 170 components of the task-set—such as the parameterisation of the response-set and stimulus-171 set and their contribution to response selection—are formalised in mathematical equations 172 which make clear predictions which can then be tested against empirical data. In addition, 173 elements of the ECTVA theory have been adapted to model task switching performance 174 in a number of studies (e.g., Logan & Schneider, 2010; Schneider & Logan, 2005), and as 175 such situating the current study within this theoretical framework provides opportunity for 176 cumulative theory development. 177

ECTVA is clear that response-set and stimulus-set are *independent* task-set parameters that can be updated in isolation. Despite this independence, in this section I show that ECTVA produces a pattern of data that have previously been interpreted as representing *non-independence* of task-set components, specifically the *integrated* view of Philipp and Koch (2010).

ECTVA states that each stimulus provides bottom-up sensory evidence for each re-183 sponse category in the set of all possible response categories. In the design reported in the 184 current research, the possible response categories are odd, even, low, and high The amount 185 of evidence provided is proportional to the similarity between the stimulus and a repre-186 sentation of each response category in memory. For example, the digit 3 provides strong 187 evidence for the response categories odd and low, and weak (but non-zero) evidence for the 188 response categories even and high This evidence, η , is then multiplied by the system's bias, 189 β , toward each response category (i.e., the response-set). In the current research, the bias 190 is determined by the relevant response-set cue. For example, if the response-set cue was 191 "odd/even", then both β_{odd} and β_{even} would be set to high, and both β_{low} and β_{high} would 192 be set to low. 193

¹⁹⁴ However, in the current research there are two stimuli presented to participants (for ¹⁹⁵ example, one orange digit and one purple digit), and participants use information from ¹⁹⁶ the stimulus-set cue to ascertain which stimulus to attend to and perform the relevant ¹⁹⁷ categorisation upon. ECTVA handles such situations by establishing an attention weight ¹⁹⁸ to each stimulus in the display. An important component of the attention weight is a ¹⁹⁹ priority parameter, π , which represents the top-down bias of selecting stimuli that have ²⁰⁰ features in the stimulus-set. For example, if the cue is in orange, then π_{orange} would be

set to high, and π_{purple} would be set to low. This is then combined multiplicatively with the sensory evidence, η , that each stimulus belongs to each colour in the stimulus-set to provide an attention weight to each stimulus in the display (for example, an orange digit provides strong evidence for the stimulus-set "orange", and weak but non-zero evidence for the stimulus-set "purple").

Bringing these together, the overall processing rate of categorising stimulus x as belonging to category i, v(x, i), is the multiplicative combination of (a) the sensory evidence that x belongs to category i, $\eta(x, i)$; (b) the top-down bias to category i, β_i ; and (c) the relative attention weight³ to stimulus x in the stimulus display D, which is influenced by attentional priority, π_k , to stimulus feature k:

$$v(x,i) = \eta(x,i)\beta_i \frac{w_x}{\sum_{z \in D} w_z},\tag{1}$$

211 where

$$w_x = \sum_{k \in S} \eta(x, k) \pi_k.$$
⁽²⁾

Here, $\eta(x,k)$ is the sensory evidence that stimulus x has characteristic k in the stimulus-set S. The probability of choosing response category i given stimulus x is also influenced by the presence of the irrelevant stimulus y, and is given by

$$p(i|x) = \frac{v(x,i) + v(y,i)}{\sum_{z \in D} \sum_{j \in R} v(z,j)}.$$
(3)

As the response categories are arbitrarily mapped onto experimental response keys, the probability of selecting response key a, $p(R_a)$, is the sum of the response category probabilities in the set M of response categories mapped to that key,

$$p(R_a|x) = \sum_{j \in M} p(j|x).$$
(4)

These rates and probabilities are then used to determine response selection speed and accuracy via a random walk process. Full details of the random walk process (and further mathematical details of ECTVA) are given in Appendix A.

 $^{^{3}}$ Note that in this Equation, the parameters representing response-set and stimulus-set are used in the same Equation by TVA. This could be considered a form of integration, but it differs from the view of Philipp and Koch (2010) to the extent that there is no explicit binding of these parameters into a single unified task-set representation, and they can be updated independently. I return to discussion of this distinction in the General Discussion.

9

Reconfiguration in ECTVA. In ECTVA, cognitive control acts to set the relevant 221 response-set parameters (i.e., β_{odd} , β_{even} , β_{low} , and β_{high}) and the stimulus-set parameters 222 (i.e., π_{orange} and π_{purple}) for the current trials. If the relevant response-set or stimulus-set 223 (or both) changes from the previous trial, the relevant parameters must be reconfigured 224 which takes time. Importantly, the reconfiguration of each parameter is assumed to be 225 independent and can therefore be executed in parallel. As such, the total reconfiguration 226 time is determined by the slowest parameter to be reconfigured, and the mean total re-227 configuration time gets longer as the number of parameters to be reconfigured increases. 228 This is because, as the number of parameters increases, the probability that at least one 229 parameter will have a longer reconfiguration time than the reconfiguration time for a single 230 parameter increases, thereby increasing the mean total reconfiguration time. 231

As reconfiguration time of a single parameter is assumed in ECTVA to be distributed exponentially with rate parameter v, the expected time, T, taken to switch n parameters is given by

$$T = \frac{1}{v} \cdot \left(\sum_{i=1}^{n} \frac{1}{i}\right).$$
(5)

By this equation, the increase in reconfiguration time is a negatively accelerated 235 function of the number of parameters to be reconfigured (Figure 3A). It is this property 236 which explains the under-additive interaction between response-set sequence and stimulus-237 set sequence in the design of the current research: (1) when there is a repetition of both 238 response-set and stimulus-set, zero parameters require reconfiguration; (2) when there is 239 a repetition of response-set but a switch of stimulus-set, two parameters require reconfig-240 uration; (3) when there is a repetition of stimulus-set but a switch of response-set, four 241 parameters require reconfiguration; and (4) when there is a switch of both response-set and 242 stimulus-set, six parameters require reconfiguration. The finishing times for these number 243 of parameters are highlighted on Figure 3A, and rearranged in Figure 3B to map on to the 244 factorial combination of response-set and stimulus-set sequence as in the current study. As 245 can be seen, an under-additive interaction pattern arises due to the non-linear increase in 246 reconfiguration time with increasing number of parameters to be reconfigured.⁴ 247

The under-additive interaction predicted by ECTVA is most similar to the pattern expected by the *integrated* view of Philipp and Koch (2010), but differs in that ECTVA

⁴That independent reconfiguration of task-set components leads to an under-additive interaction is perhaps not intuitive as it is not in line with typical expectations about parallel or independent processes, which would typically predict that reconfiguration of both components should not increase response time above that of a switch in either component alone. In ECTVA the under-additive interaction arises due to the non-linear increase in reconfiguration time with increasing number of parameters, and this in turn arises due to the exponentially distributed finishing times of individual parameter reconfiguration. As such, total reconfiguration time of all parameters is determined by the slowest parameter to reconfigure. By analogy we can think of parameters as individual (i.e., independent) runners in a race, and total race time (i.e., total reconfiguration time) is determined by the time it takes for all runners to cross the line; if the expected finishing time of each runner is described by an exponential distribution, then the total finishing time of all runners increases non-linearly as the number of total runners increases (by Equation 5, visualised in Figure 3A).



Figure 3. A. The increase in reconfiguration time with increasing number of parameters to reconfigure as given by Equation 5. Note that the arrows and labels a, b, c, and d show how these reconfiguration times map onto the current factorial design, shown in panel B assuming four β values. B. The same reconfiguration times labelled in panel A arranged as the factorial combination of response-set sequence (repetition vs. switch) and stimulus-set sequence (repetition vs. switch). C & D. The same as panels A & B but assuming categorical representations of response-set (i.e., with two β values).

predicts response-set switches in isolation (point c in Figures 3A & B) are more costly than 250 stimulus-set switches in isolation (point b in Figures 3A & B) because more parameters 251 need updating in the former scenario. The original ECTVA model of Logan and Gordon 252 (2001) had separate bias parameters for each response category and as such I implemented 253 the same choice, but there is nothing within the assumptions of TVA or ECTVA that 254 says one also could not model bias parameters as influencing higher-level categories (e.g., 255 "magnitude"; G. Logan, personal communication, June 14 2023). This reduces the number 256 of β parameters to two (i.e., $\beta_{magnitude}$ and β_{parity}), which then reproduces the pattern of 257 the integrated view (see Figures 3C &D). Whether response-set is represented at the level 258 of response options (i.e., four β values) or at the level of response categories (i.e., two β 259 values) is a point I return to in the General Discussion. 260

Compatibility. An important consideration in the current study is the impact of 261 the processing rate of the response categories (Equation 1) associated with the *irrelevant* 262 stimulus (called "crosstalk" by Logan & Gordon, 2001). As is explicit in Equation 3, the 263 probability of choosing response category i is influenced by both the processing rate of 264 choosing category i given stimulus x, v(x,i), and the processing rate of choosing category 265 i given stimulus y, v(y, i). When both stimuli provide evidence for response category i, 266 the numerator of Equation 3 increases and the probability of choosing response category i267 increases, leading to faster and more accurate response selection. 268

A visual example of how the compatibility $effect^5$ arises is shown in Figure 4, and 269 this same example is worked out numerically in Appendix A. Here, the response-set cue 270 signals an odd/even judgement, and the stimulus-set cue is orange. In both panels, π_{orange} 271 is set to high meaning the attention weight of the orange stimulus is higher than that of the 272 purple stimulus; the effect of this is that the sensory evidence for each response category 273 is more strongly weighted toward evidence coming from the orange stimulus (indicated by 274 the bold arrows coming from the orange digit) than from the purple stimulus (indicated by 275 the light arrows coming from the purple digit). The sensory evidence from each stimulus 276 is then modified by the response bias for each category and is set by the response-set cue 277 which makes β_{odd} and β_{even} high. The correct response category, "even", therefore receives 278 a high probability of being selected because the processing rate of this response category 279 given the relevant stimulus (i.e., v(even, 2)) is high and the processing rate of this response 280 category given the irrelevant stimulus (i.e., v(even, 4)) is moderate. In contrast, on the 281 incompatible trial, the irrelevant stimulus and the response-set provide non-trivial evidence 282 for response-category "odd" (i.e., v(odd, 7)). The net effect is that the overall probability 283 of selecting the correct response category (Equation 3) is lower on incompatible trials than 284 compatible trials, leading to slower and more error prone responses (Equation 4). 285

⁵This is not to be confused with the term "congruency" which pertains to a single stimulus. Congruent stimuli are those for which the relevant and the irrelevant response-set responses are the same (e.g., the digit 3 is "odd" and "low", both of which require a "left" response), in contrast to incongruent stimuli for which the response-set responses are different (e.g., the digit 4 is "even" requiring a right response, and "low" requiring a left response). Compatibility, by contrast, is the degree of accord between the relevant response to the stimulus cued by the stimulus-set and the relevant response to the non-cued stimulus. Compatibility is independent of congruency.



Figure 4. Schematic overview of how ECTVA selects a response in the paradigm reported in the current paper.

286 The current study

I conducted a simulation of ECTVA adapted to the current experimental design to 287 establish predictions that could be assessed empirically (see Simulation 1 in Appendix A for 288 full details). This simulation used the example parameters reported in Logan and Gordon 289 (2001) (see their Figure 6), so the predictions reported here arise from the architecture of 290 the model rather than by tuning specific parameter values.⁶ The results of this simulation 291 can be seen in Figure 5. ECTVA predicts an under-additive interaction between response-292 set sequence and stimulus-set sequence, and this is driven by the negatively-accelerating 293 increase in reconfiguration time of response-set and stimulus-set parameters (Figure 3). 294 Importantly, in contrast to the explanation of Philipp and Koch (2010), this under-additive 295 interaction does not arise due to the integration of response-set and stimulus-set into a single 296 task-set; rather, ECTVA is clear that response-set (β) and stimulus-set (π) are independent 297 parameters that can reconfigured in parallel and this under-additive interaction is diagnostic 298 of this parallel reconfiguration. If ECVTA is forced to reconfigure β and π parameters 299 serially, this interaction is removed (see Simulation 2 in Appendix A). 300

Note that there is no predicted effect of switching either response-set or stimulus-set in the proportion error predictions. In ECTVA, set-switching is a time-consuming process that takes place before response selection begins; as such, the negative impact of parameter reconfiguration time on switch trials is isolated to response time. However, it is a common finding that switch costs in task switching occur in both response time *and* error rates (Grange & Houghton, 2014; Kiesel et al., 2010; Monsell, 2003; Vandierendonck et al.,

⁶In Appendix B I provide an exploration of the parameter space of ECTVA's main parameters and demonstrate that the main prediction of an under-additive interaction between response-set and stimulus-set sequence is indeed a consequence of the architecture of the model, and in particular Equation 5.



Stimulus Set • repetition • switch

Figure 5. Predictions from simulations of ECTVA for the current experimental design. See Simulation 1 in Appendix A for more details.

2010), so it seems that this default version of ECTVA is not capturing a key aspect of task switching phenomena. However, switch costs in error rates appear in ECTVA when task-set parameter values from the previous trial are allowed to carryover into the current trial. I return to this point in Experiment 4 (see also Simulation 5 in Appendix A).

A second prediction is the effect of response-compatibility: On compatible trials, over-311 all response time is shorter and more accurate than on incompatible trials, although this 312 effect appears more subtle for the response times than for accuracy. This effect—largely 313 overlooked in studies on component-switching in task switching—provides valuable infor-314 mation as to the task-set representation used to select a response: Despite the stimulus-set 315 prioritising sensory evidence coming from the relevant stimulus, the irrelevant stimulus still 316 contributes to the probabilities of selecting each response category, and on incompatible 317 trials this leads to a reduction in the probability of selecting the correct response cate-318 gory causing slower and more error-prone response selection. It is important to note that 319 ECTVA predicts that response-compatibility does not interact with either the response-set 320 or stimulus-set sequence: Compatibility effects arise during response selection after recon-321 figuration of response-set and stimulus-set parameters has occurred. 322

323 The purpose of the current study is therefore to test these combined predictions of

ECTVA to understand how stimulus-set and response-set are represented and controlled 324 in task switching. An additional contribution of the current study is to examine com-325 ponent switching within a paradigm—and therefore using task-set components—perhaps 326 more typical of traditional studies of task switching than previous studies of component-327 switching have used. For example, studies supporting the flat task-set organisation view 328 (e.g., Vandierendonck et al., 2008) and the componential view (Hübner et al., 2001) used 329 global-local stimuli with components of task (e.g., odd/even vs. low/high) and stimulus 330 dimension (global vs. local); studies supporting the integrated view (e.g., Philipp & Koch, 331 2010) used visual stimuli together with visual and auditory cues (Experiments 2 & 3) 332 with components of task (odd/even vs. low/high) and response modality (manual response 333 vs. verbal response); and studies supporting the hierarchical view (e.g., Kleinsorge & Heuer, 334 1999) used visual numerical stimuli with components of judgement (numerical judgement 335 of stimulus vs. spatial judgement of stimulus) and judgement-to-response mapping (com-336 patible vs. incompatible). Thus divergent findings across these studies (Figure 2) could be 337 explained by the different experimental paradigms and task-set components being switched 338 between. In the current study, task-set components being switched between (response-set 339 and stimulus-set) are clearly mapped onto parameters in ECTVA providing a clear theo-340 retical foundation upon which to study component-switching. 341

342

Experiment 1

Experiment 1a was a pilot study; as such, I performed a direct replication in Experiment 1b. The findings of both experiments are discussed together after reporting both.

345 Experiment 1a

346 Method.

Participants. All experiments received ethical approval from the Faculty of Nat ural Science's Psychology Research Ethics Committee at the host institution (REC 0246),
 and all participants provided informed consent. Participants were recruited online via Pro lific Academic. 50 participants residing either in the United Kingdom or the United States
 of America were recruited; 7 participants were removed for failure to maintain a session-wise
 accuracy above 90%.⁷ Participants could only access the experiment using either a desktop
 or laptop computer (i.e., tablets and smartphones were disallowed).

Materials. The experiment was programmed and delivered using Gorilla (Anwyl-Irvine, Massonnié, Flitton, Kirkham, & Evershed, 2020). All stimuli were presented within a black square frame centered on the screen on a grey background (RGB values 166, 166, 166). Task cues were either the words "odd/even" or "low/high", and could either be in orange (RGB 217, 93, 1) or purple (RGB 116, 112, 178) font. The identity of the cue served as the cue for the relevant response-set, and the colour of the cue served as the cue for the relevant stimulus-set. Cue identity and colour was selected randomly (with replacement)

 $^{^{7}}$ At the request of a Reviewer, data from all Experiments were reanalysed using a more liberal exclusion criterion of 80%. The results were largely the same, and none of the small differences changed any of the interpretations or main conclusions of the paper.

on each trial, and was presented slightly above center of the square frame. Task stimuli were the digits 1–9, excluding 5. Two different digits were randomly selected on each trial and presented beneath the cue.⁸ One stimulus was in orange, and the other in purple. The left-right ordering of the orange and purple digit was randomised on every trial.

Procedure. Each trial began with a blank square frame for 250ms, followed by 365 presentation of the cue. After 100 milliseconds (ms), the two digits appeared beneath the 366 cue. The cue and the digits remained on the screen until a response from the participant 367 was registered. The task of the participant was to perform the cued response-set (either a 368 lower/higher than 5 judgement, or an odd/even judgement) on the digit that matched the 369 cued stimulus-set for the trial (indicated by the colour of the cue). Participants made their 370 responses using the "z" and "m" keys of their keyboard: If the response-set was "low/high", 371 participants were instructed to press the z key for a "lower than 5" response, and m key for a 372 "higher than 5" response; if the response-set was "odd/even", z served as the "odd" response 373 and m the "even" response. Once a response had been registered, the frame went blank and 374 the next trial began. After receiving task instructions, participants were presented with 375 a practice block consisting of 16 trials, which could be repeated upon request. The main 376 experimental section comprised 6 blocks of 64 trials each, with self-paced rest screens after 377 each block. 378

Design. The experiment manipulated three independent variables in a fully related design: Response-Set Sequence (repetition vs. switch), Stimulus-Set Sequence (repetition vs. switch), and Response-Compatibility (compatible vs. incompatible). Response time (in ms) and proportion error served as the dependent variables.

Results. All of the data wrangling, statistical modelling, and visualisation utilised 383 R (R Core Team, 2022) and various packages⁹. The first trial from each block was 384 removed as these cannot be classified into either response/stimulus-set repetitions or 385 response/stimulus-set switches. The trial following an error was also removed for the same 386 reason; for the response time analysis, error trials themselves were also removed. For the 387 response time analysis, correct RTs shorter than 150 ms were removed, as were RTs longer 388 than 2.5 SDs above each participant's mean for each cell of the experimental design. Error 389 trimming (error trials plus post-error trials) removed 8.66% of the data, and RT trimming 390

⁸This actually introduces a slight bias of 57.14% incompatible trials (16 stimulus pairs out of a total of 28 unique stimulus pairs) to 42.86% compatible trials (12 stimulus pairs out of a total of 28 unique stimulus pairs). Thank you to Iring Koch for highlighting this.

⁹Specifically, I used R (Version 4.2.1; R Core Team, 2022) and the R-packages *afex* (Version 1.2.0; Singmann, Bolker, Westfall, Aust, & Ben-Shachar, 2022), *brms* (Version 2.21.0; Bürkner, 2017, 2018, 2021), *dplyr* (Version 1.1.4; Wickham, François, Henry, Müller, & Vaughan, 2023), *forcats* (Version 1.0.0; Wickham, 2023a), *ggplot2* (Version 3.5.1; Wickham, 2016), *lme4* (Version 1.1.33; Bates, Mächler, Bolker, & Walker, 2015), *lubridate* (Version 1.9.3; Grolemund & Wickham, 2011), *MASS* (Version 7.3.57; Venables & Ripley, 2002), *Matrix* (Version 1.5.1; Bates, Maechler, & Jagan, 2022), *papaja* (Version 0.1.2; Aust & Barth, 2022), *patchwork* (Version 1.2.0.9000; Pedersen, 2022), *purrr* (Version 1.0.1; Wickham & Henry, 2023), *Rcpp* (Eddelbuettel & Balamuta, 2018; Version 1.0.12; Eddelbuettel & François, 2011), *readr* (Version 2.1.4; Wickham, Hester, & Bryan, 2023), *stringr* (Version 1.5.1.9000; Wickham, 2023b), *tibble* (Version 3.2.1; Müller & Wickham, 2023), *tidybayes* (Version 3.0.4; Kay, 2023), *tidyr* (Version 1.3.0; Wickham, Vaughan, & Girlich, 2023), *tidyverse* (Version 2.0.0; Wickham et al., 2019), *tinylabels* (Version 0.2.3; Barth, 2022), and *trimr* (Version 1.1.1; Grange, 2022)



removed 2.87% of this error-trimmed data. Mean RT^{10} and proportion error across both factors of the design are visualised in Figure 6.

Figure 6. Behavioural data from Experiment 1a. The upper two plots show mean response time (in milliseconds, ms) and the lower two plots show proportion error as a function of response-set sequence, stimulus-set sequence, and response compatibility. Error bars show one standard error around the mean.

Analytical approach. All data in the current paper was analysed via use of Bayes factors for factorial designs as set out by Rouder, Morey, Verhagen, Swagman, and Wagenmakers (2017) using model comparisons with the BayesFactor package (and its default prior settings) in R¹¹. As the current experiment has three factors, the number of models—where each model differs in terms of inclusion of a particular predictor, be it a main effect or an

¹⁰As the critical test involves the test of an interaction between response-set sequence and stimulus-set sequence (and potentially response-compatibility), data for all experiments were also analysed using log-transformations of RT to protect against so-called "removable interactions" (see Wagenmakers, Krypotos, Criss, & Iverson, 2012). The results were qualitatively identical throughout with transformed RTs.

¹¹In the supplementary material, I provide frequentist analyses (factorial ANOVAs) for all experiments reported in this paper. (For the purposes of peer-review, this is included in the manuscript as Appendix E.)

interaction—required for the model comparison procedure grows to 128 (see Rouder et al., 398 2017), making model comparison non-feasible. To overcome this, I constructed a full model 399 which contained all main effects, all two-way interactions, plus the three-way interaction as 400 predictors. I then computed Bayes factors for seven new models, each removing an individ-401 ual predictor from the full model. The evidence for a particular predictor is then the Bayes 402 factor for that *omitted* model compared to the Bayes factor for the *full* model; that is, the 403 BF expressing evidence for the predictor becomes $BF = \frac{omitted-model}{full-model}$; Bayes factors lower 404 than 1 indicate support for the full model (and therefore for inclusion of that particular 405 predictor), whereas Bayes factors larger than 1 indicate support for the omitted-model (and 406 therefore for removal of that particular predictor). Throughout I interpret BFs between 1–3 407 as anecdotal evidence, 3–10 as moderate evidence, 10–30 as strong evidence, 30–100 as very 408 strong evidence, and greater than 100 as extreme evidence (Lee & Wagenmakers, 2014). 409

Response time analysis. The model comparisons can be seen in Table 1. This 410 showed extreme evidence in favour of the predictors Response-Set Sequence and Stimulus-411 Set Sequence: Consideration of the un-transformed marginal means showed response-set 412 repetitions (M = 1458 ms, SE = 25) were responded to faster than response-set switches 413 (M = 1653, SE = 28), and stimulus-set repetitions (M = 1451, SE = 27) were responded to 414 faster than stimulus-set switches (M = 1660, SE = 26). There was an ecdotal evidence for 415 the predictor Response-Compatibility: Response-compatible trials (M = 1538, SE = 28)416 were responded to faster than response-incompatible trials (M = 1573, SE = 27). 417

There was extreme evidence for the interaction between Response-Set Sequence and Stimulus-Set Sequence. There was moderate evidence against the interactions between Stimulus-Set Sequence and Response Compatibility and between Response-Set Sequence and Response Compatibility. There was also moderate evidence against the three-way interaction.

To evaluate the predictions of the different task-set structure hypotheses, I collapsed 423 the data across compatibility, and performed three separate pairwise comparisons using 424 Bayesian t-tests: (a) The comparison of single-component switch costs (i.e., a comparison 425 between response time when there is just a switch of response-set alone and when there 426 is just a switch of stimulus-set alone); (b) the comparison between response time for a 427 single-component switch of response-set and when there is a switch of both components; 428 and (c) the comparison between response time for a single-component switch of stimulus-429 set and when there is a switch of both components. The results showed moderate evidence 430 for no difference in response time between a single-component switch of response-set and 431 a single-component switch of stimulus-set $(BF_{10} = 0.19)$, but there was extreme evidence 432 for a difference between a single-component switch of response-set and a switch of both 433 components $(BF_{10} = 2.71 \times 10^6)$, and extreme evidence for a difference between a single-434 component switch of stimulus-set and a switch of both components ($BF_{10} = 571.11$). 435

Error analysis. The error data showed extreme evidence for Response-Set Sequence: There were overall fewer errors for response-set repetitions (M = 2.73%, SE = 0.24) than for response-set switches (M = 4.45%, SE = 0.31). There was also strong evidence for Response-Compatibility, with fewer errors on response-compatible trials (M = 3.08%, SE = 0.25) than response-incompatible trials (M = 4.10%, SE = 0.32). There appeared to be

Model comparison results for the behavioural data for Experiments 1a and 1b. The Bayes factors (BF) show comparison of the full factorial model (including all main effects and all interactions) against models with particular predictors omitted (i.e., $BF = \frac{omitted - model}{full - model}$). BF values below 1 indicate evidence in favour inclusion of that predictor.

Experiment	Omission	BF(RT)	BF (Error)
1a	R-Set Seq. (RS) x S-Set Seq. (SS) x Compatibility (C)	3.37	3.99
	$C \ge SS$	5.13	1.23
	$C \ge RS$	5.10	5.28
	RS x SS	4.34×10^{-3}	4.47
	С	0.68	0.08
	SS	9.67×10^{-30}	0.77
	RS	2.30×10^{-26}	2.43×10^{-5}
11.	DQ QQ Q	2 77	4.90
10		3.77	4.28
	C x SS	5.22	3.94
	$C \ge RS$	5.80	0.40
	RS x SS	1.02×10^{-5}	0.34
	С	$2.03 imes 10^{-3}$	1.96
	SS	1.01×10^{-39}	5.77
	RS	7.03×10^{-42}	9.08×10^{-15}

Note. Seq. = sequence.

⁴⁴¹ no effect of Stimulus-Set Sequence, and although stimulus-set switch costs appeared to be ⁴⁴² larger on response-incompatible trials (M = 1.32%) than on response-compatible trials (M⁴⁴³ = 0.14\%), there was anecdotal evidence against this in the model comparisons. There was ⁴⁴⁴ moderate evidence against the interaction between Response-Set Sequence and Stimulus-Set ⁴⁴⁵ Sequence, and against the three-way interaction.

esting the predictions of task-set structure is less relevant for error rates as the predictions of task-set structure from ECTVA are based on response times (see Equation 5). However, I report them here for completeness. The results (focussing only on incompatible trials¹²) showed moderate evidence for no difference in errors between a single-component switch of response-set and a single-component switch of stimulus-set ($BF_{10} = 0.22$), anecdotal evidence for no difference between a single-component switch of response-set and a switch of both components ($BF_{10} = 0.62$), and anecdotal evidence for a difference between

¹²This was done because stimulus-set switch costs are always likely to be asbsent on compatible trials. On stimulus-set switch trials, the participant will sometimes erroneously respond to stimulus associated with the previous (now-irrelevant) stimulus-set. However, on compatible trials, because the response elicited by the relevant stimulus is the same as the response elicited by the irrelevant stimulus, the participant's response (and therefore the probability of error) will be the same regardless of which stimulus they are actually attending to. Therefore, the "cost" of failing to update stimulus-set is not captured in error rates on compatible trials.

⁴⁵³ a single-component switch of stimulus-set and a switch of both components ($BF_{10} = 2.62$).

454 Experiment 1b

Experiment 1b served as a direct replication of Experiment 1a with only minor modifications to the procedure.

457 Method.

Participants. 60 new participants were recruited from the UK and USA via Pro lific. 9 participants were removed for failure to maintain a session-wise accuracy above
 90%.

461 Materials & Procedure. The same experimental materials from Experiment 1a
 462 were used. The procedure was identical except the main experimental section comprised 8
 463 blocks of 64 trials.

Results. The data were trimmed in an identical fashion to Experiment 1a. Error
trimming removed 9.38% of the data, and RT trimming removed 3.03% of this error-trimmed
data. Mean RTs and proportion error can be seen in Figure 7.

Response time analysis. The model comparisons can be seen in Table 1. As in 467 Experiment 1a, this analysis showed extreme evidence in favour of the predictors Response-468 Set Sequence and Stimulus-Set Sequence: Response-set repetitions (M = 1298 ms, SE = 19)469 were responded to faster than response-set switches (M = 1489, SE = 22), and stimulus-470 set repetitions (M = 1301, SE = 21) were responded to faster than stimulus-set switches 471 (M = 1486, SE = 20). There was also extreme evidence for Response-Compatibility (cf., 472 Experiment 1a): Response-compatible trials (M = 1369, SE = 22) were responded to faster 473 than response-incompatible trials (M = 1418, SE = 22). 474

Importantly, there was again extreme evidence for the interaction between ResponseSet Sequence and Stimulus-Set Sequence. Again there was moderate evidence against
the interactions between Stimulus-Set Sequence and Response Compatibility and between
Response-Set Sequence and Response Compatibility. There was also moderate evidence
against the three-way interaction.

Testing the predictions of task-set structure, the results (collapsing across compatibility) showed moderate evidence for no difference in response time between a single-component switch of response-set and a single-component switch of stimulus-set ($BF_{10} = 0.16$), but there was extreme evidence for a difference between a single-component switch of responseset and a switch of both components ($BF_{10} = 2.20 \times 10^6$), and extreme evidence for a difference between a single-component switch of stimulus-set and a switch of both components ($BF_{10} = 5.71 \times 10^5$).

Error analysis. The error data again largely replicated findings from Experiment 1a. There was extreme evidence for Response-Set Sequence: There were overall fewer errors for response-set repetitions (M = 3.06%, SE = 0.21) than for response-set switches (M = 5.44%, SE = 0.26), and fewer errors for stimulus-set repetitions (M = 4.13%, SE =0.26) than for stimulus-set switches (M = 4.37%, SE = 0.25). However, unlike Experiment



Figure 7. Behavioural data from Experiment 1b. The upper two plots show mean response time (in milliseconds, ms) and the lower two plots show proportion error as a function of response-set sequence, stimulus-set sequence, and response compatibility. Error bars show one standard error around the mean.

⁴⁹² 1b there was now anecdotal evidence *against* an effect of Response-Compatibility. There ⁴⁹³ was only weak evidence in favour of the interaction between Response-Compatibility and ⁴⁹⁴ Response-Set Sequence, and the interaction between Response-Set Sequence and Stimulus-⁴⁹⁵ Set Sequence. There was moderate evidence against inclusion of the interaction between ⁴⁹⁶ Response-Compatibility and Stimulus-Set Sequence, and the three-way interaction.

Testing the predictions of task-set structure, the results (focussing only on incompatible trials) showed no evidence for a difference in errors between a single-component switch of response-set and a single-component switch of stimulus-set ($BF_{10} = 1.13$), moderate evidence for no evidence for a difference between a single-component switch of response-set and a switch of both components ($BF_{10} = 0.16$), and no evidence for a difference between a single-component switch of stimulus-set ($BF_{10} = 1.04$).

503 Discussion

In sum, both Experiments 1a and 1b showed that for response time data there was 504 clear evidence for a cost to performance when switching either response-set or stimulus-505 set, as well as an under-additive interaction: The costs associated with switching a single 506 component (either response-set or stimulus-set) were equivalent, and there was an additional 507 cost of switching both task-set components that was less than the sum of switching both 508 components individually. Such a pattern of data is compatible with the integrated view 509 of Philipp and Koch (2010), but it is also compatible with the predictions of ECTVA (if 510 it is assumed response-set is represented at the categorical level; see Figure 3D). In the 511 error data, there was clear evidence of a response-set switch cost, but no evidence for a 512 stimulus-set switch cost (and no interaction between the two). Recall that ECTVA predicts 513 no cost in error rates for switching either response-set or stimulus-set. 514

ECTVA makes additional predictions about response-compatibility that were confirmed in the data: incompatible trials led to overall slower response times, and this did not interact with either set-sequence. This effect was moderate in Experiment 1a, but extreme evidence was present in Experiment 1b. The error data in Experiment 1a showed strong evidence for more errors on incompatible trials (this effect was anecdotally null in Experiment 1b).

521

Experiment 2

Experiment 1 provided evidence generally congruent with ECTVA predictions about 522 the costs associated with switching components of the task-set. However, in Experiment 1 523 a switch in either response-set or stimulus-set was also associated with a switch in the cue 524 (e.g., "odd/even" to "low/high" in the case of response-set, and "orange" to "purple" in the 525 case of stimulus-set). It is a consistent finding that such cue switching can contribute to 526 estimates of switch cost. To overcome this, researchers use two-cues per set (e.g., "odd/even" 527 and "parity" cue the "odd/even" judgement) to disentangle the effects of switching cue from 528 the cost of switching set (Logan & Bundesen, 2003; Mayr & Kliegl, 2003). Such a 2:1 cue-529 task mapping produces three types of sequence: A cue repetition, where both the cue and 530 the relevant set repeat (e.g., "odd/even"-"odd/even"); a cue switch, where the cue switches 531 but the relevant set repeats (e.g., "odd/even"-"parity"); and a set switch (where both the 532 cue and relevant set switches (e.g., "odd/even"-"low/high"). The impact of cue switching 533 can be estimated by comparing cue switch trials with cue repetition trials, and the impact 534 of set switching independent of cue switching can be estimated by comparing set switch 535 trials with cue switch trials. Research has generally found that both cue-switching and 536 set-switching contribute to the overall switch cost (Jost, De Baene, Koch, & Brass, 2013), 537 leaving open the question of the extent to which the costs observed in Experiment 1 were 538 due to switching response-set and stimulus-set, or due to cue-switching. 539

In Experiment 2, I used a two cues to signal each response-set (e.g., "low/high" and "magnitude" indicated the "lower/higher than five" response-set) and each stimulus-

set (e.g., a dark-red cue and a light-red cue indicated the "red" stimulus-set)¹³, and cuerepetitions were disallowed by design such that response-set and stimulus-set repetition trials always involved a switch in cue so estimates of switch cost are not confounded by cuerepetition effects. All previous studies examining the organisation of task-sets in componentswitching studies have only utilised one cue per component, leaving open the question of the extent to which the different observed empirical effects (Figure 2) were driven by componentswitching or cue-switching.

549 Method

Participants. 70 new participants were recruited using the same procedure as in
 Experiment 1. 15 participants were removed due to session-wise accuracy below 90%.

Materials & Procedure. Experiment 2 used the same cues and stimuli as in 552 Experiment 1, but the colours were different. The number stimuli could be either blue 553 (RGB: 68, 118, 180) or red (RGB: 179, 39, 32). The cues were either dark-blue (RGB: 48, 554 85, 129), bright-blue (RGB: 86, 152, 232), dark-red (RGB: 129, 28, 23), or bright-red (RGB: 555 232, 50, 41). The main procedural change from Experiment 1 was that in Experiment 2 556 the relevant stimulus-set could be cued by one of two shades: If the cue was either dark-557 or bright-blue, the relevant stimulus-set was blue, and if the cue was either dark- or bright-558 red, the relevant stimulus-set was red. The colour of the cue was chosen randomly on each 559 trial with the constraint that no immediate repetitions of cue-shade could occur; that is, 560 all stimulus-set repetition trials involved a stimulus-set cue-switch. As before participants 561 received a practice block comprising 16 trials (with an option to repeat) followed by 8 blocks 562 of 64 trials. 563

564 **Results**

Error trimming removed 8.76% of the data, and RT trimming removed 2.98% of this error-trimmed data. Mean RTs and proportion error can be seen in Figure 8.

Response time analysis. The model comparisons can be seen in Table 2. There was extreme evidence in favour of Response-Set Sequence and Stimulus-Set Sequence: Response-set repetitions (M = 1704 ms, SE = 33) were responded to faster than responseset switches (M = 1770, SE = 37), and stimulus-set repetitions (M = 1669, SE = 35) were responded to faster than stimulus-set switches (M = 1805, SE = 35). There was also strong evidence also for Response-Compatibility: Response-compatible trials (M = 1715, SE = 34) were also responded to faster than response-incompatible trials (M = 1760, SE = 37).

¹³Schneider (2016) demonstrated that cue-switch costs can be caused by both *conceptual* priming (i.e., matching of the concept cued) and *perceptual* priming (i.e., matching of the physical property of the cue) on cue-repetition trials. In Experiment 2, response-set repetition trials (e.g., "low/high"–"magnitude") and stimulus-set repetition trials (e.g., "dark-red"–"light-red") involve a conceptual repetition (because both "low/high" and "magnitude" refer to the same concept, and "dark-red" and "light-red" also refer to the same concept) but a perceptual switch (because neither "low/high" and "magnitude" or "dark-red" and "light-red" share identical perceptual properties). Experiment 2, then, disallows perceptual priming of cue encoding, but conceptual priming remains. I return to this in the General Discussion.



Figure 8. Behavioural data from Experiment 2. The left plot shows mean response time (in milliseconds, ms) and the right plot shows proportion error as a function of response-set sequence and stimulus-set sequence. Error bars show one standard error around the mean.

In contrast to Experiment 1, there was anecdotal evidence against the magnitude of 574 the response-set switch cost being different for stimulus-set repetitions and switches: How-575 ever, numerically the pattern in the data was similar to that in Experiment 1: response-set 576 switch cost was 91 ms when the stimulus-set repeated, and 42 ms when the stimulus-577 set also switched.¹⁴ There was moderate evidence that this interaction did not interact 578 with Response-Compatibility. There was also moderate evidence against the two-way in-579 teractions between Response-Compatibility and Response-Seq Sequence as well as between 580 Response-Compatibility and Stimulus-Set Sequence. 581

Testing the predictions of task-set structure, the results (collapsing across compatibility) showed strong evidence for a difference in response time between a single-component switch of response-set and a single-component switch of stimulus-set ($BF_{10} = 14.48$), ex-

¹⁴In the frequentist analysis, this interaction was approaching significance, F(1, 54) = 3.02, p = .088, $\eta_g^2 = .001$.

Table 2

Model comparison results for the behavioural data for Experiment 2. The Bayes factors (BF) show comparison of the full factorial model (including all main effects and all interactions) against models with particular predictors omitted (i.e., $BF = \frac{omitted - model}{full - model}$). BF values below 1 indicate evidence in favour inclusion of that predictor.

Omission	BF(RT)	BF (Error)
R-Set Seq. (RS) x S-Set Seq. (SS) x Compatibility (C)	4.41	3.93
$C \ge SS$	5.44	3.09×10^{-5}
$C \ge RS$	4.79	5.93
RS x SS	1.58	4.02
С	0.07	8.23×10^{-12}
SS	5.01×10^{-17}	2.15×10^{-4}
RS	2.84×10^{-4}	8.41×10^{-3}

Note. Seq. = sequence.

treme evidence for a difference between a single-component switch of response-set and a switch of both components ($BF_{10} = 7.80 \times 10^5$), but no evidence for a difference between a single-component switch of stimulus-set and a switch of both components ($BF_{10} = 0.63$).

Error analysis. The error analysis largely supported the findings found in the 588 response times. There was extreme evidence for both Response-Set Sequence and Stimulus-589 Set Sequence: There were fewer errors for response-set repetitions (M = 3.25%, SE = 0.21)590 than for response-set switches (M = 4.20%, SE = 0.24), and fewer errors for stimulus-set 591 repetitions (M = 3.13%, SE = 0.21) than for stimulus-set switches (M = 4.31%, SE =592 0.24). There was also extreme evidence for Response-Compatibility, with fewer errors on 593 response-compatible trials (M = 2.75%, SE = 0.20) than response-incompatible trials (M 594 = 4.69%, SE = 0.24). There was extreme evidence that Response-Compatibility interacted 595 with Stimulus-Set Sequence: Stimulus-set switch costs were larger on incompatible trials (M 596 = 2.46%) than on compatible trials (M = -0.10\%). There was moderate evidence against 597 the interaction of Response-Set Sequence and Stimulus-Set Sequence, and against inclusion 598 of the interaction between Response-Compatibility and Response-Set Sequence, and the 599 three-way interaction. 600

Testing the predictions of task-set structure, the results (focussing on incompatible trials) showed moderate evidence for a difference in errors between a single-component switch of response-set and a single-component switch of stimulus-set ($BF_{10} = 3.80$), extreme evidence for a difference between a single-component switch of response-set and a switch of both components ($BF_{10} = 168.23$), and moderate evidence for no difference between a single-component switch of stimulus-set ($BF_{10} = 0.29$).

607 Discussion

Experiment 2 controlled for cue-switch effects in estimates of response-set and stimulus-set switch costs by using a 2:1 cue-to-set mapping. The results showed robust

response-set and stimulus-set switch costs in both RT and error rates indicating the costs 610 observed in Experiment 1 are not exclusively driven by cue-transition effects. There was 611 again an under-additive interaction in the response time data, but it was weaker than in 612 Experiment 1. The examination of the predictions of the task-set structure showed that 613 a switch of just the stimulus-set component was more costly than a switch of just the 614 response-set component. This appears to be contrary to the predictions of all of the main 615 hypotheses of the structure of task-sets during component switching, as well as to ECTVA 616 where a switch in just the response-set component is expected to be more costly (Figure 3B) 617 or just as costly (Figure 3D) as a switch in just the stimulus-set component. It's not clear 618 why the interaction is weaker in Experiment 2, but overall response time was considerably 619 higher and more variable in Experiment 2 (M = 1737 ms, SD = 520) than in Experiment 620

⁶²¹ 1a (M = 1555, SD = 359) or Experiment 1b (M = 1393, SD = 310). ⁶²² Congruent with the predictions of ECTVA, Experiment 2 demonstrated a response-(1) $H_{1}^{(1)} = 0$ ($H_{1}^{(1)} = 0$) (H_{1}

compatibility effect in both RT and error. Although this effect was present in Experiment 623 1, it was inconsistent (i.e., it was strong in error but weak in RT in Experiment 1a, and 624 strong in RT but absent in error in Experiment 1b) but here it was consistent across 625 both RT and error. There was also an interaction between response-compatibility and 626 stimulus-set sequence: there was no evidence for a stimulus-set switch cost on compatible 627 trials, but a large cost on incompatible trials. This suggests that when the stimulus-set 628 switches (e.g., from "red" to "blue"), participants sometimes erroneously respond to the 629 stimulus associated with the previous (now-irrelevant) stimulus-set. On compatible trials, 630 erroneously responding to the irrelevant stimulus carries little consequence because the 631 response is compatible with the relevant stimulus (i.e., the response elicited by both stimuli 632 is the same). On incompatible trials, though, the correct response elicited by the relevant 633 and irrelevant stimulus are different, so attending to the wrong stimulus leads to an incorrect 634 response. 635

636

Experiment 3

Experiment 3 was interested in exploring the effects of general (Experiment 3a) and component-specific (Experiments 3b and 3c) advanced preparation on the costs of switching stimulus-set and response-set. Experiment 3a manipulated the general preparation time, whereas Experiments 3b and 3c selectively cued a single-component ahead of stimulus presentation allowing advanced preparation of that component.

642 Experiment 3a

Opportunity for preparation can be manipulated in cued task switching studies by varying the time between cue onset and stimulus onset, the cue-stimulus interval (CSI). Increasing the CSI has consistently been shown to reduce switch costs in standard task switching paradigms (but see Altmann, 2004), thought to reflect the opportunity to engage in task-set reconfiguration before stimulus onset (e.g., Meiran, 1996; Meiran et al., 2000) or reflecting enhanced time for cue encoding processes (Altmann & Gray, 2008; Logan & Bundesen, 2003; Schneider & Logan, 2005).

Preparation interval has also been manipulated in some studies on component switch-650 ing. In their Experiments 2–4, Vandierendonck et al. (2008) manipulated CSI across blocks, 651 and although extended CSIs led to faster responding, it did not interact with task sequence, 652 dimension sequence, nor did it moderate the two-way interaction between task sequence and 653 dimension sequence; at both short and long CSIs, single-component switches were just as 654 costly as two-component switches, supporting their "flat" view of task-set organisation. 655 Kleinsorge et al. (2002) found at short CSIs the data supported a hierarchical organisation, 656 but at longer CSIs there was no difference in response time across all types of switch se-657 quence (i.e., the data reflected a "flat" organisation). Philipp and Koch (2010) also found 658 data congruent with a "flat" organisation at longer CSIs, but took this as evidence that 659 task-set components had been integrated into a single representation (which takes time and 660 is therefore only possible at longer CSIs); as the components are integrated into a single rep-661 resentation, a switch in any single component is just as costly as a switch in all components, 662 leading to equivalent switch costs. 663

Reductions of component switch costs arises naturally from the assumptions of 664 ECTVA. Recall that in ECTVA, when response-set or stimulus-set switches, relevant con-665 trol parameters (β for response-set and π for stimulus-set) need to be reconfigured which 666 takes time. ECTVA assumes that response-set and stimulus-set reconfiguration occurs in-667 dependently, and hence can occur in parallel. The under-additive interaction arises due to 668 the overall number of (independent) parameters that need to be reconfigured. Implicit in 669 this assumption is that greater opportunity for task preparation allows more of this recon-670 figuration to occur before stimulus onset; as such, extending the CSI leads to a reduction 671 in both the response-set switch cost, the stimulus-set switch cost, and a reduction in the 672 strength of the interaction (see Simulation 3 in Appendix A). 673

The purpose of Experiment 3a was to test the prediction of ECTVA of the effects of extended CSI on the costs of switching stimulus-set and response-set, as well as their interaction. In addition, as extending the CSI also allows greater opportunity for cue encoding processes to complete (e.g., Schneider & Logan, 2005), finding an interaction of stimulus-set and response-set sequences at longer CSIs would also provide important converging evidence that the costs are not due to cue-transition effects (cf., Experiment 2).

Method.

680

Participants. The sample size was increased as the effect of interest in this experiment is the modulation of the two-way interaction previously reported. 122 new participants were recruited as per the previous experiments. 10 participants were removed due to session-wise accuracy below 90%.

Materials & Procedure. The experimental materials were the same as in Experiment 1. Each trial began with the presentation of the cue, which was presented for either 100ms or 900ms, after which time the stimuli also appeared. The cue presentation time was randomised on each trial. Once a response was made, the frame went blank for an inter-trial interval of 250 ms, after which time the next trial began. After the practice block (see Experiment 1b) participants completed 10 blocks of 64 trials.

Results. Error trimming removed 7.11% of the data, and RT trimming removed 2.88% of this error-trimmed data. Mean RTs and proportion error can be seen in Figure 9.



Figure 9. Behavioural data from Experiment 3a. The left plot shows mean response time (in milliseconds, ms) and the right plot shows proportion error as a function of response-set sequence and stimulus-set sequence. Error bars show one standard error around the mean.

Table 3

Model comparison results for the behavioural data for Experiment 3a. The Bayes factors (BF) show comparison of the full factorial model (including all main effects and all interactions) against models with particular predictors omitted (i.e., $BF = \frac{omitted - model}{full - model}$). BF values below 1 indicate evidence in favour inclusion of that predictor.

Omission	BF(RT)	BF (Error)
CSI x Compatibility (C) x R-Set Seq. (RS) x S-Set Seq. (SS)	4.12	3.13
$C \ge RS \ge SS$	5.44	4.06
$CSI \ge C \ge SS$	5.87	3.93
$CSI \ge C \ge RS$	6.66	1.32
$CSI \ge RS \ge SS$	0.08	5.65
$C \ge SS$	7.02	$1.39 imes 10^{-3}$
C x RS	7.79	8.44
C x CSI	7.40	4.21
RS x SS	6.77×10^{-15}	0.95
CSI x SS	4.42×10^{-5}	8.81
CSI x RS	1.91×10^{-4}	5.40
С	3.17×10^{-10}	1.14×10^{-11}
SS	1.75×10^{-103}	$5.11 imes 10^{-8}$
RS	1.04×10^{-103}	2.01×10^{-23}
CSI	1.13×10^{-334}	1.01

Note. Seq. = sequence.

710

Response time analysis. The model comparisons can be seen in Table3. This 693 showed extreme evidence in favour of CSI: Consideration of the marginal means showed 694 a large effect of CSI; RTs were shorter at long CSI (M = 1112 ms, SE = 10) than short 695 CSI (M = 1444, SE = 12). As in previous experiments, there was extreme evidence in 696 favour of both Response-Set Sequence and Stimulus-Set Sequence: Response-set repetitions 697 (M = 1200 ms, SE = 11) were responded to faster than response-set switches (M = 1356, SE = 1000 ms)698 SE = 13), and stimulus-set repetitions (M = 1200, SE = 12) were responded to faster 699 than stimulus-set switches (M = 1356, SE = 12). There was extreme evidence that these 700 switch costs were reduced at longer CSIs: Response-set switch cost was 186 ms at short 701 CSI and 125 ms at longer CSIs, and stimulus-set switch cost was 188 ms at short CSI and 702 123 ms at longer CSIs. There was extreme evidence for Response Compatibility: Response 703 times were shorter for response-compatible trials (M = 1255, SE = 12) than for response-704 incompatible trials (M = 1301, SE = 13). However, there was moderate evidence that 705 Response Compatibility was not involved in any interactions. 706

Critically, there again was extreme evidence for an interaction between Response-Set
 Sequence and Stimulus-Set Sequence. There was strong evidence that this interaction was
 further modulated by CSI.

Testing the predictions of task-set structure, the results for the 100ms CSI condition

(collapsing across compatibility) showed no difference in response time between a single-711 component switch of response-set and a single-component switch of stimulus-set $(BF_{10} =$ 712 (0.11), but there was extreme evidence for a difference between a single-component switch of 713 response-set and a switch of both components $(BF_{10} = 1.20 \times 10^{12})$, and extreme evidence 714 for a difference between a single-component switch of stimulus-set and a switch of both 715 components $(BF_{10} = 2.47 \times 10^9)$. For the 900ms CSI condition, there was no difference in 716 response time between a single-component switch of response-set and a single-component 717 switch of stimulus-set $(BF_{10} = 0.11)$, but there was extreme evidence for a difference be-718 tween a single-component switch of response-set and a switch of both components (BF_{10}) 719 $= 4.11 \times 10^{10}$), and extreme evidence for a difference between a single-component switch of 720 stimulus-set and a switch of both components $(BF_{10} = 1.24 \times 10^7)$ 721

Error analysis. The model comparison showed no evidence for an effect of CSI, but extreme evidence for an effect of Response-Set Sequence, Stimulus-Set Sequence, and Response-Compatibility: Response-set repetitions (M = 2.29%, SE = 0.10) produced lower error than response-set switches (M = 3.83%, SE = 0.14); stimulus-set repetitions (M = 2.58%, SE = 0.11) produced lower error than stimulus-set switches (M = 3.53%, SE = 0.13); and compatible trials (M = 2.52%, SE = 0.11) produced lower error than incompatible trials (M = 3.60, SE = 0.13).

There was an ecdotal-to-moderate evidence against all interactions with the exception of the interaction between Response Compatibility and Stimulus-Set Sequence, for which there was extreme evidence in support: stimulus-set switch costs were smaller on compatible trials (0.35%) than on incompatible trials (M = 1.55%).

Testing the predictions of task-set structure, the results (focussing on incompatible 733 trials) for the 100ms CSI condition (collapsing across compatibility) showed no difference 734 in errors between a single-component switch of response-set and a single-component switch 735 of stimulus-set $(BF_{10} = 2.18)$, strong evidence for a difference between a single-component 736 switch of response-set and a switch of both components $(BF_{10} = 46.68)$, and strong evi-737 dence for a difference between a single-component switch of stimulus-set and a switch of 738 both components $(BF_{10} = 92.88)$. For the 900ms CSI condition, there was no evidence 739 for a difference in errors between a single-component switch of response-set and a single-740 component switch of stimulus-set $(BF_{10} = 0.12)$, no evidence for a difference between a 741 single-component switch of response-set and a switch of both components $(BF_{10} = 0.47)$, 742 and no evidence for a difference between a single-component switch of stimulus-set and a 743 switch of both components $(BF_{10} = 0.21)$. 744

Discussion. The experiment replicated the finding of an under-additive interaction 745 between response-set and stimulus-set sequence. In addition, extended preparation reduced 746 overall response times and it reduced the magnitude of the interaction between response-set 747 and stimulus-set sequence. Tests of the task-set structure showed that—at both short- and 748 long-CSIs—single-component switch costs were equivalent, and both were smaller than costs 749 associated with switching both components. These findings are in contrast to the results 750 of previous work showing the task-set structure appears "flat" at extended preparation 751 intervals (Kleinsorge et al., 2002; Philipp & Koch, 2010; Vandierendonck et al., 2008). 752 However, the observed preparation effects are predicted by ECTVA due to the assumption 753

754

of parallel reconfiguration of response-set and stimulus-set parameters that can occur ahead of stimulus onset with extended CSIs, although this reconfiguration is not complete by

of stimulus onset with extended CSIs, although this reconfiguration is not complete by
stimulus onset and a residual switch cost for both response-set and stimulus-set remains
(which can also be accounted by ECTVA; see Figure 16 in Appendix A).

An additional ECTVA prediction replicated here in both response times and error 758 rates is that of response-compatibility: Responses were overall slower and less accurate on 759 incompatible trials, driven by the influence of the irrelevant stimulus on response selec-760 tion. ECTVA predicts that CSI does not interact with response-compatibility as advanced 761 preparation exclusively affects task-set parameter reconfiguration, whereas the effects of 762 response compatibility arise during response selection after the task-set parameters have 763 been established. Experiment 3a also replicated the interaction of response-compatibility 764 and stimulus-set sequence in errors, where stimulus-set switch costs only appeared in in-765 compatible trials. 766

767 Experiment 3b

ECTVA assumes that response-set and stimulus-set are control parameters that can 768 be reconfigured independently and in parallel; the under-additive interaction is diagnostic 769 of this parallel reconfiguration. In Experiment 3a, extending the CSI afforded more time 770 for this parallel reconfiguration to occur before stimulus onset. Experiments 3b and 3c are 771 interested in the question of whether participants can *selectively* reconfigure either response-772 set or stimulus-set independently of the other. This was achieved by cuing one component 773 of the task-set in advance before providing the cue for the other component of the task-set 774 (which appears at the same time as the stimuli; see left panel of Figure 10): For example, 775 when the response-set is cued first, the task cue indicates the relevant response-set (e.g., 776 "low/high") but provides no information about the relevant stimulus-set (and vice-versa 777 when the stimulus-set is cued first); the relevant stimulus-set (e.g., "purple") is cued at the 778 same time as stimulus onset. 779

Such a design has been conducted previously by Philipp and Koch (2010) (Experiment 780 3), where participants were provided opportunity for advanced preparation of either (a) the 781 type of judgement to make on a stimulus (odd/even vs. lower/higher), (b) the response 782 modality (vocal vs. manual), (c) both components, or (d) no opportunity for preparation. 783 Results showed overall shorter RTs when participants had opportunity to prepare for both 784 components, a result the authors suggest supports the account that the task-set components 785 become integrated into a single representation which can be achieved during the preparation 786 interval when both components are cued. However, specific preparation (of either judgement 787 or modality components) did not lead to a reduction in the switch cost associated with that 788 component; rather preparation of a single component led to a reduction in the switch cost 789 associated with the other component. Acknowledging this surprising result, the authors 790 speculate that the non-prepared component might receive more attention when it is cued 791 just before stimulus onset. 792

⁷⁹³ In contrast to these findings, ECTVA predicts that selective task-set component ⁷⁹⁴ preparation is possible, and that such preparation leads to a selective reduction in the

switch cost associated with the component that was prepared for (see Simulation 4 in Ap-795 pendix A). This prediction arises naturally from the assumption of independent parameters 796 controlling response-set and stimulus-set. The purpose of Experiments 3b and 3c was to 797 assess the effect of providing opportunity for selective task-set component preparation on 798 resulting switch costs. Both experiments were identical except for one modification: In 799 Experiment 3b information about the first cued component remained visible during stimu-800 lus onset, and in Experiment 3c information about the first cued component was removed 801 before stimulus onset (see Figure 10). Removal of the information about the first cued com-802 ponent was removed to encourage participants to engage in advanced task-set preparation 803 (Verbruggen, Liefooghe, Vandierendonck, & Demanet, 2007; but see Schneider, 2016). The 804 findings of both experiments are discussed together after reporting both. 805

Method.

806

Participants. 115 new participants were recruited as per previous experiments. 12
 participants were removed as their session-wise accuracy was below 90%.

Materials & Procedure. The experimental procedure is shown in Figure 10. The 809 cue in this experiment did not initially provide information about both the response-set and 810 the stimulus-set. Instead, information about each relevant set was provided in stages: In the 811 first stage, only information about either the response-set or the stimulus-set (dependent 812 upon the current condition) was presented; after 900 ms, the cue added information about 813 the other relevant set, and at this stage the cue appeared as in previous experiments (i.e., 814 a compound cue showing information about the relevant response-set and stimulus-set). In 815 the first stage, in the condition where the response-set was cued first the cue could either be 816 the words "odd/even" or "low/high" in white. After 900 ms, the word cue changed colour 817 to either orange or purple which signalled the relevant stimulus-set. In the condition where 818 the stimulus-set was cued first, the cue was the placeholder "---/---" in either orange or 819 purple. After 900 ms, the placeholder was replaced with either the words "odd/even" or 820 "low/high" were presented in the same colour as the stimulus-set. The imperative stimuli 821 for the trial appeared at stage 2 when the compound cue was present. 822

Participants initially worked though 16 trials of practice with the compound cues 823 (i.e., the initial stage of practice was identical to previous experiments). After this initial 824 practice, participants received specific instructions for the current experiment, dependent 825 upon the current condition (the order of which was randomised across participants). In 826 one half of the experiment, participants received instructions for the response-set cued first 827 before receiving 16 trials of practice with that procedure (which could be repeated). After 828 this, participants received 5 blocks of 64 trials. The next half of the experiment presented 829 instructions for the stimulus-set cued first condition before receiving 16 trials of practice 830 with that procedure (which again could be repeated). After this, participants received 5 831 blocks of 64 trials. There were self-paced rest screens after each block. 832

Results. Error trimming removed 8.59% of the data, and RT trimming removed
2.84% of this error-trimmed data. Mean RTs and proportion error can be seen in Figure
11.

Response time analysis. The model comparisons can be seen in Table 4. This
 showed moderate evidence in favour of Order: RTs were shorter when the stimulus-set cue



Figure 10. Overview of Experiments 3b and 3c.

appeared first (M = 1320 ms, SE = 16) than when the response-set cue appeared first (M 838 = 1353, SE = 17). As in previous experiments, there was extreme evidence in favour of 839 both Response-Set Sequence and Stimulus-Set Sequence: Response-set repetitions (M =840 1281 ms, SE = 16) were responded to faster than response-set switches (M = 1392, SE 841 = 18), and that stimulus-set repetitions (M = 1298, SE = 17) were responded to faster 842 than stimulus-set switches (M = 1375, SE = 17). There was moderate evidence that these 843 switch costs did not interact with Set Order. Response-set switch cost was 118 ms when the 844 response-set cue was presented first, and 105 ms when the stimulus-set cue was presented 845 first. Stimulus-set switch costs were 82 ms when the response-set cue was presented first, 846 and 71 ms when the stimulus-set cue was presented first. There was extreme evidence for 847 Response Compatibility: Response times were shorter for response-compatible trials (M =848 1306, SE = 17) than for response-incompatible trials (M = 1366, SE = 17). However, there 849 was anecdotal-to-moderate evidence that Response Compatibility was not involved in any 850 interactions. 851

There again was very strong evidence for an interaction between Response-Set Sequence and Stimulus-Set Sequence. There was moderate evidence that this interaction was not further modulated by Set Order or Compatibility, and there was moderate evidence against the four-way interaction.

Testing the predictions of task-set structure, the results (collapsing across compatibility and set-order) showed anecdotal evidence for no difference in response time between a single-component switch of response-set and a single-component switch of stimulus-set $(BF_{10} = 2.74)$, but there was extreme evidence for a difference between a single-component switch of response-set and a switch of both components ($BF_{10} = 679.80$), and extreme evidence for a difference between a single-component switch of stimulus-set and a switch of



Stimulus Set - repetition - switch

Figure 11. Behavioural data from Experiment 3b as a function of whether response-set was cued first ("rs_first") or stimulus-set was cued first ("ss_first"). Error bars show one standard error around the mean.

Table 4

Model comparison results for the behavioural data for Experiment 3b. The Bayes factors (BF) show comparison of the full factorial model (including all main effects and all interactions) against models with particular predictors omitted (i.e., $BF = \frac{omitted - model}{full - model}$). BF values below 1 indicate evidence in favour inclusion of that predictor.

Omission	BF(RT)	BF (Error)
Order (O) x Compatibility (C) x R-Set Seq. (RS) x S-Set Seq. (SS)	4.08	3.62
$C \ge RS \ge SS$	4.61	6.93
$O \ge C \ge SS$	5.91	5.49
$O \ge C \ge RS$	3.08	8.75
$O \ge RS \ge SS$	5.12	3.75
$C \ge SS$	5.91	0.05
$C \ge RS$	8.02	2.37
C x O	1.11	8.59
RS x SS	0.03	0.98
O x SS	7.48	12.34
O x RS	7.55	8.17
С	$1.17 imes 10^{-5}$	1.77×10^{-6}
SS	2.06×10^{-9}	0.02
RS	2.39×10^{-19}	3.84×10^{-30}
0	0.16	20.05

Note. Seq. = sequence.

⁸⁶² both components $(BF_{10} = 5.68 \times 10^6)$.

Error analysis. The error analysis largely mirrored the outcomes found for re-863 sponse time. There was extreme evidence in favour of Response-Set Sequence and very 864 strong evidence for Stimulus-Set Sequence: Response-set repetitions (M = 2.59%, SE = 865 (0.12) were responded to more accurately than response-set switches (M = 4.56%, SE = 866 (0.15), and that stimulus-set repetitions (M = 3.28%, SE = 0.13) were responded to more 867 accurately than stimulus-set switches (M = 3.87%, SE = 0.15). There was moderate-to-868 strong evidence that these switch costs did not interact with Set Order. However, there 869 was strong evidence against an effect of Set Order overall, and there was no evidence of an 870 interaction between Response-Set Sequence and Stimulus-Set Sequence. 871

Unlike the RT analysis, there was strong evidence for an interaction between Response Compatibility and Stimulus-Set Sequence: stimulus-set switch costs were 0.03% on compatible trials, and 1.14% on incompatible trials. There was anecdotal-to-moderate evidence against all other interaction effects

Testing the predictions of task-set structure, the results (focussing on incompatible trials and collapsing across set-order) showed extreme evidence for a difference in errors between a single-component switch of response-set and a single-component switch of stimulus-set ($BF_{10} = 10.99$), no evidence for a difference between a single-component switch

of response-set and a switch of both components ($BF_{10} = 0.95$), but there was extreme evidence for a difference between a single-component switch of stimulus-set and a switch of both components ($BF_{10} = 1.19 \times 10^5$).

883 Experiment 3c

Experiment 3b showed no evidence for specific reductions in single-component switch 884 costs when opportunity for advanced preparation of that component was possible. However, 885 one possible explanation for this lack of a preparation effect could be that participants did 886 not engage in advanced preparation; this is possible in the current design because although 887 information about one component (e.g., the response-set) is provided in advance of the 888 stimuli, the information remains present throughout the trial (see Figure 10). A possible 889 strategy, then, is to wait until information about both task-set components is present before 890 commencing preparation. To remove this possible strategy, Experiment 3c replicated the 891 design of 3b but information about the task-set component provided first was removed at 892 stimulus onset. 893

894 Method.

Participants. 130 new participants were recruited as per previous experiments. 29
 participants were removed due to failure to maintain a session-wise accuracy above 90%.

Materials & Procedure. The experiment was similar to Experiment 3b (see Fig-897 ure 10) except the relevant response-set and stimulus-set was cued serially, and no joint 898 compound-cue was presented. In the first stage of cue presentation, only information about 899 either the response-set or the stimulus-set (dependent upon the current condition) was pre-900 sented (as in Experiment 3b). However, after 900 ms, this information was removed and 901 information about only the other set was presented. In both stages of cue presentation, the 902 information about the relevant response-set was provided by either the words "odd/even" 903 or "low/high" in white, and the information about the relevant stimulus-set was provided 904 by the placeholder "---/---" presented in either orange or purple. As in Experiment 3b, 905 the imperative stimuli for the trial appeared at the same time as the second stage of cue 906 presentation. 907

Although compound cues were not presented in the main experiment, as in Experi-908 ment 3b participants initially worked though 16 trials of practice with the compound cues 909 in order to learn the task and the response-set and stimulus-set cues. After this initial 910 practice, participants received specific instructions for the current experiment, dependent 911 upon the current condition (the order of which was randomised across participants). In 912 one half of the experiment, participants received instructions for the response-set cued first 913 before receiving 16 trials of practice with that procedure (which could be repeated). After 914 this, participants received 5 blocks of 64 trials. The next half of the experiment presented 915 instructions for the stimulus-set cued first condition before receiving 16 trials of practice 916 with that procedure (which again could be repeated). After this, participants received 5 917 blocks of 64 trials. There were self-paced rest screens after each block. 918

Results. Error trimming removed 10.08% of the data, and RT trimming removed
2.82% of this error-trimmed data. The response times and error rates can be seen in Figure
12



Stimulus Set - repetition - switch

Figure 12. Behavioural data from Experiment 3c as a function of whether response-set was cued first ("rs_first") or stimulus-set was cued first ("ss_first"). Error bars show one standard error around the mean.

Response time analysis. The model comparisons can be seen in Table 5. This 922 showed extreme evidence in favour of Order: RTs were shorter overall when the response-923 set cue was presented first (M = 1253 ms, SE = 14) than when the stimulus-set cue was 924 presented first (M = 1327, SE = 14). As in previous experiments, there was extreme 925 evidence in favour of both Response-Set Sequence and Stimulus-Set Sequence: Response-926 set repetitions (M = 1238 ms, SE = 13) were responded to faster than response-set switches 927 (M = 1342, SE = 14), and stimulus-set repetitions (M = 1254, SE = 13) were responded 928 to faster than stimulus-set switches (M = 1326, SE = 14). There was moderate evidence 929 that these switch costs did not interact with Set Order. Response-set switch cost was 114 930
Table 5

Model comparison results for the behavioural data for Experiment 3c. The Bayes factors (BF) show comparison of the full factorial model (including all main effects and all interactions) against models with particular predictors omitted (i.e., $BF = \frac{omitted - model}{full - model}$). BF values below 1 indicate evidence in favour inclusion of that predictor.

Omission	BF(RT)	BF (Error)
Order (O) x Compatibility (C) x R-Set Seq. (RS) x S-Set Seq. (SS)	2.92	3.44
$C \ge RS \ge SS$	4.65	4.16
$O \ge C \ge SS$	6.13	0.53
$O \ge C \ge RS$	6.14	2.91
$O \ge RS \ge SS$	5.29	2.00
$C \ge SS$	6.60	$2.01 imes 10^{-3}$
$C \ge RS$	2.92	7.37
C x O	3.13×10^{-3}	3.63
RS x SS	0.03	5.62
O x SS	8.35	3.79
O x RS	4.06	6.98
С	2.23×10^{-10}	$3.83 imes 10^{-11}$
SS	8.74×10^{-18}	2.85×10^{-5}
RS	6.50×10^{-36}	3.31×10^{-30}
0	1.57×10^{-18}	0.94

Note. Seq. = sequence.

ms when the response-set cue was presented first, and 94 ms when the stimulus-set cue was presented first. Stimulus-set switch costs were 72 ms when the response-set cue was presented first, and 73 ms when the stimulus-set cue was presented first.

There was extreme evidence for Response Compatibility: Response times were shorter for response-compatible trials (M = 1262, SE = 13) than for response-incompatible trials (M = 1318, SE = 14). There was very strong evidence that Response Compatibility interacted with Set Order, but there was anecdotal-to-moderate evidence that Response Compatibility was not involved in any other interactions.

There again was very strong evidence for an interaction between Response-Set Sequence and Stimulus-Set Sequence. There was moderate evidence that this interaction was not further modulated by Set Order or Compatibility, and there was anecdotal evidence against the four-way interaction.

Testing the predictions of task-set structure, the results (collapsing across compatibility and set-order) showed strong evidence for a difference in response time between a single-component switch of response-set and a single-component switch of stimulus-set $(BF_{10} = 17.94)$, extreme evidence for a difference between a single-component switch of response-set and a switch of both components $(BF_{10} = 7.90 \times 10^5)$, and extreme evidence for a difference between a single-component switch of both

949 components $(BF_{10} = 3.58 \times 10^9)$.

Error analysis. There was no evidence for an effect of Set Order, but there was 950 extreme evidence in favour of both Response-Set Sequence and Stimulus-Set Sequence: 951 Response-set repetitions (M = 3.18%, SE = 0.12) were responded to more accurately than 952 response-set switches (M = 5.38%, SE = 0.17), and stimulus-set repetitions (M = 3.83%, 953 SE = 0.14) were responded to more accurately than stimulus-set switches (M = 4.75%, SE 954 = 0.16). There was moderate evidence that these switch costs did not interact with Set 955 Order. However, there was no evidence for an effect of Set Order overall, and there was 956 moderate evidence against an interaction between Response-Set Sequence and Stimulus-Set 957 Sequence. 958

There was extreme evidence for an overall effect of Response Compatibility: accuracy was better on compatible trials (M = 3.62%, SE = 0.15) than on incompatible trials (M = 4.95%, SE = 0.16). There was extreme evidence for an interaction between Response Compatibility and Stimulus-Set Sequence: stimulus-set switch costs were 0.19% on compatible trials, and 1.66% on incompatible trials. There was anecdotal-to-moderate evidence against all other interaction effects.

Testing the predictions of task-set structure, the results (focussing on incompatible trials and collapsing across set order) showed no evidence for a difference in errors between a single-component switch of response-set and a single-component switch of stimulus-set $(BF_{10} = 0.44)$, very strong evidence for a difference between a single-component switch of response-set and a switch of both components $(BF_{10} = 60.59)$, and extreme evidence for a difference between a single-component switch of stimulus-set alone and a switch of both components $(BF_{10} = 8.91 \times 10^5)$.

Discussion. The results of both Experiments 3b and 3c were clear: Despite being 972 provided opportunity for advanced preparation of specific components of the task-set (either 973 response-set or stimulus-set) there was no reduction in the switch cost associated with this 974 component. This replicates the finding of Philipp and Koch (2010) who also reported no 975 effects of component-specific preparation, which they suggested was consistent with the 976 idea that task-set components become integrated into a single representation. In their 977 Experiment 3, they had an additional condition where opportunity for preparing both task-978 set components was provided, which was not done in the current study's Experiments 979 3b or 3c (but see the long CSI condition of Experiment 3a). Philipp and Koch (2010) 980 reported overall shorter RTs when both components could be prepared compared with 981 when preparation was only possible for a single-component, which the authors propose was 982 caused by the ability to begin the integration of the task-set components when preparation 983 was possible for both. Experiments 3b and 3c also replicated the finding of response-984 compatibility, and in particular its interaction with stimulus-set sequence. 985

Experiments 3b and 3c again replicated the finding of an under-additive interaction between response-set and stimulus-set sequence, with costs of a single-component switch (of either response-set or stimulus-set) being smaller than the cost of switching both components. Whilst this interaction pattern is compatible with the task-set integration view of Philipp and Koch (2010), such an interaction pattern is also predicted by ECTVA which assumes independent task-set components that can be reconfigured in parallel. However, the

39

finding of a lack of an effect of selective preparation on the component-specific switch costs is in contrast to the predictions of ECTVA, which makes a clear prediction that componentspecific preparation is possible and it leads to a selective reduction in the associated switch cost.

One possibility is that component-specific preparation is possible in the current ex-996 perimental design, but participants do not consistently take advantage of this. Such an 997 argument has been used to explain the residual switch cost in task switching (the cost that 998 remains even after ample time to prepare for a task switch). For example, the *failure to* 999 engage hypothesis (DeJong, 2000; Nieuwenhuis & Monsell, 2002) proposes that advanced 1000 task-set reconfiguration is possible with sufficient preparation time, but participants fail 1001 to engage with such reconfiguration on a proportion of trials. In exploratory analysis of 1002 Experiment 3b and 3c's data (reported in Appendix C), I found some evidence to support 1003 this account. Specifically, using response time distributional analysis the component switch 1004 costs were smaller when their respective component was cued first (e.g., response-set switch 1005 cost was smaller when response-set was cued first) at the shortest response time percentiles. 1006 If one assumes the shortest percentiles reflect those trials that participants engaged in ad-1007 vanced preparation (hence why they are the shorter RTs; DeJong, 2000), this could be taken 1008 as evidence supporting ECTVA's prediction that component-specific preparation is possible. 1009 In Appendix C I report an additional experiment that sought to encourage advanced prepa-1010 ration further; adapting Experiment 3c's design, the cue for the component that was cued 1011 first was only presented very briefly (see e.g., Verbruggen et al., 2007). However, there was 1012 still no evidence for a reduction in component-specific switch cost, and the distributional 1013 analysis did not clearly replicate the reduction of switch cost at shorter RT percentiles.¹⁵ 1014

1015

Experiment 4

For completeness, in Experiment 4 I examined the effect of varying the response-cue 1016 interval (RCI)—the time between the response to the previous trial and the onset of the 1017 cue for the current trial—on component-switch costs. The RCI does not impact prepara-1018 tion processes because the next task is unknown during this period when task sequence 1019 is random. Despite this, studies have shown that the task switch cost is larger at short 1020 RCIs than at long RCIs (see e.g., Meiran et al., 2000; but see Altmann, 2005). The decay 1021 account of this increase in switch cost centers on the idea that there is increased proactive 1022 interference from the previous trial's task-set at short RCIs, but at longer RCIs the previous 1023 trial's task-set activation will have decayed leading to less proactive interference on switch 1024 trials (Allport et al., 1994; Altmann & Gray, 2008; Meiran et al., 2000). Other accounts 1025 (Horoufchin, Philipp, & Koch, 2011b, 2011a; Grange, 2016; see also Grange & Cross, 2015) 1026 propose that the RCI effects are not caused by passive decay, but rather due to a reduction 1027 of the *temporal distinctiveness* of a previous trial's episodic memory trace at shorter RCIs, 1028 which leads to a loss of repetition benefits (and as such, an increased switch cost). 1029

1030

To my knowledge, no study on component switching has manipulated the RCI so it

¹⁵Note that Philipp and Koch (2010) reported that preparation of a single component led to a reduction in the switch cost associated with the *other* component, which was to some extent replicated in the additional experiment.

remains unclear whether the RCI impacts response-set switching, stimulus-set switching, or both in the current design. However, Horoufchin et al. (2011a) provided evidence from a standard task switching design showing strong RCI effects when stimulus valency was manipulated (Experiment 2), and weak RCI effects when response valency was manipulated (Experiment 1). Interpreting their findings from a temporal distinctiveness account of RCI effects, Horoufchin et al. concluded that manipualtion of the RCI impacts cue-based retrieval of stimulus-set.

There is evidence from the previous experiments of the current study that suggest 1038 there is interference in the current design from the previous-trial's stimulus-set representa-1039 tion, driven by the interaction of stimulus-set sequence and response compatibility in error 1040 rates (see Experiments 2 and 3a-3c). This interaction—driven by stimulus-set switch costs 1041 only being present on response-incompatible trials—suggests participants sometimes attend 1042 to the stimulus associated with the previous-trial on stimulus-set switch trials. This could 1043 either reflect a failure to reconfigure this parameter or persisting activation of—and proac-1044 tive interference from—the previous (but now irrelevant) stimulus-set. That this interaction 1045 was not moderated by the cue-stimulus interval in Experiment 3a suggests that it might 1046 not be due to a failure to update the relevant parameters, but rather due to persisting 1047 carry-over of the previous trial's task-set parameters. 1048

In terms of ECTVA, carry-over effects arise if the assumption is added that a pro-1049 portion of the previous trial's response-set and stimulus-set parameters carry-over into the 1050 next trial and influence the current trial's parameter settings. On component repetition 1051 trials this can facilitate performance because the relevant component parameters receive 1052 a boost; on stimulus-set repetition trials this would lead to enhanced priority to the rel-1053 evant stimulus-set leading to greater attentional weight to the relevant stimulus, and on 1054 1055 response-set repetition trials this would lead to enhanced bias toward the relevant response categories. If RCI impacts the proportion of the previous trial's response-set and stimulus-1056 set parameter values that carry-over into the next trial (e.g., via task-set decay) then the 1057 impact of the parameter carry-over reduces with increased RCI. Indeed, under certain as-1058 sumptions (see Simulation 5 in Appendix A) ECTVA predicts a reduction in response-set 1059 and stimulus-set switch cost with increasing RCI for response times, but the reduction in 1060 stimulus-set switch cost with increasing RCI is larger on response-incompatible trials. For 1061 the error data, ECTVA now predicts the presence of both a response-set and stimulus-set 1062 switch cost, the first simulation to demonstrate a component switch cost in error data. Both 1063 costs reduce with increasing RCI. Interestingly, ECTVA predicts stimulus-set switch costs 1064 only on response-incompatible trials, which has been observed in the majority of experi-1065 ments reported in the current paper. This effect decreases with increasing RCI, suggesting 1066 it is caused by the carry-over of the irrelevant stimulus-set from the previous trial at short 1067 RCIs leading to greater attentional weight on the irrelevant stimulus (see Equation 2). 1068

1069 These collective predictions were tested in Experiment 4.

1070 Method

Participants. 120 new participants were recruited as per previous experiments. 18
 participants were removed due to failure to maintain session-wise accuracy above 90%.

Materials & Procedure. The materials were identical to Experiment 3a (where CSI was manipulated). In the current experiment, on each trial the compound cue was presented for 250 ms before the imperative stimuli appeared. Once a response was registered, the frame went blank for an inter-trial interval (i.e, response-cue interval) of either 100 ms or 900 ms (which was randomised on each trial). Participants received 16 trials of practice before completing 10 blocks of 64 trials.

1079 **Results**

Error trimming removed 7.52% of the data, and RT trimming removed 2.91% of this error-trimmed data. The response time and error data can be seen in Figure 13.



Figure 13. Behavioural data from Experiment 4. The left plot shows mean response time (in milliseconds, ms) and the right plot shows proportion error as a function of response-set sequence and stimulus-set sequence. Error bars show one standard error around the mean.

Response time analysis. The model comparisons can be seen in Table 6. This 1082 showed extreme evidence in favour of RCI: RTs were slower overall at the short RCI (M =1083 1401 ms, SE = 14) than at the long RCI (M = 1314, SE = 14). As in previous experiments, 1084 there was extreme evidence in favour of both Response-Set Sequence and Stimulus-Set 1085 Sequence: Response-set repetitions (M = 1277 ms, SE = 13) were responded to faster than 1086 response-set switches (M = 1438, SE = 15), and stimulus-set repetitions (M = 1284, SE =1087 13) were responded to faster than stimulus-set switches (M = 1431, SE = 14). There was 1088 moderate evidence that the response-set switch cost reduced at longer RCIs, from 181 ms 1089 at short RCIs to 143 ms at long RCIs. There was extreme evidence that the stimulus-set 1090 switch cost reduced at longer RCIs, from 199 ms at the short RCIs to 95 ms at the long 1091 RCIs. 1092

There was extreme evidence for Response Compatibility: Response times were shorter for response-compatible trials (M = 1336, SE = 14) than for response-incompatible trials (M = 1379, SE = 14). However, there was anecdotal-to-moderate evidence that Response Compatibility was not involved in any interactions.

There again was extreme evidence for an interaction between Response-Set Sequence and Stimulus-Set Sequence. There was moderate evidence that this interaction was not further modulated by RCI or Compatibility, and there was moderate evidence against the four-way interaction.

Testing the predictions of task-set structure, the results (collapsing across compatibility and RCI) showed no difference in response time between a single-component switch of response-set and a single-component switch of stimulus-set ($BF_{10} = 0.21$), but there was extreme evidence for a difference between a single-component switch of response-set and a switch of both components ($BF_{10} = 2.45 \times 10^{15}$), and extreme evidence for a difference between a single-component switch of stimulus-set and a switch of both components ($BF_{10} = 1.09 \times 10^{13}$).

There was moderate evidence for no overall effect of RCI, but there Error analysis. 1108 was extreme evidence in favour of both Response-Set Sequence and Stimulus-Set Sequence: 1109 Response-set repetitions (M = 2.33%, SE = 0.11) were responded to more accurately than 1110 response-set switches (M = 4.30%, SE = 0.15), and stimulus-set repetitions (M = 2.97%, 1111 SE = 0.13) were responded to more accurately than stimulus-set switches (M = 3.65%, SE 1112 = 0.14). There was moderate evidence that Response-Set did not interact with RCI, but 1113 extreme evidence for an interaction between Stimulus-Set and RCI: The stimulus-set switch 1114 cost was 1.30% at short RCIs but was 0.07% at long RCIs. 1115

There was extreme evidence for an overall effect of Response Compatibility: accuracy was better on compatible trials (M = 2.85%, SE = 0.13) than on incompatible trials (M = 3.77%, SE = 0.14). There was moderate evidence for an interaction between Response Compatibility and Stimulus-Set Sequence: stimulus-set switch costs were 0.26% on compatible trials, and 1.10% on incompatible trials. There was anecdotal-to-moderate evidence against all other interaction effects.

Testing the predictions of task-set structure, the results (focussing on incompatible trials and collapsing across RCI) showed moderate evidence for a difference in errors between a single-component switch of response-set and a single-component switch of stimulusTable 6

Model comparison results for the behavioural data for Experiment 4. The Bayes factors (BF) show comparison of the full factorial model (including all main effects and all interactions) against models with particular predictors omitted (i.e., $BF = \frac{omitted-model}{full-model}$). BF values below 1 indicate evidence in favour inclusion of that predictor.

Omission	BF(RT)	BF (Error)
Compatibility (C) x R-Set Seq. (RS) x RCI x S-Set Seq. (SS)	4.07	5.66
$C \ge RS \ge SS$	3.87	2.27
$C \ge RCI \ge SS$	6.44	1.40
$C \ge RS \ge RCI$	5.04	2.20
$RS \ge RCI \ge SS$	5.40	6.18
$C \ge SS$	3.94	0.30
$C \ge RS$	2.41	8.36
C x RCI	4.13	8.43
$RS \ge SS$	8.33×10^{-9}	3.91
RCI x SS	9.41×10^{-12}	4.25×10^{-3}
RCI x RS	0.22	10.54
С	$5.20 imes 10^{-8}$	$2.91 imes 10^{-7}$
SS	2.94×10^{-86}	$9.81 imes 10^{-4}$
RS	9.51×10^{-101}	3.35×10^{-34}
RCI	5.33×10^{-33}	6.04

Note. Seq. = sequence.

set $(BF_{10} = 3.72)$, strong evidence for a difference between a single-component switch of response-set and a switch of both components $(BF_{10} = 24.89)$, and extreme evidence for a difference between a single-component switch of stimulus-set and a switch of both components $(BF_{10} = 1.30 \times 10^7)$.

1129 Discussion

Experiment 4 replicated all of the main findings from the previous experiments. In 1130 addition, Experiment 4 demonstrated an effect of RCI on component-switching. In par-1131 ticular, increasing the RCI led to a decrease in both the response-set switch cost and the 1132 stimulus-set switch cost in response times, although the effect was more pronounced for 1133 stimulus-set costs. The error data showed a reduction of the stimulus-set switch cost with 1134 increasing RCI, but no effect on the response-set switch cost. Together, these results provide 1135 strong evidence that the stimulus-set representation—and to a lesser extent, the response-1136 set representation—from the previous trial carries over at short RCIs and influences per-1137 formance on the current trial, either by facilitating performance on component-repetition 1138 trials or interfering on component-switch trials (or both). These results extend previous 1139 examinations of the effect of RCI on task switching performance (e.g., Meiran et al., 2000) 1140

by partitioning the RCI's effects on the different task-set components in multi-component switching, and are congruent with the conclusions of Horoufchin et al. (2011a).

These effects of RCI can be explained—with certain (possibly strong) assumptions— 1143 by ECTVA by allowing task-set parameters to decay across the RCI and contribute to the 1144 parameter values on the next trial; this reduces their impact on the next trial at longer RCIs 1145 for both response times and errors, although ECTVA predicted a reduction in response-1146 set switch costs in the error data at longer RCIs which was not observed. That RCI 1147 had a weaker effect on response-set switch costs (it had a moderate effect in RTs, but 1148 no evidence for an effect in errors) could be explained by the implementation of ECTVA 1149 in Simulation 5 (see Appendix A), where response-set parameter values were inhibited by 1150 reducing their activation by a fixed value at the end of every trial; thus, the effect of RCI 1151 might be reduced as there is less residual activation to act upon. In contrast, stimulus-set 1152 parameters were not inhibited. The original ECTVA model (Logan & Gordon, 2001) had 1153 a form of response inhibition (although it was implemented differently; see Appendix A), 1154 and response inhibition has been suggested to be a feature of control during task switching 1155 to avoid response perseveration (e.g., Druey, 2014; Druey & Hübner, 2008; Koch, Frings, & 1156 Schuch, 2018), so this is a plausible explanation. Further work is required on this. 1157

1158

1141

1142

General Discussion

The goal of the current research was to explore the nature of task-set representations, and in particular how response-set (i.e., how to respond) and stimulus-set (i.e., what to respond to) components of the task-set are represented and controlled when the environment signals a change in one or more component. Previous work has explored switching of stimulus-set alone (Gopher et al., 2000; Rushworth et al., 2005) and switching of responseset alone (Rushworth et al., 2002), but fewer studies have explored switching of response-set and stimulus-set (Kieffaber et al., 2013; Kleinsorge et al., 2004).

Component switching of this type allows testing of the behavioural predictions of the 1166 different accounts of the organisation of the task-set (see Figure 1): the *flat*-view (Vandieren-1167 donck et al., 2008) proposes that a change in any component requires reconfiguration of the 1168 complete task-set; the *integrated*-view (Philipp & Koch, 2010) also proposes that a change 1169 in any component requires complete reconfiguration, but task-set components are integrated 1170 first into a single task-representation before response selection; the *componential*-view (Hüb-1171 ner et al., 2001) proposes components are represented and switched independently with 1172 additive effects of component-switching; and the *hierarchical*-view (Kleinsorge et al., 2004) 1173 suggests components are organised hierarchically, with a switch at the higher level of the 1174 hierarchy leading to a switch at all levels. 1175

The current study took a different approach by grounding it within the theoretical framework of ECTVA (Logan & Gordon, 2001) where task-set components are modelled as independent control parameters that can be reconfigured in parallel. Simulations from ECTVA showed an under-additive interaction between response-set and stimulus-set sequences, mimicking the predictions of the integrated view of Philipp and Koch (2010). However, in contrast to the integrated-view, the under-additive interaction in ECTVA is

1182

1183

1184

1185

1186

1187

1188

1189

1190

diagnostic of parallel reconfiguration of the independent response-set and stimulus-set parameters coupled with the negatively-accelerating increase in reconfiguration time as more parameters require updating. In addition, the current study took advantage of the explicit specification of task-set components in ECTVA and their influence on response selection by exploring response-compatibility effects, which—to my knowledge—has not been explored in studies of component switching. ECTVA predicts slower and more error-prone responses when the response associated with the irrelevant stimulus is different (i.e., *incompatible*) to the response associated with the relevant stimulus. Such compatibility effects (or "crosstalk" in Logan & Gordon, 2001) provide valuable information as to the task-set rep-

resentation: despite the task-set prioritising the relevant stimulus, the irrelevant stimulus is still represented (albeit weakly) and thus influences response selection speed and accuracy.

1193 Summary of Results

Experiments 1a and 1b sought to establish a baseline test of response-set and stimulus-1194 set component-switching. An under-additive interaction between response-set and stimulus-1195 set sequence was observed, as well as effects of response-compatibility. Experiment 2 sought 1196 to establish the extent to which the component-switch costs were driven by cue-switching 1197 by using two cues per set (Logan & Bundesen, 2003; Mayr & Kliegl, 2003). Previous studies 1198 of component-switching have only used one cue per set, meaning the effects of component-1199 switching cannot be disentangled from the effects of cue-switching. Experiment 2 showed 1200 clear response-set and stimulus-set switch costs, suggesting these were not driven by cue-1201 switching effects. However, the interaction—although present and under-additive—was 1202 weaker than in Experiments 1a and 1b. There were clear effects of response-compatibility 1203 in both RT and error data. 1204

Experiment 3 examined the effects of general and component-specific preparation on 1205 switching performance. As the data from Experiments 1 and 2 were largely compatible with 1206 both the predictions of the integrated view and of ECTVA, the manipulations of Experiment 1207 3 provide a stronger test of the predictions of ECTVA which predicts that opportunity for 1208 general preparation reduces both the stimulus-set and response-set switch costs, as well 1209 as their interaction (Simulation 3, Appendix A), and that component-specific preparation 1210 reduces the switch-cost selectively for the prepared component (Simulation 4, Appendix A). 1211 In Experiment 3a, general preparation via the cue-stimulus interval showed a reduction in 1212 response-set and stimulus-set switch costs—as well as a reduction in their interaction—at 1213 longer CSIs, as predicted by ECTVA. Experiments 3b and 3c provided opportunity for 1214 component-specific preparation by cuing one component of the task-set ahead of stimulus 1215 onset. In contrast to predictions from ECTVA, opportunity for preparation of a task-set 1216 component did not lead to a selective reduction in the switch cost associated with that 1217 component (see also Appendix B). However, this finding could be reconciled with ECTVA 1218 if the assumption is added that parallel parameter reconfiguration does not commence until 1219 both task-set components have been encoded. It is interesting to note that this is also 1220 an assumption that is implicitly made by the integrated-view of Philipp and Koch (2010). 1221 as both task-set components need to be encoded before they can be integrated into a 1222 single representation. This additional assumption requires further empirical and theoretical 1223

testing. Note that there were no observed effects of general (Experiment 3a) or specific preparation (Experiments 3b–3c) on the compatibility effect, as predicted by ECTVA: Taskset component preparation does not impact the compatibility effect in ECTVA because preparation affects the *setting* of task-set components, and compatibility effects arise at a later stage of processing (i.e., during response selection).

Experiment 4 examined the effects of interference from the previous task-set on the 1229 current trial's performance via the response-cue interval. No study on component-switching 1230 has examined the effect of RCI on component switch costs, but Experiments 2 and 3a-1231 3c reported here provided some evidence that the previous trial's stimulus-set influenced 1232 performance on the current trial (driven by the observation of stimulus-set switch costs only 1233 on incompatible trials). This is congruent with the conclusions of Horoufchin et al. (2011a) 1234 ho reported stronger RCI effects for stimulus-related components of the task-set. With 1235 certain additional assumptions, ECTVA predicts a reduction of response-set and stimulus-1236 set switch cost in both the RT and error data, driven by decay of the task-set components 1237 from the previous trial which therefore leads to a reduction in interference on component-1238 switch trials. Interestingly, ECTVA predicted stimulus-set switch costs in the error data 1239 only for incompatible trials, as had been observed previously. The results of Experiment 4 1240 largely confirmed the predictions of ECTVA: For response-set switches, there was a small 1241 reduction in RT switch cost at longer RCI, but no reduction in the error switch cost; in 1242 contrast, for stimulus-set switches, there was a large reduction in both RT and error switch 1243 cost. Thus, in the current design, there is evidence that the previous trial's stimulus-set 1244 (and to a lesser extent, the response-set) parameters carry-over and influence performance 1245 on the current trial. 1246

1247 On Task-Set Representation

Overall, the results of this study are compatible with the view that control of stimulus-1248 set and response-set in task switching is achieved by the task-set holding independent repre-1249 sentations of each component, and that these components can be independently reconfigured 1250 (Kieffaber et al., 2013), although there are limits on the extent to which specific prepara-1251 tion of a single-component is possible. That RCI had differential effects on response-set 1252 and stimulus-set switch costs provides converging evidence for the independence of these 1253 components in the task-set representation, and therefore speak against the integration of 1254 these components into a single task-set representation. There was no evidence in any of the 1255 experiments for the *flat*-view, the *componential*-view, or the *hierarchical*-view. Instead, the 1256 data are broadly compatible with ECTVA and independent component-representations. 1257

To provide a further test of the predictions of task-set structure, I conducted a series of 1258 Bayesian meta analyses to combine data across all Experiments (see Appendix D for details). 1259 The results showed no difference in response times between a single-component switch of 1260 response-set and a single-component switch of stimulus-set (Figure 14A). This is predicted 1261 by the flat, integrated, and componential view of task-set structure, but is not predicted 1262 by the hierarchical-view. It is also predicted by ECTVA if response-set is represented at 1263 a higher level of task categories (e.g., "magnitude") rather than at individual response 1264 options (e.g., "lower than 5"). This reduces the number of response-set β parameters to 1265

two (one for each task category), equating it with the number of stimulus-set π parameters; 1266 the net effect is that response-set switches are just as costly as stimulus-set switches (see 1267 Figure 3). There is precedence in the literature for representing response-set at the task 1268 category level (Logan & Schneider, 2010), but ECTVA originally represented response-set 1269 at the individual response option level (Logan & Gordon, 2001). Further work is needed 1270 to understand whether response-set is represented at the task level or the response option 1271 level, but the data in the current study appear compatible with it being represented at the 1272 task category level. 1273



Figure 14. Forest plots of the Bayesian meta-analyses (see Appendix D for details). Each plot shows the meta-analysis for a different response time (RT) contrast: (A) a single-component switch of response-set vs. a single-component switch of stimulus-set; (B) a single-component switch of response-set vs. a switch of both response-set and stimulus-set, and (C) a single-component switch of stimulus-set vs. a switch of both components. The density plots show the posterior distribution of the estimate of each experiment / condition's contrast, the filled points and error bars show the mean and 95% credible intervals of the distribution, respectively, and the open red circles show the observed contrast for each experiment / condition. ss = stimulus-set; rs = response-set.

The meta analyses also showed that a switch of both response-set and stimulus-1274 set was more costly than (a) just a switch of the response-set component (Figure 14B), 1275 and (b) just a switch of the stimulus-set component (Figure 14C). This pattern of results 1276 is not compatible with either the flat-view or the hierarchical-view of task-set structure. 1277 However, it is compatible with the componential-view, as well as the integrated-view and 1278 the predictions of ECTVA. The consistent presence of an under-additive interaction, though, 1279 rules out the componential-view (Hübner et al., 2001) which proposes that the switch cost 1280 associated with a change in both components is additive to the effects of costs associated 1281 with single-component switches. 1282

1283 Limitations & Future Research

Component-specific preparation. One clear limitation with the conclusions of 1284 the current study is that not all data confirmed the predictions of ECTVA. Most notably, 1285 as reported in Experiments 3b and 3c, opportunity for advanced preparation of a switch of 1286 either response-set or stimulus-set did not lead to a reduction in the switch cost associated 1287 with that component. Philip and Koch (2010) observed a similar lack of component-specific 1288 preparation effects in their study, and interpreted this as evidence in support of individual 1289 components being integrated in the task-set. ECTVA's prediction of component-specific 1290 preparation arises naturally from the assumption of independent parameters controlling 1291 response-set and stimulus-set, but this leads to a clear prediction of component-specific 1292 preparation leading to a reduction in switch cost associated with that component. One 1293 possibility is that response-set and stimulus-set are independent control parameters—and 1294 therefore component-specific preparation is possible—but on a proportion of trials the sys-1295 tem fails to take advantage of the potential for component-specific preparation. Some 1296 evidence in favour of this "failure to engage" hypothesis (DeJong, 2000) was found in dis-1297 tributional analysis (see Appendix C), but failed to replicate in an additional Experiment 1298 (Experiment 3d reported in Appendix C). Another possibility is that as these experiments 1299 require frequent updating of response-set and stimulus-set, it becomes beneficial for the cog-1300 nitive system to wait until information about both response-set and stimulus-set is available 1301 before commencing independent and parallel reconfiguration of these parameters. Such an 1302 account, of course, requires rigorous testing in new experiments, so this is an important 1303 area for future research and theoretical development. 1304

On the integration of task-set components. Regarding integration of task-1305 set components, on the surface it may appear that ECTVA also assumes integration of 1306 independent task-set components into a single representation, and is therefore equivalent 1307 to the integrated-view of Philipp and Koch (2010). Although in the current study I have 1308 discussed ECTVA as a single model, it comprises multiple parts: ECTVA is an abstract 1309 specification of the computations that an executive control process must execute in order to 1310 program TVA to select relevant objects and the response to those objects. In terms of the 1311 current study, ECTVA must interpret the task cues and set appropriate parameter values for 1312 response-set and stimulus-set. These parameter values must then be transmitted to TVA. 1313 independently and in parallel (Logan & Gordon, 2001). The task-set is represented at the 1314 propositional level in working memory, where representations of the current trial's task 1315 requirements are held. This propositional representation could be considered "integrated" 1316 in that it must contain information relating to which object to respond to (i.e., stimulus-set) 1317 and how to respond to that object (response-set). 1318

The task-set is specified at the parameter level in TVA, and it could be argued there is a degree of "intergration" here, too.¹⁶ Recall from Equation 1 that the probability of categorising a stimulus as belonging to response categories is the multiplicative combination of sensory evidence from the stimulus, attentional bias to each response category (i.e., response-set), and the relative attention weight to the stimulus (which is influenced by stimulus-set). However, response-set and stimulus-set contribute independently in this

¹⁶I am grateful to Gordon Logan for clarification of these points.

equation; there is no explicit binding of the task-set components into a single task representation (cf., Philipp & Koch, 2010). That said, they could be considered "integrated" in as much as they work together (via Equation 1) to select objects and actions to these objects.

However, the key theoretical distinction between the integrated account and the 1328 ECTVA account centers on how task-set components are updated. In Philipp and Koch's 1329 (2010) integration account, response-set and stimulus-set are integrated into a single task-1330 set representation; when the environment cues a switch in either component, the entire 1331 task-set must be updated. In contrast, in ECTVA, response-set and stimulus-set are up-1332 dated independently, meaning that a change in one component does not necessitate the 1333 updating of the other; rather, ECTVA transmits to TVA only the parameter values that 1334 need updating. Indeed, that not all parameters require updating when a single compo-1335 nent switches in ECTVA is what gives rise to the under-additive interaction observed in 1336 all experiments in the current study (see Figure 3A & B). This assumption of independent 1337 updating of task-set components is what differentiates ECTVA from the view of Philipp and 1338 Koch (2010). Further theoretical work is required to arbitrate between these accounts. This 1339 would require the development of a detailed (mathematical) specification of the integrated 1340 view so that predictions can be derived and tested empirically. This remains an important 1341 avenue for future research. 1342

Discrepant findings across studies. Further work is also needed to understand 1343 the discrepant findings of task-set structure across studies. Although the current study 1344 found no evidence to support the flat-, hierarchical-, or componential-view, previous studies 1345 have reported evidence to support each. Experimental paradigm differences—and conse-1346 quently differences in relevant task-set components—may explain these discrepant findings. 1347 Whilst the majority of the studies have one component similar to the response-set ma-1348 nipulated in the current study, the nature of the other component that can change from 1349 trial-to-trial differs across studies. The nature of the component obviously changes how 1350 that component is represented in the task-set, which could explain the differences across 1351 studies. For example, evidence in support of the flat-view (Vandierendonck et al., 2008) 1352 used Stroop-like stimuli and global-local stimuli; one of the components that was varied 1353 was which stimulus dimension to attend to (e.g., attend to global, Allport et al., 1994). In 1354 contrast, the majority of evidence for the hierarchical view comes from paradigms where 1355 one of the components was the judgement-response mapping (compatible vs. incompatible, 1356 Kleinsorge & Heuer, 1999; Kleinsorge, Heuer, & Schmidtke, 2001, 2002). Further still, ev-1357 idence for the integrated view comes from a study where one of the components was the 1358 response modality required (vocal vs. manual responses, Philipp & Koch, 2010). It remains 1359 possible that such such differences can explain the discrepant findings, so a fruitful avenue 1360 for future research would be to explore the effect of these paradigm choices in a single study. 1361

However, Kleinsorge (2004) reported evidence supporting the hierarchical-view in a paradigm that factorially manipulated response-set (odd/even vs. lower/higher than five) and stimulus-set (attend to numerical value of digit vs. attend to number of digits presented) in a way not so disimilar to the current study. Although there are paradigm differences between the current study and that of Kleinsorge (2004), it is not clear why these (minor) differences should lead to different findings. One possibility is that the current study used transparent cues for both response-set and stimulus-set, whilst Kleinsorge used non-

1369

and response-set was cued by the location of the stimuli). Cue transparency has been shown to impact other task switching phenomena (e.g., Grange & Houghton, 2010b, 2010a; Houghton, Pritchard, & Grange, 2009; Jost et al., 2013) so it may also influence how the task-set is organised. In addition, the study of Kleinsorge (2004) presented greater opportunity for practice effects (e.g., 20 blocks of 67 trials, compared to 6–10 blocks of 64 trials in the current study) which may also influence outcomes. Both of these possibilities remain interesting avenues for future research.¹⁷

Task-sets and different modalities. Future research could also explore the extent 1377 to which task-set components are updated independently in other modalities. For example, 1378 Seibold and colleagues (Seibold, Nolden, Oberem, Fels, & Koch, 2018, 2019) have utilised a 1379 component switching paradigm in a dichotomous listening task where one component per-1380 tained to auditory attention selection criteria (e.g., attending to the male vs. female speaker, 1381 or attending to the left vs. the right ear) and the other component of type of "judgement" 1382 (i.e., response set; odd vs. even judgement on number stimuli, and vowel vs. consonant 1383 judgement on letter stimuli). Results generally showed an under-additive interaction as 1384 in the current study, with smaller response-set switch costs when the attention compo-1385 nent switched than when it repeated. However, the authors interpreted this interaction 1386 as reflecting non-independence of the processing of task-set components. In addition, the 1387 authors related their findings to the binding of task-set components: if the attention se-1388 lection component (e.g., "attend left") is bound with the judgement (e.g., "odd vs. even"), 1389 then a switch in either component on the next trial can lead to so-called partial repetition 1390 costs (e.g., Hommel, 2004), reflecting the interference caused by the mismatch between the 1391 demands of the current trial and the retrieved episode from the previous trial; however, 1392 when both components switch there is no interference between the demands of the current 1393 trial and the previous trial's episode, leading to faster responding. Such an account would 1394 predict faster response times when both components switch than when either component 1395 switches in isolation, producing the interaction ascribed to the hierarchical view (see Fig-1396 ure 2D)), which was observed in some (Experiment 3 of Seibold et al., 2019) but not all 1397 experiments. 1398

However, the under-additive pattern observed in many of the experiments of Seibold 1399 and colleagues (Seibold et al., 2018, 2019) has in the current study been shown to also 1400 be produced by independent updating of task-set components. In addition, none of the 1401 experiments reported in the current study produced data suggestive of episodic retrieval 1402 effects. It may be that task-set organisation and the updating of task-set components— 1403 and the involvement of episodic binding—depends on the modality of the input of task 1404 components. Exploring the interplay of component modality and episodic retrieval could 1405 prove valuable in unpicking the nature of task-set representation. 1406

1407

Open Research Practices Statement

All raw data, analysis scripts, and model code can be downloaded from https://osf. io/ayutg/.

¹⁷Thank you to Thomas Kleinsorge for raising these possibilities.

1410	Acknowledgements
1411 1412 1413 1414 1415	I am grateful to Gordon Logan, Thomas Kleinsorge, and Iring Koch for their helpful comments during the review process. I also am grateful to Stefanie Schuch for fruitful discussions about various aspects of this work. Components of the work were presented at the 2022 and 2023 Spring meetings of the Experimental Psychology Society (at Keele University UK and Plymouth University UK, respectively)
1416	Funding
1417	This work was supported by a grant from the Experimental Psychology Society, UK.
1418	Declaration of Competing Interests
1419 1420	The author declared that there are no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In Attention and Performance Series. Attention and performance
 1424 15: Conscious and nonconscious information processing (pp. 421–452). Cambridge,
- 1425 MA, US: The MIT Press.
- Altmann, E. M. (2004). Advance Preparation in Task Switching: What Work Is Being Done? *Psychological Science*, 15(9), 616–622. https://doi.org/10.1111/j.09567976.2004.00729.x
- Altmann, E. M. (2005). Repetition priming in task switching: Do the benefits dissipate?
 Psychonomic Bulletin & Review, 12(3), 535–540. https://doi.org/10.3758/BF03193801
- Altmann, E. M., & Gray, W. D. (2008). An integrated model of cognitive control in
 task switching. *Psychological Review*, 115(3), 602–639. https://doi.org/10.1037/0033295X.115.3.602
- Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020).
 Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, 52(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x
- Aust, F., & Barth, M. (2022). papaja: Prepare reproducible APA journal articles with R
 Markdown. Retrieved from https://github.com/crsh/papaja
- Barth, M. (2022). tinylabels: Lightweight variable labels. Retrieved from https://cran.r project.org/package=tinylabels
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models
 using lme4. Journal of Statistical Software, 67(1), 1–48. https://doi.org/10.18637/jss.
 v067.i01
- Bates, D., Maechler, M., & Jagan, M. (2022). Matrix: Sparse and dense matrix classes and
 methods. Retrieved from https://CRAN.R-project.org/package=Matrix
- Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, 97(4), 523–547.
 https://doi.org/10.1037/0033-295X.97.4.523
- Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan.
 Journal of Statistical Software, 80(1), 1–28. https://doi.org/10.18637/jss.v080.i01
- ¹⁴⁵⁰ Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. ¹⁴⁵¹ The R Journal, 10(1), 395–411. https://doi.org/10.32614/RJ-2018-017
- Bürkner, P.-C. (2021). Bayesian item response modeling in R with brms and Stan. Journal
 of Statistical Software, 100(5), 1–54. https://doi.org/10.18637/jss.v100.i05
- DeJong, 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.
- Druey, M. D. (2014). Stimulus-category and response-repetition effects in task switching:
 An evaluation of four explanations. Journal of Experimental Psychology. Learning,
 Memory, and Cognition, 40(1), 125–146. https://doi.org/10.1037/a0033868
- Druey, M. D., & Hübner, R. (2008). Response inhibition under task switching: Its strength
 depends on the amount of task-irrelevant response activation. *Psychological Research*,
 72(5), 515–527. https://doi.org/10.1007/s00426-007-0127-1
- ¹⁴⁶³ Eddelbuettel, D., & Balamuta, J. J. (2018). Extending extitR with extitC++: A Brief ¹⁴⁶⁴ Introduction to extitRcpp. *The American Statistician*, 72(1), 28–36. https://doi.org/

1421

- 10.1080/00031305.2017.1375990 1465
- Eddelbuettel, D., & François, R. (2011). Rcpp: Seamless R and C++ integration. Journal 1466 of Statistical Software, 40(8), 1–18. https://doi.org/10.18637/jss.v040.i08 1467
- Gopher, D., Armony, L., & Greenshpan, Y. (2000). Switching tasks and attention policies. 1468 Journal of Experimental Psychology: General, 129(3), 308–339. https://doi.org/10. 1469

1037/0096-3445.129.3.308 1470

- Grange, J. A. (2016). Temporal distinctiveness in task switching: Assessing the mixture-1471 distribution assumption. Frontiers in Psychology, 7. https://doi.org/10.3389/fpsyg. 1472 2016.002511473
- Grange, J. A. (2022). Trimr: An implementation of common response time trimming 1474 *methods*. Retrieved from https://CRAN.R-project.org/package=trimr 1475
- Grange, J. A., & Cross, E. (2015). Can time-based decay explain temporal distinctiveness 1476 effects in task switching? The Quarterly Journal of Experimental Psychology, 68(1), 1477 19-45. https://doi.org/10.1080/17470218.2014.934696 1478
- Grange, J. A., & Houghton, G. (2010a). Cue-switch costs in task-switching: Cue priming 1479 or control processes? Psychological Research, 74(5), 481–490. https://doi.org/10.1007/ 1480 s00426-009-0270-y 1481
- Grange, J. A., & Houghton, G. (2010b). Heightened conflict in cue-target translation 1482 increases backward inhibition in set switching. Journal of Experimental Psychol-1483 ogu. Learning, Memory, and Cognition, 36(4), 1003–1009. https://doi.org/10.1037/ 1484 a0019129 1485
- Grange, J. A., & Houghton, G. (2011). Task preparation and task inhibition: A comment on 1486 koch, gade, schuch, & philipp (2010). Psychonomic Bulletin & Review, 18(1), 211–216. 1487 https://doi.org/10.3758/s13423-010-0023-3 1488
- Grange, J. A., & Houghton, G. (Eds.). (2014). Task Switching and Cognitive Control. New 1489 York, NY: Oxford University Press. 1490
- Grolemund, G., & Wickham, H. (2011). Dates and times made easy with lubridate. Journal 1491 of Statistical Software, 40(3), 1–25. Retrieved from https://www.jstatsoft.org/v40/i03/ 1492
- Hommel, B. (2004). Event files: Feature binding in and across perception and action. 1493
- Trends in Cognitive Sciences, 8(11), 494–500. https://doi.org/10.1016/j.tics.2004.08. 1494 0071495
- Horoufchin, H., Philipp, A. M., & Koch, I. (2011a). Temporal distinctiveness and repe-1496 tition benefits in task switching: Disentangling stimulus-related and response-related 1497 contributions. Quarterly Journal of Experimental Psychology, 64(3), 434–446. https: 1498 //doi.org/10.1080/17470218.2010.496857 1499
- Horoufchin, H., Philipp, A. M., & Koch, I. (2011b). The dissipating task-repetition 1500 benefit in cued task switching: Task-set decay or temporal distinctiveness? Journal 1501 of Experimental Psychology. Human Perception and Performance, 37(2), 455–472. 1502 https://doi.org/10.1037/a0020557
- 1503
- Houghton, G., Pritchard, R., & Grange, J. A. (2009). The role of cue-target translation in 1504 backward inhibition of attentional set. Journal of Experimental Psychology. Learning, 1505 Memory, and Cognition, 35(2), 466–476. https://doi.org/10.1037/a0014648 1506
- Hübner, R., Futterer, T., & Steinhauser, M. (2001). On attentional control as a source of 1507 residual shift costs: Evidence from two-component task shifts. Journal of Experimental 1508
- Psychology: Learning, Memory, and Cognition, 27(3), 640–653. https://doi.org/10. 1509

- 1510 1037/0278-7393.27.3.640
- Jost, K., De Baene, W., Koch, I., & Brass, M. (2013). A review of the role of cue processing
 in task switching. Zeitschrift für Psychologie, 221(1), 5–14. https://doi.org/10.1027/
 2151-2604/a000125
- Kay, M. (2023). tidybayes: Tidy data and geoms for Bayesian models. https://doi.org/10.
 5281/zenodo.1308151
- Kieffaber, P. D., Kruschke, J. K., Cho, R. Y., Walker, P. M., & Hetrick, W. P. (2013).
 Dissociating stimulus-set and response-set in the context of task-set switching. *Journal of Experimental Psychology: Human Perception and Performance*, 39(3), 700–719. https://doi.org/10.1037/a0029545
- Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch,
 I. (2010). Control and interference in task switching—A review. *Psychological Bulletin*,
 136(5), 849–874. https://doi.org/10.1037/a0019842
- Kleinsorge, T. (2004). Hierarchical switching with two types of judgment and two stimulus
 dimensions. *Experimental Psychology*, 51, 145–149. https://doi.org/10.1027/1618-3169.
 51.2.145
- Kleinsorge, T., & Heuer, H. (1999). Hierarchical switching in a multi-dimensional task
 space. Psychological Research, 62(4), 300–312. https://doi.org/10.1007/s004260050060
- Kleinsorge, T., Heuer, H., & Schmidtke, V. (2001). Task-set reconfiguration with binary
 and three-valued task dimensions. *Psychological Research*, 65(3), 192–201. https://doi.
 org/10.1007/s004260000051
- Kleinsorge, T., Heuer, H., & Schmidtke, V. (2002). Processes of task-set reconfiguration:
 Switching operations and implementation operations. Acta Psychologica, 111(1), 1–28.
 https://doi.org/10.1016/S0001-6918(01)00076-2
- Kleinsorge, T., Heuer, H., & Schmidtke, V. (2004). Assembling a task space: Global
 determination of local shift costss. *Psychological Research*, 68(1), 31–40. https://doi.
 org/10.1007/s00426-003-0134-9
- Koch, I., Frings, C., & Schuch, S. (2018). Explaining response-repetition effects in task
 switching: Evidence from switching cue modality suggests episodic binding and response
 inhibition. *Psychological Research*, 82(3), 570–579. https://doi.org/10.1007/s00426017-0847-9
- Koch, I., Poljac, E., Müller, H., & Kiesel, A. (2018). Cognitive structure, flexibility, and
 plasticity in human multitasking-An integrative review of dual-task and task-switching
 research. *Psychological Bulletin*, 144 (6), 557–583. https://doi.org/10.1037/bul0000144
- Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian cognitive modelling: A practical course*. Cambridge University Press.
- Logan, G. D. (2003). Executive control of thought and action: In search of the wild
 homunculus. Current Directions in Psychological Science, 12(2), 45–48. https://doi.
 org/10.1111/1467-8721.01223
- Logan, G. D., & Bundesen, C. (2003). Clever homunculus: Is there an endogenous act
 of control in the explicit task-cuing procedure? Journal of Experimental Psychology:
 Human Perception and Performance, 29, 575–599. https://doi.org/10.1037/0096-1523.
 29.3.575
- Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dualtask situations. *Psychological Review*, 108(2), 393–434. https://doi.org/10.1037/0033-

1555 295x.108.2.393

- Logan, G. D., & Schneider, D. W. (2010). Distinguishing reconfiguration and compound-cue
 retrieval in task switching. 50(3-4), 413. https://doi.org/10.5334/pb-50-3-4-413
- Mayr, U., & Kliegl, R. (2003). Differential effects of cue changes and task changes on
 task-set selection costs. Journal of Experimental Psychology: Learning, Memory, and
 Cognition, 29(3), 362–372. https://doi.org/10.1037/0278-7393.29.3.362
- Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22(6), 1423–1442. https://doi.org/10.1037/0278-7393.22.6.1423
- Meiran, N. (2000). Modeling cognitive control in task-switching. *Psychological Research*, 63(3), 234–249. https://doi.org/10.1007/s004269900004
- Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. Cog *nitive Psychology*, 41(3), 211–253. https://doi.org/10.1006/cogp.2000.0736
- ¹⁵⁶⁸ Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134–140. https: ¹⁵⁶⁹ //doi.org/10.1016/S1364-6613(03)00028-7
- Monsell, S., & Mizon, G. A. (2006). Can the task-cuing paradigm measure an endogenous
 task-set reconfiguration process? Journal of Experimental Psychology: Human Perception and Performance, 32, 493–516. https://doi.org/10.1037/0096-1523.32.3.493
- ¹⁵⁷³ Müller, K., & Wickham, H. (2023). *Tibble: Simple data frames*. Retrieved from https: ¹⁵⁷⁴ //CRAN.R-project.org/package=tibble
- Nieuwenhuis, S., & Monsell, S. (2002). Residual costs in task switching: Testing the failureto-engage hypothesis. *Psychonomic Bulletin & Review*, 9(1), 86–92. https://doi.org/
 10.3758/BF03196259
- Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, 104, 266–300. https://doi.org/10.1037/0033-295X.
 104.2.266
- Pedersen, T. L. (2022). Patchwork: The composer of plots. Retrieved from https://CRAN.
 R-project.org/package=patchwork
- Philipp, A. M., & Koch, I. (2010). The integration of task-set components into cognitive task
 representations. *Psychologica Belgica*, 50(3-4), 383–411. https://doi.org/10.5334/pb50-3-4-383
- R Core Team. (2022). R: A language and environment for statistical computing. Vi enna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R project.org/
- Rangelov, D., Töllner, T., Mueller, H., & Zehetleitner, M. (2013). What are task-sets:
 A single, integrated representation or a collection of multiple control representations? *Frontiers in Human Neuroscience*, 7.
- Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124, 207–231. https://doi.org/10. 1037/0096-3445.124.2.207
- Rouder, J. N., Morey, R. D., Verhagen, J., Swagman, A. R., & Wagenmakers, E.-J. (2017).
 Bayesian analysis of factorial designs. *Psychological Methods*, 22(2), 304–321. https: //doi.org/10.1037/met0000057
- Rushworth, M. F. S., Passingham, R. E., & Nobre, A. C. (2002). Components of switching
- intentional set. Journal of Cognitive Neuroscience, 14(8), 1139–1150. https://doi.org/

1600 10.1162/089892902760807159

Rushworth, M. F. S., Passingham, R. E., & Nobre, A. C. (2005). Components of attentional
set-switching. *Experimental Psychology*, 52(2), 83–98. https://doi.org/10.1027/16183169.52.2.83

- Schneider, D. W. (2016). Investigating a method for reducing residual switch costs in cued
 task switching. Memory & Cognition, 44(5), 762–777. https://doi.org/10.3758/s13421016-0590-2
- 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. https://doi.org/10.1037/0096-3445.134.3.343
- Schneider, D. W., & Logan, G. D. (2007). Defining task-set reconfiguration: The case
 of reference point switching. *Psychonomic Bulletin & Review*, 14(1), 118–125. https:
 //doi.org/10.3758/BF03194038
- Schneider, D. W., & Logan, G. D. (2014). Tasks, task sets, and the mapping between
 them. In *Task switching and cognitive control* (pp. 27–44). New York, NY, US: Oxford
 University Press. https://doi.org/10.1093/acprof:osobl/9780199921959.003.0002
- Seibold, J. C., Nolden, S., Oberem, J., Fels, J., & Koch, I. (2018). Auditory attention switch ing and judgment switching: Exploring multicomponent task representations. Attention,
 Perception, & Psychophysics, 80(7), 1823–1832. https://doi.org/10.3758/s13414-018 1557-0
- Seibold, J. C., Nolden, S., Oberem, J., Fels, J., & Koch, I. (2019). The binding of an auditory target location to a judgement: A two-component switching approach. *Quarterly Journal of Experimental Psychology*, 72(8), 2056–2067. https://doi.org/10.1177/1747021819829422
- Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2022). Afex: Analy sis of factorial experiments. Retrieved from https://CRAN.R-project.org/package=afex
- Vandierendonck, A., Christiaens, E., & Liefooghe, B. (2008). On the representation of task
 information in task switching: Evidence from task and dimension switching. *Memory*& Cognition, 36(7), 1248–1261. https://doi.org/10.3758/MC.36.7.1248
- Vandierendonck, A., Liefooghe, B., & Verbruggen, F. (2010). Task switching: Interplay
 of reconfiguration and interference control. *Psychological Bulletin*, 136(4), 601–626.
 https://doi.org/10.1037/a0019791
- Venables, W. N., & Ripley, B. D. (2002). Modern applied statistics with s (Fourth). New
 York: Springer. Retrieved from https://www.stats.ox.ac.uk/pub/MASS4/
- Verbruggen, F., Liefooghe, B., Vandierendonck, A., & Demanet, J. (2007). Short cue
 presentations encourage advance task preparation: A recipe to diminish the residual
 switch cost. Journal of Experimental Psychology: Learning, Memory, and Cognition,
 33, 342–356. https://doi.org/10.1037/0278-7393.33.2.342
- ¹⁶³⁸ Vuorre, M. (2016). Bayesian meta-analysis with R, Stan, and brms.
- Wagenmakers, E.-J., Krypotos, A.-M., Criss, A. H., & Iverson, G. (2012). On the interpre tation of removable interactions: A survey of the field 33 years after Loftus. *Memory & Cognition*, 40(2), 145–160. https://doi.org/10.3758/s13421-011-0158-0
- Wickham, H. (2016). ggplot2: Elegant graphics for data analysis. Springer-Verlag New
 York. Retrieved from https://ggplot2.tidyverse.org
- ¹⁶⁴⁴ Wickham, H. (2023a). Forcats: Tools for working with categorical variables (factors). Re-

- trieved from https://CRAN.R-project.org/package=forcats
- ¹⁶⁴⁶ Wickham, H. (2023b). Stringr: Simple, consistent wrappers for common string operations.
- ¹⁶⁴⁷ Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., ... Yutani,
- H. (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686.
 https://doi.org/10.21105/joss.01686
- Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). Dplyr: A grammar
 of data manipulation. Retrieved from https://CRAN.R-project.org/package=dplyr
- Wickham, H., & Henry, L. (2023). Purre: Functional programming tools. Retrieved from
 https://CRAN.R-project.org/package=purre
- Wickham, H., Hester, J., & Bryan, J. (2023). Readr: Read rectangular text data. Retrieved
 from https://CRAN.R-project.org/package=readr
- Wickham, H., Vaughan, D., & Girlich, M. (2023). *Tidyr: Tidy messy data*. Retrieved from
 https://CRAN.R-project.org/package=tidyr

Appendix A — Executive Control of the Theory of Visual Attention (ECTVA) Simulations

¹⁶⁶⁰ Further details of ECTVA

The main details of ECTVA was outlined in the introduction. Here more detail is provided. In the current experiment design, categorisation of a stimulus into one of the potential response categories (*odd*, *even*, *low*, and *high*) is modelled as a race among all alternatives, where the speed (or "rate", in ECTVA terminology) with which each potential categorisation¹⁸ is made is determined by the multiplicative combination of sensory evidence and response-set bias. Formally, the rate, v, at which a single stimulus x is categorised as belonging to response category i is given by

$$v(x,i) = \eta(x,i)\beta_i.$$
(6)

The mean finishing time for the race is the reciprocal of the processing rates of all possible categorisations entering the race, given by

$$t_{win} = \frac{1}{\sum_{j \in R} \eta(x, j)\beta_j},\tag{7}$$

and the probability p of selecting response category i in the race rather than each of the other response categories j in the set of possibilities R is given by

$$p(i|x) = \frac{\eta(x,i)\beta_i}{\sum_{j \in R} \eta(x,j)\beta_j}.$$
(8)

Dealing with two stimuli. As explained in the introduction, categorisation of the stimulus is complicated by the fact two stimuli are present (x and y). ECTVA handles such situations by establishing an attention weight, w, to each stimulus in the display D, which is partly determined by sensory evidence η from the environment that stimulus x has a particular feature k (i.e., $\eta(x, k)$). As an example, for the orange digit $\eta(x, orange)$ would

¹⁸For clarity it is important to emphasise that ECTVA attempts to categorise the stimulus into *all* possible response categories, but the sensory evidence and response-set bias will ensure that the correct categorisation has the highest processing rate and therefore the highest chance of winning the race. As an example, when presented with the digit 7 and the cue "odd/even", there will be a race among alternatives to categorise the digit as belonging to the response categorise odd, even, low, and high As the sensory evidence for the response categories *odd* and *high* will be higher than the sensory evidence for *even* and *low*, and because—due to the cue—response-set bias will be higher for the response categories *odd* and *even*, when combined via Equation 1 the rate parameter will be higher for the response category *odd*—i.e., v(7, odd)—and therefore this categorisation will have a higher probability to win the race.

¹⁶⁷⁷ be high and $\eta(x, purple)$ would be low. This sensory evidence is combined multiplicatively ¹⁶⁷⁸ with the attentional priority (i.e., top-down bias) of selecting stimuli with features in the ¹⁶⁷⁹ relevant stimulus-set, π_k . This combination provides the *absolute* attention weight to the ¹⁶⁸⁰ stimulus x:

$$w_x = \sum_{k \in S} \eta(x, k) \pi_k.$$
(9)

The *relative* attention weight given to stimulus x is the ratio of this absolute weight to the total absolute attention weights of all stimuli z in the display D, given by

$$\frac{w_x}{\sum\limits_{z\in D} w_z} = \frac{\sum\limits_{k\in S} \eta(x,k)\pi_k}{\sum\limits_{z\in D} \sum\limits_{k\in S} \eta(z,k)\pi_k},\tag{10}$$

and is used to ascertain the overall weight of attention paid to stimulus x. The overall processing rate of categorising stimulus x as belonging to category i is then the multiplicative combination of (a) the sensory evidence that x belongs to category i, $\eta(x, i)$; (b) the topdown bias to category i, β_i ; and (c) the relative attention weight to stimulus x

$$v(x,i) = \eta(x,i)\beta_i \frac{w_x}{\sum_{z \in D} w_z}.$$
(11)

Importantly, categorisations from the other stimulus, y, also contribute to the overall probability of selecting response category i, but as stimulus y will have lower attention weight its processing rate (i.e., v(y, i)) will be lower. The probability of choosing category i when presented with stimuli x and y becomes

$$p(i|x) = \frac{\eta(x,i)\beta_i \frac{w_x}{\sum_{z \in D} w_z} + \eta(y,i)\beta_i \frac{w_y}{\sum_{z \in D} w_z}}{\sum_{z \in D} \sum_{j \in R} \eta(z,j)\beta_j \frac{w_z}{\sum_{z \in D} w_z}}.$$
(12)

1691 **Response selection.** Response selection in ECTVA is modelled as a random walk 1692 process. There is one counter for each possible response category; at each step of the 1693 random walk, the probability of incrementing the counter representing response category i1694 is provided by Equation 12. The random walk is terminated when the number of counts of 1695 one response category is greater than the counter with the next-largest number of counts by

a certain criterion set by the parameter K; at this stage this response category is said to be selected and response execution occurs. As the response categories are arbitrarily mapped onto experimental response keys, the probability of selecting response key a, $p(R_a)$, is the sum of the response category probabilities in the set M of response categories mapped to that key,

$$p(R_a|x) = \sum_{j \in M} p(j|x).$$
(13)

The time taken to complete each step of the random walk process is the reciprocal of the summed processing rates for all stimuli in the display plus a fixed cost, α , representing the time taken to increment the counter and make the comparison to determine whether response selection can terminate. Thus the time per step is given by

$$T_{Step} = \frac{1}{\sum_{z \in D} \sum_{j \in R} v(z, j)} + \alpha, \tag{14}$$

and the number of steps required in the random walk process—from the exemplar-based random walk model of Nosofsky and Palmeri (1997)—is given by

$$N_{Step}(x,y) = \frac{1}{p(R_a|x,y) - p(R_b|x,y)} [\theta_1(2K) - \theta_2(K)]$$
(15)

1707 where

$$\theta_1 = \frac{[p(R_a|x,y)/p(R_b|x,y)]^{2K} + 1}{[p(R_a|x,y)/p(R_b|x,y)]^{2K} - 1}$$

1708 and

$$\theta_2 = \frac{[p(R_a|x,y)/p(R_b|x,y)]^K + 1}{[p(R_a|x,y)/p(R_b|x,y)]^K - 1}$$

Assuming the correct response key is Key a, the probability of a correct response, $p(R_c)$, is given by

$$p(R_C|x,y) = \frac{1 - [p(R_b|x,y)/p(R_a|x,y)]^K}{1 - [p(R_b|x,y)/p(R_a|x,y)]^2 K}.$$
(16)

Parameter reconfiguration. In ECTVA, the time taken for reconfiguration of 1711 each parameter is assumed to be distributed exponentially with a rate parameter, μ , of 1712 1. Reconfiguration of each parameter in ECTVA is assumed to be independent and can 1713 therefore be executed in parallel. As such, the total reconfiguration time is determined 1714 by the slowest parameter to be reconfigured, and the mean total reconfiguration time gets 1715 longer as the number of parameters to be reconfigured increases. Reconfiguration time was 1716 simulated by taking the maximum value of n samples from an exponential distribution with 1717 rate parameter 1, where n is the number of parameters to be reconfigured and is determined 1718 by the current trial's response-set sequence and stimulus-set sequence. 1719

The model's response time for the current trial is therefore the reconfiguration time plus number of steps in the random walk multiplied by the time taken per step:

$$RT = T_{\text{Reconfiguration}} + (N_{\text{Step}} \cdot T_{\text{Step}}).$$
(17)

1722 **Response compatibility.** A worked out example of how compatibility effects arise 1723 in ECTVA is given in Table 7.

1724 Simulation 1. Simulating the Basic Experiment Design

In Simulation 1 I explored ECTVA's predictions when adapted to the current experimental design. I used the same parameter values as reported in Logan and Gordon (2001) for all Simulations in this appendix, which can be seen in Table 8. All predictions therefore are not a result of parameter-fitting routines optimising the model predictions to fit observed data, but rather are a consequence of the architecture of the model.

I simulated 100 repetitions of the full experiment design; each repetition therefore had 256 trials comprising every possible combination of (a) all stimulus identities avoiding repetitions; (b) current response-set; and (c) current stimulus-set. Trials were organised post-hoc into response-set and stimulus-set repetitions and switches, and were categorised as compatible or incompatible based on the response-compatibility to the two selected stimuli on each trial.

1736 The results are shown in Figure 5 in the main body of the paper.

1737 Simulation 2. Parallel vs. serial reconfiguration

A central assumption in ECTVA is that reconfiguration of control parameters occurs 1738 in parallel, and total reconfiguration time is determined by the slowest parameter to be 1739 reconfigured. Therefore, in the current research the assumption is that when both response-1740 set and stimulus-set switches, the four bias parameters $(\beta_{odd}, \beta_{even}, \beta_{low}, \text{ and } \beta_{high})$ and the 1741 two priority parameters (π_{orange} and π_{purple}) are reconfigured concurrently. As described in 1742 the main body of the text, it is this assumption of independent and parallel reconfiguration 1743 that explains the under-additive interaction between response-set sequence and stimulus-set 1744 sequence, and the mean expected reconfiguration time is given by 1745

Table 7 $\,$

Effects of compatibility on response evidence. For congruent calculations, x = 2, y = 4. For incongruent calculations, x = 2, y = 7.

ECTVA Component	Parameter / Term	Compatible	Incompatible
Sensory Evidence	$\eta_{associated}$	10	10
	$\eta_{unassociated}$	1	1
Response Set	β_{odd}	1	1
	β_{even}	1	1
	β_{low}	0.1	0.1
	eta_{high}	0.1	0.1
Stimulus Set	π	1	1
Stillulus Set	π_{orange}	0.1	0.1
	^{<i>n</i>} purple	0.1	0.1
Equation 1	v(x, odd)	0.902	0.902
-	v(y, odd)	0.098	0.982
	v(x, even)	9.02	9.02
	v(y, even)	0.982	0.098
	v(x, low)	0.902	0.902
	v(y, low)	0.098	0.01
	v(x, high)	0.09	0.09
	v(y, high)	0.01	0.098
	v_{total}	12.1	12.1
Emerican 9		0.002	0.156
Equation 3	p_{odd}	0.083	0.150
	p_{even}	0.820	0.755
	p_{low}	0.083	0.075
	p_{high}	0.008	0.016
Equation 4	p_a	0.166	0.231
*	p_b	0.834	0.769

Table 8

ECTVA parameter values from Logan and Gordon (2001) used in all Simulations unless otherwise stated.

Parameter	Description	Value
$\eta_{\rm rs-associated}$	Evidence for response sets associated with presented stimuli	10.00
$\eta_{\rm rs-unassociated}$	Evidence for response sets unassociated with presented stimuli	1.00
$\beta_{\rm associated}$	Attentional bias to response sets associated with presented cue	1.00
$\beta_{\mathrm{unassociated}}$	Attentional bias to response sets unassociated with presented cue	0.10
$\eta_{\rm ss-associated}$	Evidence for stimulus sets associated with presented stimuli	10.00
$\eta_{\rm ss-unassociated}$	Evidence for stimulus sets unassociated with presented stimuli	1.00
$\pi_{\rm associated}$	Attentional priority for stimulus sets associated with presented stimuli	1.00
$\pi_{\mathrm{unassociated}}$	Attentional priority for stimulus sets unassociated with presented stimuli	0.10
v	Rate parameter for parameter reconfiguration finishing time	1.00
α	Constant time cost for incrementing response counters	0.30
K	Response thereshold for response selection	3.00

$$T_{parallel} = \frac{1}{v} \sum_{i=1}^{n} \frac{1}{i}.$$
 (18)

However, an alternative could be that reconfiguration of the response-set parameters and stimulus-set parameters cannot occur in parallel; that is, stimulus-set reconfiguration must wait until response-set reconfiguration is complete (or vice-versa). In this case, total reconfiguration time is the sum of the slowest reconfiguration time from all n response-set parameters and the slowest reconfiguration time from all m stimulus-set parameters:

$$T_{serial} = \frac{1}{v} \left(\sum_{i=1}^{n} \frac{1}{i} + \sum_{j=1}^{m} \frac{1}{j} \right).$$
(19)

This leads to an additive effect of response-set and stimulus-set sequence. To demon-1751 strate this I conducted two simulations of ECTVA using the design in the current study. 1752 In one simulation parameter reconfiguration could occur in parallel (Equation 18) and re-1753 configuration time was determined as the slowest parameter reconfiguration time simulated 1754 by taking the maximum value from n draws from an exponential distribution with rate 1755 parameter 1. In the other simulation, I enforced a type of serial reconfiguration in that one 1756 set of control parameter (i.e., β parameters or π parameters) were reconfigured first before 1757 starting on reconfiguration of the other set of control parameters; within each set reconfig-1758 uration could occur in parallel (i.e., all four β parameters could be reconfigured in parallel 1759 with reconfiguration time determined by the slowest parameter; Equation 19). This was 1760 simulated by taking the maximum value from n draws from an exponential distribution with 1761

rate parameter 1 (where n equals the number of response-set parameters to reconfigure) and adding this to the maximum value from m draws from an exponential distribution with rate parameter 1 (where m equals the number of stimulus-set parameters to reconfigure).

In each simulation I simulated 100 repetitions of the experiment design. The results of these simulations are shown in Figure 15. Both parallel and serial reconfiguration models each predict a response-set and stimulus-set switch cost in response times, but the underadditive interaction is only present in the parallel model. Note that this effect is absent in predictions of proportion error, where only a main effect of response compatibility is predicted.



Figure 15. Predicted mean response time (in arbitrary units) and proportion error from ECTVA simulations of the current experimental design (with parallel and serial reconfiguration of control parameters).

1771 Simulation 3. Effects of Extended Cue–Stimulus Intervals

Simulation 3 used the same parameter values as before and assumed parallel reconfiguration of response-set and stimulus-set parameters. Total number of parameters to reconfigure were again determined by the current response-set and stimulus-set sequence.

To model the effects of cue-stimulus interval (CSI), I assumed that reconfiguration time could be absorbed into the CSI; as such, if the CSI is longer than reconfiguration time, parameter reconfiguration does not contribute to the response time. That is,

$$T = \max(0, T_{\text{Reconfiguration}} - \text{CSI}).$$
(20)

As time is in arbitrary units in the model, I conducted four simulations with CSI at 0 units, 1 unit, 2 units, and 4 units. The results are shown in Figure 16: As CSI increases, the effect of response-set sequence, stimulus-set sequence, plus their interaction diminishes for response times. CSI has no effect on error rates as reconfiguration occurs before response selection commences.

1783 Simulation 4. Effects of Cueing Order

I modelled the effects of cueing order in a similar way to that of CSI, but advanced reconfiguration was only possible for the component of the task-set that was cued first: If response-set was cued first, then reconfiguration of the β parameters could be absorbed into the CSI. I set the time between the onset of the first cue and the onset of the second cue plus the stimuli at a value of 2 arbitrary model time units.

The simulation was run twice, once with response-set being cued first and the other 1789 with stimulus-set being cued first. The results of the simulation are shown in Figure 17. 1790 In the response time data, a clear pattern emerged: If response-set was cued first, the 1791 response-set switch cost was drastically reduced but a stimulus-set switch cost remained; 1792 if stimulus-set was cued first, the stimulus-set switch cost was drastically reduced but a 1793 response-set switch cost remained. That is, advanced preparation of a component of the 1794 task-set led to a reduction in the switch cost associated with that component. Cueing order 1795 had no effects on the error rates. 1796

¹⁷⁹⁷ Simulation 5. Effect of Extended Response–Cue Intervals

ECTVA by default has only one source of carryover effect from the previous trial, which is the value of the response counters. Recall that response selection proceeds via a random walk where each response category has a counter value; these counters are incremented on each step of the random walk, and the counter that is incremented is determined probabilistically by Equation 12 until the number of counts in one response counter is greater than that in the counter with the next-largest number of counts by criterion K. At this stage the response category is selected. In the simulations reported by Logan and



Figure 16. Predicted response times (upper panel, in arbitrary units) and proportion error (lower panel) from ECTVA as the cue–stimulus interval is increased (with values of 0, 1, 2, and 4 units of arbitrary time).



Figure 17. Predicted response times (upper panel, in arbitrary units) and proportion error (lower panel) from ECTVA as a function of whether response-set is cued first ("rs_first") or stimulus-set is cued first ("ss_first").

Gordon (2001), the values in the response counters at the end of response selection were 1805 inhibited by reducing their counter values by 90%; therefore 10% of the "activation" of the 1806 response categories carry-over into the next trial. This carry-over could prime performance 1807 if the response-category required on the next trial matches that of the previous trial, as 1808 the correct response counter will have a head-start in the random walk process; this could 1809 also lead to a cost if the response-category required on the next trial is different to that 1810 of the previous trial as the new response category will take longer to surpass the criterion 1811 K due to persisting activation of an incorrect response category. Note though that this 1812 response inhibition mechanism was not implemented in the current work because I used the 1813 analytical approach to calculating response selection time and accuracy (Equations 14-16). 1814

In exploring the effects of RCI within ECTVA, I operationalised the concept of residual activation differently via the strength of the previous task set, represented by β and π values. I allowed a proportion of these values to carry-over into the next trial by adding this proportion to the parameter values for the current trial, representing carry-over of task-set. Carry-over of the response-set parameters will benefit performance on response-set repetition trials (because the set-relevant β values will receive an additive boost) and will inter-

fere on response-set switch trials (because the set-*irrelevant* β values will receive a boost). Similarly, carry-over of stimulus-set parameters will benefit performance on stimulus-set repetition trials (because the set-relevant π values will receive an additive boost) and will interfere on stimulus-set switch trials (because the set-*irrelevant* π values will receive a boost).

1826

We can assume that this residual activation decays over time, for example:

$$y(t) = y_0 \cdot \exp^{-\tau * t} \tag{21}$$

where y(t) is the activation of set-parameter y after the RCI time t, y_0 is the initial activation 1827 of y, and τ is a decay constant (set here at 1 for both response-set and stimulus-set). 1828 To keep a version of ECTVA's response inhibition concept, response-set parameter values 1829 were reduced before the remaining activation began to decay as a function of RCI. In the 1830 original ECTVA, response counters were inhibited by 90%, but in the simulation reported 1831 here I reduced this to 70% so that there was sufficient residual activation for the RCI 1832 to act upon. To constrain unlimited growth of the parameters they were clipped to the 1833 maximum parameter value shown in Table 8, so RCI effects are operationalised as the 1834 absolute difference between β and π parameter values to the relevant and irrelevant response-1835 sets and stimulus-sets, respectively. 1836

The predictions across various RCIs are shown in Figure 18. For the response times, 1837 there is a small reduction of response-set switch cost (costs of 1.60, 1.51, 1.48, and 1.46 1838 arbitrary units for RCIs of 0.25, 1, 2, and 4) a small reduction in stimulus-set switch cost 1839 with increasing RCI, although the reduction in stimulus-set switch cost is only evident on 1840 response-incompatible trials (costs of 1.09, 1.02, 0.987, and 0.969; costs were 0.908 in all 1841 RCIs on compatible trials). In addition, at short RCIs there is an under-additive interaction 1842 between response-set and stimulus-set sequence in the response times but this does not 1843 appear to reduce at longer RCIs. 1844

In the error data, ECTVA predicts a response-set and stimulus-set switch cost at the shortest RCI. Note that this is the first observation of switch cost in ECTVA error rates. There is a reduction of response-set switch cost that appears to be larger for incompatible trials (4.28%, 1.91%, 0.67%, 0.10%) than on compatible trials (1.77%, 0.60%, 0.17%, 0.02%). Similarly, the stimulus-set switch cost also reduced with increasing RCI, but was exclusively present only on incompatible trials (1.58%, 0.83%, 0.33%, 0.06%) and not on compatible trials (-0.05%, -0.01%, 0.00%, 0.00%).



Figure 18. Predicted response times (upper panel, in arbitrary units) and proportion error (lower panel) from ECTVA as the response-cue interval is increased (with values of 0.25, 1, 2, and 4 units of arbitrary time).

1852

Appendix B — Parameter Space Simulations

To demonstrate that the under-additive interaction between response-set and stimulus-set sequence predicted by ECTVA is a feature of the model architecture rather than the specific parameter values used in the simulations, in the current Section I conduct a series of simulations that explore the parameter space of the model. The under-additive interaction is exclusively due to the non-linear increase in reconfiguration time with increasing number of parameters to be reconfigured, as formalised in Equation 5, and as such should arise regardless of the actual parameter values.

To demonstrate this, I selectively varied one key parameter at a time across 12 values 1860 whilst retaining all other parameters at their default values (see Table 9) and plotted the 1861 model's response time predictions for each. Specifically, I varied the key control parameters 1862 $(\beta_{associated} \text{ and } \pi_{associated})$ as well as the key sensory evidence parameters $(\eta_{rs-associated})$ 1863 and $\eta_{ss-associated}$; I also varied the parameter reconfiguration time, which was the rate 1864 parameter of the exponential distribution (v). Note that I did not vary "unassociated" 1865 parameters. For example, I only varied $\beta_{associated}$, which represents the attentional bias 1866 to response-sets associated with the presented cue on the current trial (e.g., response-sets 1867 "odd" and "even" if the presented cue is "Odd/Even"); I did not vary $\beta_{unassociated}$, which 1868 represents the attentional bias to response-sets unassociated with the presented cue (e.g., 1869 response-sets "low" and "high" if the presented cue is "Odd/Even"; see Table 8 in Appendix 1870 A for definitions of all parameters). What matters with the associated–unassociated pair-1871 ings is the relative balance of weight assigned to each, and—by the very definition of the 1872 parameter—associated parameters must always be set to larger values of unassociated. So 1873 the parameter values varied in the current simulations varied the relative weight given to 1874 associated parameters in relation to unassociated parameters. 1875

For each parameter that was the target of exploration in Table 9, I established 12 equally spaced parameter values within the range specified in the table. Then, for each parameter, I simulated 100 repetitions of the full experimental design (see Appendix A) using the target parameter value together with the default settings for the other parameters. I then plotted the mean simulated response time for response-set repetitions and switches and stimulus-set repetitions and switches.

The results of these simulations can be seen in Figure 19 ($\beta_{associated}$), Figure 20 ($\pi_{associated}$), Figure 21 ($\eta_{ss-associated}$), Figure 22 ($\eta_{rs-associated}$), and Figure 23 (v). Note in all Figures and for all parameter settings there is an under-additive interaction between response-set sequence and stimulus-set sequence.



Figure 19. Predicted mean response time (in arbitrary units) for response-set and stimulusset sequences (i.e., repetition vs., switch) from ECTVA simulations where the β

set sequences (i.e., repetition vs. switch) from ECTVA simulations where the $\beta_{associated}$ parameter was varied between 0.2 and 1.0.



Stimulus Set - repetition - switch

Figure 20. Predicted mean response time (in arbitrary units) for response-set and stimulusset sequences (i.e., repetition vs. switch) from ECTVA simulations where the $\pi_{associated}$ parameter was varied between 0.2 and 1.0.


Figure 21. Predicted mean response time (in arbitrary units) for response-set and stimulusset sequences (i.e., repetition vs. switch) from ECTVA simulations where the $\eta_{ss-associated}$ parameter was varied between 3.0 and 10.0.



Stimulus Set - repetition - switch

Figure 22. Predicted mean response time (in arbitrary units) for response-set and stimulusset sequences (i.e., repetition vs. switch) from ECTVA simulations where the $\eta_{rs-associated}$ parameter was varied between 3.0 and 10.0.



Stimulus Set - repetition - switch

Figure 23. Predicted mean response time (in arbitrary units) for response-set and stimulusset sequences (i.e., repetition vs. switch) from ECTVA simulations where the rate parameter v of reconfiguration time was varied between 0.5 and 3.0.

Table 9		
ECTVA default	paramete	r values from
Logan and Gor	don (2001	1) as well as
the parameter v	alue range	$es \ explored \ in$
the parameter s	pace simu	lations.
Parameter	Default	Range
$\eta_{ m rs-associated}$	10.00	3.00-10.00
$\eta_{\rm rs-unassociated}$	1.00	
$\beta_{\rm associated}$	1.00	0.20 - 1.00
$\beta_{\mathrm{unassociated}}$	0.10	

10.00

1.00

3.00 - 10.00

Table 0

 $\eta_{\text{ss-associated}}$

 $\eta_{\rm ss-unassociated}$

 $\begin{array}{ccccc} \pi_{\rm associated} & 1.00 & 0.20 - 1.00 \\ \pi_{\rm unassociated} & 0.10 & - \\ v & 1.00 & 0.50 - 3.00 \\ \alpha & 0.30 & - \\ K & 3.00 & - \end{array}$

1886

Appendix C — Exploratory Analysis of Preparation Effects

To examine further the effects of preparation on switch cost, I examined the com-1887 ponent switch costs across the entire response time distribution in Experiments 3b and 1888 3c. To do so, I constructed delta plots which visualise component switch costs across vari-1889 ous percentiles of the response-time distribution, from fastest responses to the slowest. To 1890 construct the delta plots, I collapsed the data across response-compatibility and for each 1891 participant I calculated their response times at various percentiles (0.1, 0.3, 0.5, 0.7, and1892 (0.9) for each level of response-set and stimulus-set sequence and calculated the component 1893 switch costs at each percentile. This was done separately for the different levels of set-order. 1894 When constructing the delta plots for response-set switch costs, I calculated the switch cost 1895 as the difference between trials with a single-component switch of response-set (i.e., trials 1896 with a response-set switch but a repetition of stimulus-set) and trials with a repetition of 1897 both sets. For the stimulus-set delta plots, I calculated the switch cost as the difference 1898 between trials with a single-component switch of stimulus-set (i.e., trials with a stimulus-set 1899 switch but a repetition of response-set) and trials with a repetition of both sets. This was 1900 done separately for the levels of set-order (i.e., response-set cued first and stimulus-set cued 1901 first). 1902

For response-set costs (see the first column of Figure 24A), response-set switch costs were smaller when the response-set was cued first for the first percentile, but at later percentiles the switch cost was larger when response-set was cued first. For stimulus-set switch costs (second column of Figure 24A), stimulus-set switch costs were smaller when the stimulus-set was cued first for the first percentiles, but at later percentiles the switch cost was larger when stimulus-set was cued first. For completeness, I also constructed delta plots for the switch costs when both components switched (final column of Figure 24A)) which showed switch costs were consistently smaller across percentiles when the stimulus-set was cued first.

To examine these observations statistically, three separate Bayesian linear models 1912 were conducted (one for each type of switch cost) predicting the switch cost from set order 1913 and percentile. For these models, the data from Experiment 3b and 3c were combined.¹⁹ For 1914 response-set costs, the model showed an effect of set-order (b = 38.12, 95% credible interval 1915 [CI] 21.18, 67.11), and effect of percentile (b = 213.84 [173.58, 253.53]), plus an interaction 1916 (b = -119.41 [-176.00, -62.02]). To examine this interaction, contrasts were taken from the 1917 model by comparing the posterior estimates for the difference in switch cost between each 1918 level of set-order (calculated by subtracting the cost for stimulus-set first from the cost for 1919 response-set first conditions). These contrasts can be seen in the first column of Figure 24B 1920 together with their 95% CIs. As can be seen, the contrasts showed that the switch cost 1921 was reliably smaller when the response-set was cued first at the shortest percentile, but this 1922 pattern was reversed at longer percentiles. 1923

For stimulus-set costs the model showed an effect of set-order (b = -37.10, [-71.67, -2.94]), an effect of percentile (b = 54.19 [11.55, 96.67]), plus an interaction (b = 61.71 [2.23, 120.80]). The contrasts (see second column of Figure 24B) showed that the switch cost was reliably smaller when the stimulus-set was cued first at the shortest percentile, but this pattern was reversed at longer percentiles.

For the costs when both components switched the model showed no effect of set-order (b = -5.17, [-40.88, 29.61]), there was an effect of percentile (b = 189.12 [145.40, 231.66]), but no interaction (b = -37.56 [-97.64, 24.40]). The contrasts (see final column of Figure 24B) showed that the switch cost was smaller when the stimulus-set was cued first at all percentiles.

1934 Experiment 3d

The exploratory distributional analysis has provided some support for ECTVA's pre-1935 dictions of reduced component-switch cost when that component is sufficiently prepared, 1936 but this conclusion relies on the assumption of the failure to engage hypothesis that re-1937 sponses in the shortest RT percentiles are those in which participants have engaged in 1938 advanced preparation (DeJong, 2000; see also Grange & Houghton, 2011). In an attempt 1939 to behaviourally encourage participants to engage in advanced preparation of task-set com-1940 ponents, I conducted an additional experiment which modified Experiment 3c's design, but 1941 the cue presented first was only presented for 300 ms before disappearing. Reducing the 1942 cue presentation time should encourage advanced preparation (see Verbruggen et al., 2007; 1943 but see Schneider, 2016). 1944

1945 Method.

Participants. 127 new participants were recruited from the same pool as the main
 experiments. I removed participants who failed to maintain a session-wise accuracy above

¹⁹I attempted to implement a Bayesian linear model with experiment as a random factor but experienced convergence issues that could not be resolved.



Figure 24. A. Delta plots showing mean response-set switch costs (first column), stimulusset switch costs (second column) and switch costs when both components switch (final column) in Experiment 3b (upper row) and Experiment 3c (lower row). In each column, the separate lines show the set-order conditions when the response-set is cued first ("rs_first"), and when the stimulus-set is cued first ("ss_first")). Error bars denote one standard error around the mean switch costs. B. Contrasts from the Bayesian linear models comparing the posterior estimates for the difference in switch cost between each level of set-order for response-set switch costs (first panel), stimulus-set switch costs (second panel), and switch costs when both components switch (third panel). Points show the mean posterior estimates for each contrast, and error bars denote 95% credible intervals. Points above the horizontal dashed line show switch costs that are smaller when the stimulus-set is cued first, and below the line show switch costs that are smaller when the response-set is cued first.

1963

85%, which removed 29 participants. Note that this is a more liberal accuracy criterion than
the main experiments which used 90%; however, this stricter criterion led to a removal of 46
participants, indicating that participants found this experiment more challenging (possibly
due to the shorter cue-presentation times).

Materials & Procedure. The experiment was identical to Experiment 3c with the 1952 exception of the following. In the first stage of cue presentation, only information about 1953 either the response-set or the stimulus-set (dependent upon the current condition) was pre-1954 sented (as in Experiment 3b), but here it was only presented for 300 ms. After this time, 1955 the cue became a white placeholder (e.g., "low/high" changed to (-/-) which provided 1956 no information about either response-set or stimulus-set. This white placeholder was pre-1957 sented for a further 600 ms before the placeholder changed to cue the other component not 1958 previously cued (i.e., if the firs cue provided information about the response-set, the final 1959 cue only provided information about the stimulus-set). The trial stimuli appeared at the 1960 same time as this second cue. Importantly, as in Experiment 3c, the task-set component 1961 cued at the beginning of the trial was not present during stimulus onset. 1962

Results. The mean response times and proportion error can be seen in Figure 25.

Response time analysis. The model comparisons can be seen in Table 10. This 1964 showed extreme evidence in favour of Order: RTs were shorter overall when the response-1965 set cue was presented first (M = 1497 ms, SE = 17) than when the stimulus-set cue was 1966 presented first (M = 1550, SE = 16). As in previous experiments, there was extreme 1967 evidence in favour of both Response-Set Sequence and Stimulus-Set Sequence: Response-1968 set repetitions (M = 1465 ms, SE = 16) were responded to faster than response-set switches 1969 (M = 1582, SE = 17), and stimulus-set repetitions (M = 1479, SE = 16) were responded 1970 to faster than stimulus-set switches (17). There was no evidence that stimulus-set sequence 1971 interacted with Set Order. However, the was strong evidence that response-set sequence 1972 interacted with Set Order: The response-set switch cost was 84 ms when the stimulus-set 1973 cue was presented first, but it was *larger* at 150 ms when the response-set cue was presented 1974 first. 1975

There was extreme evidence for Response Compatibility: Response times were shorter for response-compatible trials (M = 1485, SE = 16) than for response-incompatible trials (M = 1562, SE = 16). However, there was moderate evidence that Response Compatibility was not involved in any interactions except for with Set Order (which was moderate): The compatibility effect was 104 ms when the response-set was cued first, and 48 ms when the stimulus-set was cued first.

There again was moderate evidence for an interaction between Response-Set Sequence and Stimulus-Set Sequence. There was moderate evidence that this interaction was not further modulated by Set Order or Compatibility, and there was moderate evidence against the four-way interaction.

Testing the predictions of task-set structure, the results (collapsing across compatibility and set-order) showed no evidence for a difference in response time between a singlecomponent switch of response-set and a single-component switch of stimulus-set ($BF_{10} =$ 2.15), extreme evidence for a difference between a single-component switch of response-set and a switch of both components ($BF_{10} = 5.93 \times 10^5$), and extreme evidence for a difference



Figure 25. Behavioural data from Experiment 3d as a function of whether response-set was cued first ("rs_first") or stimulus-set was cued first ("ss_first"). Error bars show one standard error around the mean.

Table 10

Model comparison results for the behavioural data for Experiment 3d. The Bayes factors (BF) show comparison of the full factorial model (including all main effects and all interactions) against models with particular predictors omitted (i.e., $BF = \frac{omitted - model}{full - model}$). BF values below 1 indicate evidence in favour inclusion of that predictor.

Omission	BF(RT)	BF (Error)
Order (O) x Compatibility (C) x R-Set Seq. (RS) x S-Set Seq. (SS)	3.65	3.51
$C \ge RS \ge SS$	4.20	5.92
$O \ge C \ge SS$	3.13	0.03
$O \ge C \ge RS$	5.71	6.36
$O \ge RS \ge SS$	4.45	4.34
$C \ge SS$	6.00	$1.43 imes 10^{-4}$
$C \ge RS$	5.94	11.95
$C \ge O$	0.26	0.02
$RS \ge SS$	0.24	7.00
O x SS	0.53	9.52
O x RS	0.09	4.22
С	3.78×10^{-10}	2.00×10^{-14}
SS	3.72×10^{-14}	4.65×10^{-9}
RS	1.78×10^{-23}	1.56×10^{-12}
0	8.98×10^{-5}	1.98×10^{-3}

Note. Seq. = sequence.

between a single-component switch of stimulus-set and a switch of both components $(BF_{10})_{1992} = 4.74 \times 10^8$.

Error analysis. The results of the error analysis largely mirrored those for the 1993 response times, with the following exceptions. There was no evidence for an interaction 1994 between Response-Set Sequence and Set-Order, and no evidence for an interaction between 1995 Response-Set Sequence and Stimulus-Set Sequence. As in previous experiments, there was 1996 an interaction between Stimulus-Set Sequence and Response Compatibility which was fur-1997 ther moderated by Set Order: When response-set was cued first, the stimulus-set switch 1998 cost for compatible trials was 1.26%, and was 1.96% for incompatible trials; when stimulus-1999 set was cued first, there was a stimulus-set switch *benefit* for compatible trials of 0.24%, 2000 but there was a stimulus-set switch *cost* of 4.06% for incompatible trials. 2001

Testing the predictions of task-set structure, the results (focussing on incompatible trials and collapsing across set order) showed no evidence for a difference in errors between a single-component switch of response-set and a single-component switch of stimulus-set $(BF_{10} = 1.86)$, extreme evidence for a difference between a single-component switch of response-set and a switch of both components $(BF_{10} = 4.62 \times 10^5)$, and strong evidence for a difference between a single-component switch of stimulus-set alone and a switch of both components $(BF_{10} = 30.25)$. **Distributional analysis.** As with the exploratory analysis of Experiments 3b and 3c, I constructed delta plots of the relevant component-switch costs. These can be seen in the upper row of Figure 26.

For response-set costs the model showed no effect of set-order (b = 16.98, [-37.53, 2012 (1.04]), there was an effect of percentile (b = 164.42 [97.88, 231.77]), and there was an 2013 interaction $(b = -150.67 \ [-245.60, -54.56])$. The contrasts (see second column of Figure 2014 26A) showed that the switch cost equivalent between when response-set was cued first and 2015 when stimulus-set was cued first at the shortest percentile, but was reliably smaller when 2016 stimulus-set was cued first at longer percentiles. For stimulus-set costs the model showed a 2017 small effect of set-order (b = -45.62, [-98.69, 9.30]), no effect of percentile (b = -21.15, [-86.62, -20.25, -20.25])2018 (45.61), but there was an interaction (b = 221.04 [126.16, 313.21]). The contrasts (see second 2019 column of Figure 24B) showed that the switch cost equivalent between when stimulus-set 2020 was cued first and when response-set was cued first at the shortest two percentiles, but was 2021 reliably smaller when response-set was cued first at longer percentiles. For the costs when 2022 both components switched the model showed no effect of set-order (b = -10.57, [-68.80, 2023 (47.81), there was an effect of percentile (b = 190.68 [118.34, 263.35]), but there was also 2024 no interaction $(b = -6.63 \ [-108.16, 94.97])$. The contrasts (see final column of Figure 24B) 2025 showed that the switch cost was equivalent at all percentiles. 2026

Discussion. Despite providing more incentive to engage in advanced preparation of task-set components, this new experiment still failed to find a specific effect of component preparation on component-switch costs. In addition, the RT distributional analysis did not clearly replicate the finding from the exploratory analysis of Experiments 3b and 3c.



Figure 26. A. Delta plots showing mean response-set switch costs (first column), stimulusset switch costs (second column) and switch costs when both components switch (final column) from Experiment 3d. In each column, the separate lines show the set-order conditions when the response-set is cued first ("rs_first"), when the stimulus-set is cued first ("ss_first")). Error bars denote one standard error around the mean switch costs. **B.** Contrasts from the Bayesian linear models comparing the posterior estimates for the difference in switch cost between each level of set-order for response-set switch costs (first panel), stimulus-set switch costs (second panel), and switch costs when both components switch (third panel). Points show the mean posterior estimates for each contrast, and error bars denote 95% credible intervals. Points above the horizontal dashed line show switch costs that are smaller when the stimulus-set is cued first.

SET SWITCHING

2031

Appendix D — Bayesian Meta Analyses

In this appendix I describe the Bayesian meta analyses reported in the main paper 2032 which adopts the implementation suggested by Vuorre (2016). For each experiment (and 2033 each condition in each experiment, where relevant; e.g., "short CSI" and "long CSI" of 2034 Experiment 3a were analysed separately), three contrasts were calculated from the response 2035 time data, reflecting: (1) a single-component switch of response-set vs. a single-component 2036 switch of stimulus-set; (2) a single-component switch of response-set vs. a switch of both 2037 response-set and stimulus-set, and (3) a single-component switch of stimulus-set vs. a switch 2038 of both components. For each contrast and each experiment/condition, the mean and 2039 standard error of the contrast was calculated. 2040

One Bayesian meta-analysis was then conducted per contrast. For each meta-analysis, the mean of the contrast of each experiment/condition is modelled as

$$y_i \sim \text{Normal}(\theta_i, \sigma_i),$$
 (22)

where y_i is the mean estimate for experiment/condition *i*, which is modelled as a draw from a normal distribution centered on the true effect size of that experiment/condition, θ_i , and standard deviation σ_i . Here, σ_i is taken directly from the standard error of that experiment/study's contrast. The model also assumes that each study's true effect size, θ_i , is itself a draw from a population of studies centered on the true effect size in the population:

$$\theta_i \sim \operatorname{Normal}(\mu, \tau),$$
(23)

where μ is the true effect size of the contrast in the population, and τ is the standard deviation of this effect size in the population.

Each model was fit using **brms** (Bürkner, 2017, 2018) together with regularising priors for μ and τ ; parameter estimation proceeded by taking 5,000 samples from the posterior distribution from each of four chains (with 2,000 samples being taken as warm up per chain). Visual inspection of the chains showed good convergence, and all \bar{R} values were very close to 1.

2055

Appendix E — Frequentist Analysis of Experiments

2056 Experiment 1a

Mean Response Time Analysis. There was a significant main effect of Response-Set Sequence, F(1, 42) = 63.91, p < .001, $\eta_g^2 = .081$, Stimulus-Set Sequence, F(1, 42) = 168.78, p < .001, $\eta_g^2 = .093$, and Response-Compatibility, F(1, 42) = 4.58, p = .038, $\eta_g^2 = .003$. There was a significant two-way interaction between Response-Set Sequence and Stimulus-Set Sequence, F(1, 42) = 22.49, p < .001, $\eta_g^2 = .009$. Response-Compatibility did not interact with either Response-Set Sequence, F(1, 42) = 0.19, p = .667, $\eta_g^2 < .001$, or Stimulus-Set Sequence, F(1, 42) = 0.17, p = .685, $\eta_g^2 < .001$. The three-way interaction was also not significant, F(1, 42) = 0.83, p = .368, $\eta_g^2 < .001$.

Error Analysis. There was a significant main effect of Response-Set Sequence, F(1, 2066 + 42) = 18.97, p < .001, $\eta_g^2 = .055$, Stimulus-Set Sequence, F(1, 42) = 4.84, p = .033, $\eta_g^2 = .010$, and Response-Compatibility, F(1, 42) = 5.86, p = .020, $\eta_g^2 = .020$. The two-way interaction between Response-Set Sequence and Stimulus-Set Sequence was not significant, F(1, 42) = 0.49, p = .487, $\eta_g^2 < .001$. Response-Compatibility did not interact with either Response-Set Sequence, F(1, 42) = 0.05, p = .825, $\eta_g^2 < .001$, or Stimulus-Set Sequence, F(1, 42) = 3.49, p = .069, $\eta_g^2 = .007$. The three-way interaction was also not significant, F(1, 42) = 0.00, p = .994, $\eta_g^2 < .001$.

2073 Experiment 1b

Mean Response Time Analysis. There was a significant main effect of Response-Set Sequence, F(1, 50) = 126.98, p < .001, $\eta_g^2 = .107$, Stimulus-Set Sequence, F(1, 50) = 154.07, p < .001, $\eta_g^2 = .101$, and Response-Compatibility, F(1, 50) = 15.63, p < .001, $\eta_g^2 = .008$. There was a significant two-way interaction between Response-Set Sequence and Stimulus-Set Sequence, F(1, 50) = 52.25, p < .001, $\eta_g^2 = .013$. Response-Compatibility did not interact with either Response-Set Sequence, F(1, 50) = 0.02, p = .895, $\eta_g^2 < .001$, or Stimulus-Set Sequence, F(1, 50) = 0.57, p = .453, $\eta_g^2 < .001$. The three-way interaction was also not significant, F(1, 50) = 0.42, p = .520, $\eta_g^2 < .001$.

Error Analysis. There was a significant main effect of Response-Set Sequence, F(1,2082 $(50) = 64.68, p < .001, \eta_q^2 = .113$, but no significant main effect of Stimulus-Set Sequence, 2083 $F(1, 50) = 1.27, p = .265, \eta_g^2 = .001, \text{ or } Response-Compatibility, F(1, 50) = 2.20, p =$ 2084 .144, $\eta_q^2 = .005$. The two-way interaction between Response-Set Sequence and Stimulus-Set 2085 Sequence was significant, F(1, 50) = 10.11, p = .003, $\eta_q^2 = .010$. Response-Compatibility 2086 interacted with Response-Set Sequence, F(1, 50) = 4.82, p = .033, $\eta_g^2 = .009$, but it did not 2087 interact with Stimulus-Set Sequence, F(1, 50) = 0.64, p = .428, $\eta_g^2 = .001$. The three-way 2088 interaction was also not significant, F(1, 50) = 0.01, p = .921, $\eta_q^2 < .001$. 2089

2090 Experiment 2

Mean Response Time Analysis. There was a significant main effect of Response-Set Sequence, F(1, 54) = 14.37, p < .001, $\eta_q^2 = .004$, Stimulus-Set Sequence, F(1, 54) =

Table 11		
Experiment	3a	RTs.

Effect	F(1, 111)	p	η_g^2
CSI	948.65	<.001	.252
R-Set Sequence (RS)	219.35	< .001	.064
S-Set Sequence (SS)	418.04	<.001	.071
Compatibility (C)	27.14	<.001	.005
$CSI \ge RS$	11.98	<.001	< .001
$CSI \ge SS$	10.10	.002	< .001
RS x SS	147.97	< .001	.012
CSI x C	0.58	.449	< .001
RS x C	0.97	.326	< .001
SS x C	0.30	.587	< .001
$CSI \ge RS \ge SS$	11.80	<.001	< .001
$CSI \ge RS \ge C$	0.16	.692	< .001
$CSI \ge SS \ge C$	0.49	.485	< .001
$RS \ge SS \ge C$	0.10	.753	<.001
CSI x RS x SS x C	0.07	.799	<.001

84.65, p < .001, $\eta_g^2 = .017$, and Response-Compatibility, F(1, 54) = 5.41, p = .024, $\eta_g^2 = .002$. The two-way interaction between Response-Set Sequence and Stimulus-Set Sequence was close to significance, F(1, 54) = 3.02, p = .088, $\eta_g^2 < .001$. Response-Compatibility did not interact with either Response-Set Sequence, F(1, 54) = 1.20, p = .277, $\eta_g^2 < .001$, or Stimulus-Set Sequence, F(1, 54) = 0.37, p = .547, $\eta_g^2 < .001$. The three-way interaction was also not significant, F(1, 54) = 0.04, p = .839, $\eta_g^2 < .001$.

Error Analysis. There was a significant main effect of Response-Set Sequence, F(1, 1)2099 54) = 8.77, p = .005, $\eta_g^2 = .023$, a significant main effect of *Stimulus-Set Sequence*, F(1, 54) = 27.70, p < .001, $\eta_g^2 = .035$, and a significant main effect of *Response-Compatibility*, 2100 2101 $F(1, 54) = 57.59, p < .001, \eta_q^2 = .090$. The two-way interaction between Response-Set 2102 Sequence and Stimulus-Set Sequence was not significant, F(1, 54) = 1.05, p = .311, η_a^2 2103 =.001. Response-Compatibility did not interact with Response-Set Sequence, F(1, 54) = 2104 $0.04, p = .849, \eta_g^2 < .001$, but it did interact with *Stimulus-Set Sequence*, F(1, 54) = 26.75, 2105 $p < .001, \eta_g^2 = .041$. The three-way interaction was not significant, F(1, 54) = 0.37, p = 0.372106 .544, $\eta_a^2 < .001$. 2107

2108 Experiment 3a

2109 Mean Response Time Analysis. See Table 11.

2110 Error Analysis. See Table 12.

Experiment 3a acc.			
Effect	F(1, 111)	p	η_g^2
CSI	5.14	.025	.002
R-Set Sequence (RS)	69.09	<.001	.046
S-Set Sequence (SS)	37.67	<.001	.018
Compatibility (C)	51.63	<.001	.023
CSI x RS	1.13	.290	<.001
CSI x SS	0.00	.975	<.001
RS x SS	4.53	.035	.002
CSI x C	1.28	.261	<.001
RS x C	0.01	.907	<.001
SS x C	16.91	< .001	.007
$CSI \ge RS \ge SS$	0.03	.853	<.001
$CSI \ge RS \ge C$	2.55	.113	.001
$CSI \ge SS \ge C$	1.52	.220	<.001
$RS \ge SS \ge C$	0.59	.444	<.001
$\mathrm{CSI} \ge \mathrm{RS} \ge \mathrm{SS} \ge \mathrm{C}$	0.79	.375	<.001

Table 12Experiment 3a ad

2111 Experiment 3b

- 2112 Mean Response Time Analysis. See Table 13.
- **Error Analysis.** See Table 14.
- 2114 Experiment 3c
- ²¹¹⁵ Mean Response Time Analysis. See Table 15.
- **Error Analysis.** See Table 16.

2117 Experiment 4

- ²¹¹⁸ Mean Response Time Analysis. See Table 17.
- 2119 Error Analysis. See Table 18.

Experiment 3b RTs.			
Effect	F(1, 102)	p	η_g^2
Order (O)	0.94	.336	.001
R-Set Sequence (RS)	124.90	<.001	.013
S-Set Sequence (SS)	113.08	<.001	.006
Compatibility (C)	34.59	<.001	.004
O x RS	0.78	.378	<.001
O x SS	0.77	.381	<.001
$RS \ge SS$	44.05	<.001	.002
O x C	7.19	.009	<.001
RS x C	0.11	.736	<.001
$SS \ge C$	0.71	.400	<.001
$O \ge RS \ge SS$	0.46	.500	<.001
$O \ge RS \ge C$	4.37	.039	<.001
O x SS x C	0.40	.528	<.001
$RS \ge SS \ge C$	1.01	.318	<.001
$O \ge RS \ge SS \ge C$	0.27	.604	<.001

Table 13 Experiment ²h BT

Table 14
Experiment 3b acc.
Effect

Effect	F(1, 102)	p	η_g^2
Order (O)	0.32	.575	<.001
R-Set Sequence (RS)	122.05	<.001	.060
S-Set Sequence (SS)	16.79	<.001	.006
Compatibility (C)	17.32	<.001	.014
O x RS	1.13	.290	<.001
O x SS	0.01	.943	<.001
RS x SS	5.68	.019	.002
O x C	0.52	.473	< .001
RS x C	4.83	.030	.002
SS x C	11.95	< .001	.005
O x RS x SS	1.16	.208	<.001
$O \ge RS \ge C$	0.01	.921	<.001
O x SS x C	1.42	.236	<.001
$RS \ge SS \ge C$	0.38	.537	<.001
$O \ge RS \ge SS \ge C$	2.27	.135	<.001

Experiment 3c RTs.			
Effect	F(1, 100)	p	η_g^2
Order (O)	14.78	<.001	.012
R-Set Sequence (RS)	188.55	<.001	.019
S-Set Sequence (SS)	182.59	<.001	.010
Compatibility (C)	48.85	<.001	.006
O x RS	4.27	.041	<.001
O x SS	0.98	.326	<.001
RS x SS	59.03	< .001	.002
O x C	42.67	< .001	.002
RS x C	4.93	.029	<.001
$SS \ge C$	0.02	.881	<.001
$O \ge RS \ge SS$	0.72	.398	<.001
$O \ge RS \ge C$	0.60	.440	<.001
O x SS x C	0.23	.629	<.001
$RS \ge SS \ge C$	0.50	.479	<.001
$O \ge RS \ge SS \ge C$	1.50	.223	< .001

Table 15Experiment 3c RTs

Table 16		
Experiment	3c	acc

Effect	F(1, 100)	p	
Order (O)	2.36	.128	
R-Set Sequence (RS)	106.05	< .001	
S-Set Sequence (SS)	27.88	<.001	
Compatibility (C)	47.24	<.001	
O x RS	0.46	.501	<
O x SS	1.97	.164	~
RS x SS	0.67	.413	<
O x C	1.61	.208	<
RS x C	0.18	.669	<
SS x C	16.76	<.001	
O x RS x SS	2.31	.132	<
O x RS x C	1.70	.195	
O x SS x C	6.85	.010	
$RS \ge SS \ge C$	0.59	.446	<
$O \ge RS \ge SS \ge C$	0.44	.507	<

Experiment 4 RIS.			
Effect	F(1, 101)	p	η_g^2
RCI	85.97	<.001	.017
R-Set Sequence (RS)	319.39	< .001	.049
S-Set Sequence (SS)	373.84	<.001	.045
Compatibility (C)	21.95	<.001	.004
RCI x RS	11.61	<.001	<.001
RCI x SS	79.75	<.001	.005
$RS \ge SS$	101.29	<.001	.006
RCI x C	1.42	.236	<.001
RS x C	2.28	.134	<.001
$SS \ge C$	2.66	.106	<.001
$RCI \ge RS \ge SS$	0.22	.638	<.001
RCI x RS x C	0.02	.890	<.001
RCI x SS x C	0.05	.827	<.001
$RS \ge SS \ge C$	0.08	.781	<.001
RCI x RS x SS x C	0.02	.901	< .001

Table 17Experiment 4 RTs

Table 18	
Experiment 4	Acc

Effect	F(1, 101)	p	η
RCI	2.49	.118	.00
R-Set Sequence (RS)	114.18	<.001	.07
S-Set Sequence (SS)	19.94	< .001	.00
Compatibility (C)	38.33	<.001	.01
RCI x RS	0.01	.920	<.(
RCI x SS	18.50	< .001	.00
RS x SS	1.88	.174	<.(
RCI x C	0.36	.550	<.(
RS x C	0.52	.473	<.(
SS x C	6.92	.010	.00
RCI x RS x SS	0.40	.528	<.0
RCI x RS x C	2.43	.122	.00
RCI x SS x C	5.12	.026	.00
$RS \ge SS \ge C$	2.86	.094	.00
RCI x RS x SS x C	0.00	.947	<.0