

# Unique motor plans facilitate learning during task switching, but at the expense of greater switch costs

Matthew J. Crossley<sup>1, 2</sup>, J. Vincent Filoteo<sup>3, 4</sup>, W. Todd Maddox<sup>5</sup>, and  
F. Gregory Ashby<sup>6</sup>

<sup>1</sup>School of Psychological Sciences, Macquarie University, Sydney, Australia

<sup>2</sup>Macquarie University Performance and Expertise Research Centre, Macquarie University,  
Sydney, Australia

<sup>3</sup>Research Service, VA San Diego HealthCare System

<sup>4</sup>Departments of Psychiatry and Neuroscience, School of Medicine, University of  
California San Diego

<sup>5</sup>Cognitive Design and Statistical Consulting, Austin, Texas

<sup>6</sup>Department of Psychological & Brain Sciences, University of California, Santa Barbara

---

Correspondence: Matthew J. Crossley, PhD, School of Psychological Sciences, Macquarie University, Australian Hearing Hub, 16 University Ave, Macquarie University, NSW 2109, Australia. Email: matthew.crossley@mq.edu.au

**Author Notes:** The research described in this article was supported in part by AFOSR grant FA9550-06-1-0204 PI: Maddox.

## Abstract

Many factors influence task switching, including attention, motor resources, and the different types of memory systems required for each component task. Even so, prior task-switching studies have focused almost exclusively on tasks that do not require significant learning. An experiment is reported that investigates how attention, motor planning, and memory systems influence the learning of two categorization tasks when those tasks are: (1) completely novel to participants, and (2) must be switched between on a random trial-by-trial basis during the learning process. Results indicate that associating each category label with a unique motor response facilitates the learning of each task during task switching. Participants were more accurate on switch trials with unique motor responses, but at the cost of longer response times. A novel cognitive control model is proposed that successfully accounts for these results. Finally, these results are discussed within the context of existing category-learning and cognitive-control theories, none of which predict the observed pattern of results.

## Introduction

In task-switching experiments, participants are asked to perform two distinct tasks in a pseudorandom, interleaved order (Kiesel et al., 2010; Monsell, 2003). These studies reveal that task switching is costly – for example, switch trials reliably increase response times (RTs) and often decrease accuracy relative to stay trials. Many factors are known to influence switch costs, including the number and identity of response options (Philipp & Koch, 2010, 2011; Philipp et al., 2013), the complexity of the stimuli (Witt & Stevens, 2013), the abstractness of the rules (Stelzel et al., 2011), the perceptual and attentional demands of the component tasks (Arrington et al., 2003; Chiu & Yantis, 2009; Nagahama et al., 2001; Ravizza & Carter, 2008; Rushworth et al., 2002), and the underlying memory systems supporting performance of each task (Crossley et al., 2018; Turner et al., 2017). The task-switching literature, however, has mostly focused on switching between tasks that are well-learned and can be performed with high accuracy when in a single-task context. To our knowledge, no study has examined task switching between two tasks that must be simultaneously learned from scratch. Thus, little is known about how the control processes that drive task switching interface with the processes that drive learning of the tasks themselves.

Models of cognitive control applied to task switching paradigms do not explicitly address this absence (Abrahamse et al., 2016; Blais et al., 2007; Botvinick et al., 2001; Verguts & Notebaert, 2008). Rather, since the tasks they are leveraged to model are always well-learned, these models assume learning of the tasks has already occurred and does not proceed further during epochs of task switching. On the other hand, models of task learning do make clear assumptions both about learning processes and about switching processes. For example, current multiple-systems models of category learning assume that trial-by-trial switching routinely occurs either between competing rule-sets (e.g., one explicit rule versus some other explicit rule), or else between competing memory systems (e.g., an explicit rule versus an implicit stimulus-response map) depending on which strategy carries the most confidence on each trial (Ashby et al., 1998; Erickson & Kruschke, 1998). Furthermore, these models assume independent learning in each sub-system, and thus frequent switching

between systems or strategies is predicted to have no ill effect on task learning. In short, these models predict that learning during task switching should be no more difficult than learning under single-task conditions.

The available empirical data, although sparse, suggest that switching during category learning is considerably more difficult than predicted by any current model (Crossley et al., 2018; Erickson, 2008; H  lie & Fansher, 2018). Moreover, even after current models are allowed to learn each task to asymptote, they fail to predict the switch costs routinely observed in task switching paradigms using well-learned tasks. Each of these failures imply a major error in the assumptions of existing category-learning models. It seems clear that the cognitive control processes that drive task switching interface with the processes that drive task learning in a non-trivial way. However, to build models that are imbued with both processes, a more robust empirical database of what drives task learning during task switching is required. Here, our goal is to begin populating this database. We investigate how attention, motor planning, and memory systems influence the learning of two novel category learning subtasks when those tasks must be switched between on a random trial-by-trial basis.

## Methods

### Design

We examined category learning while participants switched on a random trial-by-trial basis between two different categorization tasks. Each task was either a rule-based (RB) category-learning task, in which the optimal strategy is a simple explicit rule, or an information-integration (II) category-learning task, in which the optimal strategy is similarity based, and has no simple verbal description (Ashby & Gott, 1988; Ashby et al., 1998). On every trial in both tasks, the participant was instructed to assign the single presented stimulus to its correct category. The stimuli in both tasks were ellipses that varied across trials in the length and orientation of the major axis. All stimuli were presented on a gray background. Ellipses from the first task were presented in white, whereas ellipses from the second task were presented in black. The experiment included 10 different conditions that are described in Table . The various conditions were distinguished from each other in the following ways:

1. Attention (one versus two relevant stimulus dimensions). Some categories required attention to only a single dimension (1D categories) whereas others required attention to both dimensions (2D categories) in order for perfect accuracy to be achieved.
2. Motor planning (same motor plan versus different motor plans). Some conditions used the same two response keys for each task (“same” conditions) whereas others used different response keys for the two tasks (“different” conditions).
3. Memory system best suited for each task (procedural memory versus declarative memory). Much prior research suggests that RB categories are learned via declarative memory, whereas II categories are learned procedurally (e.g., Ashby and Maddox, 2010; Ashby and Valentin, 2017; Smith, Berg, et al., 2012). Therefore, we investigated between-system switching by requiring participants to switch between RB and

II categories, and we examined within-system switching by requiring participants to switch between two RB tasks or two II tasks.

Each participant completed only one condition. Condition names follow a simple format. The first two letters abbreviate the category structure used for the first task, and the second two letters abbreviate the category structure used for the second task (“ud” stands for “unidimensional”, “cj” stands for “conjunction”, and “ii” stands for “information-integration”). The number appended to the end of the condition name indicates the number of response keys used (2 for the “same” conditions and 4 for the “different conditions”).

**Table 1**

*Experimental conditions.*

Attention	Motor Plan	Memory System	Name	task 1	task 2
1D	different	between	udii4	1D RB	II
		within	udcj4	1D RB	2D RB
	same	between	udii2	1D RB	II
		within	udcj2	1D RB	2D RB
2D	different	between	cjii4	2D RB	II
		within	cj4cr	2D RB	2D RB
		within	ii4cr	II	II
	same	between	cjii2	2D RB	II
		within	cj2cr	2D RB	2D RB
		within	ii2cr	II	II

## Participants

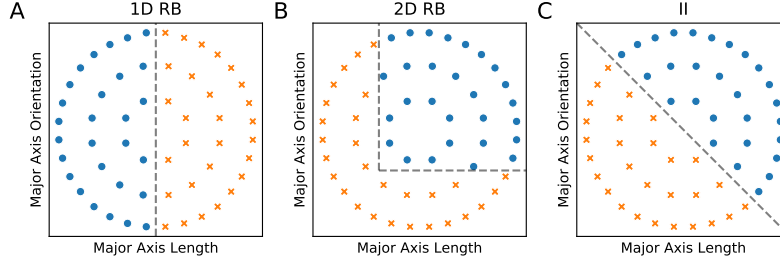
We recruited 150 participants from the University of Texas at Austin undergraduate population to serve as participants. Each participant was randomly assigned to one of the 10 possible conditions outlined in Table . There were 14 participants in condition cjii2, 16 in cjii4, 10 in udcj2, 16 in udcj4, 14 in udii2, 16 in udii4, 16 in cj2, 16 in cj4, 16 in ii2, and 16 in ii4. All participants completed the study and received course credit for their participation. All participants had normal or corrected-to-normal vision.

## Stimuli and Categories

The stimuli were ellipses in which the major axis varied across trials in length and orientation. The length of the major axis varied between 1.5 and 2.5 times the length of the minor axis (which was always 40 pixels), and the orientation of the major axis varied between  $\pi/16$  and  $7\pi/16$  radians counterclockwise rotation from horizontal. All stimuli were presented against a gray background. Stimuli from one task were presented in white, whereas stimuli in the other task were presented in black.

All categories were constructed from the same 56 stimuli (28 for each category) that were arranged in the three concentric circles shown in Figure 1. From these stimuli we created 3 distinct category structures: 1D RB structures were created by labeling all stimuli to one side of a vertical line “A” and all the others “B”; 2D RB structures were

created by labeling each stimulus consistent with a conjunction rule of the form “respond A if the stimulus has a large value on both stimulus dimensions; respond B otherwise”; II categories were constructed by labeling as category “A” all stimuli for which the numerical value that identified orientation was greater than the numerical value that identified length, and all other stimuli were assigned to category “B”.



**Figure 1**

**A:** The 1D RB categories used as a task in the *udcj2*, *udcj4*, *udii2*, and *udii4* conditions. **B:** The 2D RB categories used as a task in the *cjii4*, *udcj2*, *udcj4*, *cj2*, and *cj4* conditions. **C:** The II categories used as a task in the *cjii4*, *udcj2*, *udcj4*, *udii2*, *udii4*, *ii2*, and *ii4* conditions. Different categories are indicated both by different colors and by different marker types.

In conditions where participants switched from one 2D RB category structure to another 2D RB category structure, we used the same category structure shown in panel Figure 1B, except the response assignments were reversed in the two tasks. Similarly, in conditions where participants switched between two II tasks, we used the same II category structure shown in panel Figure 1C for both tasks, except we reversed the response assignments.

## Procedure

Participants in all conditions were told that they were to categorize ellipses on the basis of their major axis length and orientation, and that each category was equally likely. They were also instructed that black and white ellipses may or may not require a different response policy. Each participant completed a single session consisting of 560 trials, with each task interleaved psuedo-randomly. Each trial began with the presentation of a response-terminated stimulus. Participants made their response via the “c”, “v”, “b”, and “n” keys for conditions that used different response keys and category labels for each subtask, and the “z” and “m” keys for conditions that used the same response keys for each task. Feedback was presented immediately after stimulus offset. The word “correct” was shown in green font for correct responses, and the word “wrong” was shown in red font following incorrect responses. The duration of each was 750 ms. The intertrial interval between feedback offset and the next stimulus presentation was 1500 ms.

### Learning curve models

Data from each participant were aggregated across all 560 trials into 20 blocks of 28 trials each. To examine the effects of the various independent variables on performance, we fit the following hyperbolic tangent function to these block-by-block data separately for each participant:

$$P(n) = A \tanh[B(n - 1)] + C, \quad (1)$$

where  $P(n)$  is proportion correct on block  $n$ , and  $A$ ,  $B$ , and  $C$  are free parameters (all bounded between 0 and 1). This function equals  $C$  when  $n = 1$ , asymptotes at  $A + C$ , and increases at a rate determined by  $B$ . The best fitting parameters for each participant were estimated via least squares minimization using the Trust Region Reflective algorithm (Branch et al., 1999). Statistical significance of differences in any dependent measure between conditions was assessed using factorial ANOVA and independent samples t-tests.

### Decision-bound models

To identify the decision strategies that participants learned in each subtask, we fit decision-bound models to the trial-by-trial response data of each participant during the final four blocks (112 trials) of training separately for each subtask (Ashby & Gott, 1988; Maddox & Ashby, 1993). There were three types of models. One type assumed a 1-dimensional rule-based strategy, one type assumed a 2-dimensional rule-based strategy (some type of conjunction rule), and the third type assumed a 2-dimensional information-integration strategy. For details, see Ashby and Valentin (2017).

Briefly, the 1-dimensional rule models (or unidimensional rules) assumed that participants set a single criterion on one stimulus dimension and then gave one response if the stimulus had a value on this dimension that exceeded the criterion and otherwise gave the contrasting response. These models had two free parameters – namely, the value of the response criterion and the variance of perceptual and criterial noise.

The 2-dimensional rule models assumed participants used some type of conjunction rule created by setting single criteria on both stimulus dimensions. Note that this divides the stimulus space into four quadrants. The models, called general conjunctive classifiers (GCC), assumed participants gave one response to all stimuli falling in one of these four quadrants and otherwise gave the contrasting response. This strategy is consistent with some types of a conjunction rule (e.g., respond A if the major axis is small and has a steep orientation; otherwise respond B). The GCC has three free parameters: two response criteria (one of each dimension) and a noise variance.

The 2-dimensional information-integration models – i.e., the general linear classifier (GLC) – assumed that participants used a strategy that is consistent with a linear decision boundary of arbitrary slope and intercept. Stimuli falling on one side of this boundary are assigned to one category and stimuli falling on the other side are assigned to the contrasting category. Boundaries of this type are easily described mathematically (e.g., respond A if the numerical value of the length of the major axis is greater than the numerical value of the orientation of the major axis; otherwise respond B) but difficult to describe verbally and difficult to implement since length and orientation are non-commensurable. The GLC has three free parameters: the slope and intercept of the linear decision boundary and a noise variance.

In each condition, we computed the proportion of participants whose responses were best fit by each type of model, and then tested for significant differences in these proportions using  $\chi^2$  tests.

### A New Computational Model of Task Switching

Current computational models of task switching assume inhibition only at the cue or stimulus level (Abrahamse et al., 2016; Blais et al., 2007; Botvinick et al., 2001; Verguts & Notebaert, 2008). This allows the models to account for switch costs, but none of these models include a mechanism that would allow them to account for the facilitated learning we observed when the two tasks had unique, rather than identical motor responses. To overcome this limitation of the literature, we developed and tested a new computational model of task switching that assumed: (1) categories are already well learned; (2) on each trial, evidence for each response alternative accumulates over time via a diffusion process; (3) one form of cognitive control operates at the level of the cue by inhibiting response options irrelevant to the current task more than it inhibits the relevant response options; and (4) another form of cognitive control operates at the level of the motor plan, with the motor plan for each response option inhibiting the motor plan for every other response option (e.g., as in fully interconnected lateral inhibition).

The model is illustrated in Figure 2 for an application to the 1D RB conditions in which stimuli  $S_1 - S_{27}$  belong to category A in the first subtask and category C in the second subtask, and stimuli  $S_{28} - S_{56}$  belong to categories B and D in subtasks 1 and 2, respectively. Panel (a) illustrates how the model operates in the same motor plan conditions, and panel (b) shows an application of the model when the two subtasks have different motor plans.

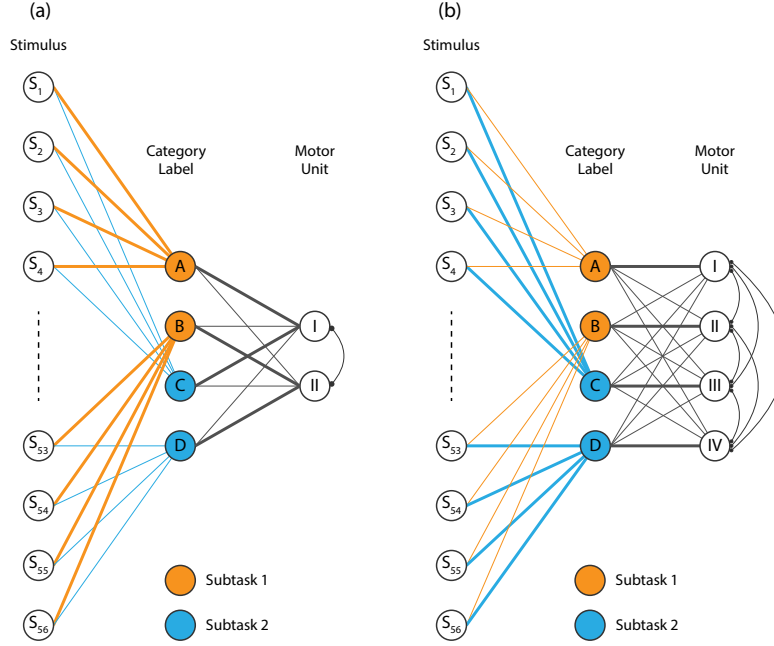
The thickness of the connections depict connection strength (i.e., not activation). The difference in connection strength between the stimulus layer and the category label layer (blue and orange lines) illustrate the effect of cue-level cognitive control. The connections among motor units are all inhibitory, which illustrates that the model assumes lateral inhibition among all output units. The thick lines connecting category label units to the appropriate motor units illustrates that the categories are assumed to be well learned.

Panel (a) depicts a subtask 1 trial, and the model assumes that the fact that the stimulus is presented in white inhibits all subtask 2 connection strengths between the stimulus and category label representations. Panel (b) depicts a subtask 2 trial, and the model assumes that the black stimulus color inhibits all subtask 1 stimulus-to-category-label connection strengths.

The model assumes that evidence for a particular response option accumulates on each trial in proportion to how well the current stimulus is learned, and is inhibited in proportion to both the cue-level and motor-level cognitive control terms. In particular, on any given trial, the evidence for response  $i$  at time  $t$ , denoted  $E_i(t)$ , is given by:

$$E_i(t) = E_i(t-1) + \alpha s_i - \beta c_i - \gamma \sum_{j \neq i} E_j(t) + \epsilon(t), \quad (2)$$

where  $s_i$  is the stimulus-driven evidence,  $c_i$  is the amount of cognitive inhibition,  $\gamma$  is the strength of the lateral inhibition from other response channels, and  $\epsilon(t)$  is white noise. The free parameters  $\alpha$ ,  $\beta$ , and  $\gamma$  were assigned the values  $\alpha = 0.1$ ,  $\beta = 0.1$ , and  $\gamma = 0.001$ ,

**Figure 2**

*Architecture of the cognitive control computational model for the 1D RB categorization task. Connections ending in a circle are inhibitory. (a) Connection strengths on a subtask 1 trial in the same motor plan condition. (b) Connection strengths on a subtask 2 trial in the different motor plan condition.*

except when reporting results from the model without one or the other forms of cognitive control, in which case they were set to zero.

The stimulus-driven evidence,  $s_i$ , was defined as  $s_i = 0.8$  if  $i$  was the correct response, and  $s_i = 0.2$  if response  $i$  was incorrect. The amount of cognitive inhibition,  $c_i$ , was different for switch and stay trials and for responses associated with the active and inactive tasks. On switch trials,  $c_i = 0.05$  for responses associated with the active task and  $c_i = 0.6$  for inactive task responses. On stay trials, these values were 0.01 and 0.8, respectively. Lateral inhibition from other response channels is modeled via the  $\sum_{j \neq i} E_j(t)$  term. In the conditions with two response alternatives, this sum included only a single term, whereas in the four-response conditions, this sum included three terms. Responses were triggered when  $E_i(t) > \Theta$ , where  $\Theta = 50$ .

Note that switch cost in the model emerges from cognitive inhibition (e.g., because  $c_i$  was different for switch and stay trials) and not lateral inhibition between motor units. Rather, lateral inhibition between motor units merely amplifies the switch cost originating with cognitive inhibition. This can be seen by recursively expanding the  $E_j(t)$  term for all available motor units. Doing this would reveal that there are effectively two  $\beta c_i$  terms in the same response conditions and four  $\beta c_i$  terms in the different response conditions.



## Results

### Learning Curves

Figures 3A and 3C show mean accuracy per block averaged over participants for each level of attention, memory system, and motor plan. Panel A shows accuracy in conditions that required participants to switch between a 1D rule and a subtask that required attention to two stimulus dimensions, and panel C shows accuracy in conditions in which participants switched between two subtasks that both required attention to two stimulus dimensions. Note that the largest effect, which is apparent in Figure 3C, is that providing four response alternatives, rather than two, appears to facilitate learning.

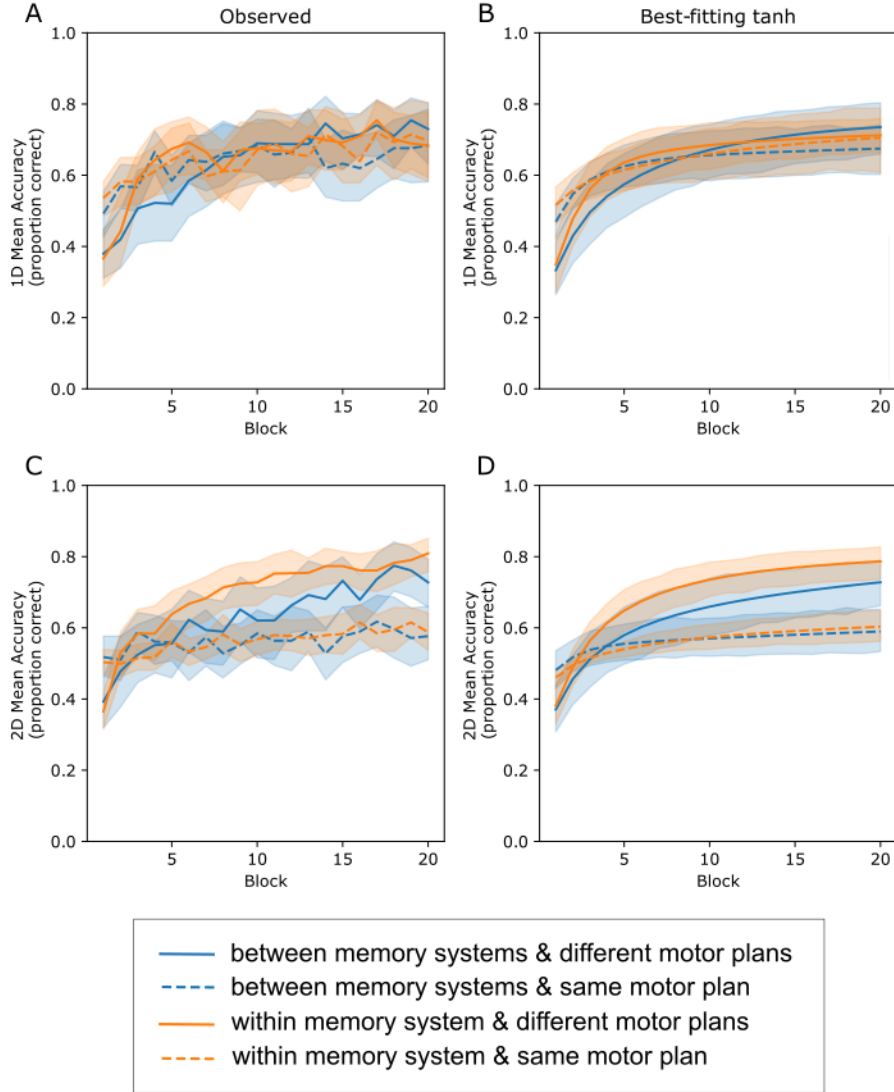
Figures 3B and 3D show the predictions of the best-fitting learning-curve models (see Equation 1), which accounted for 93% of the variance in the group level mean learning curves. One advantage of fitting this model is that it allows a more fine-grained investigation of how the three independent variables described in Table affect learning. Figure 4 shows how levels of attention (1D in left column versus 2D in right column), memory system (switching between systems in blue versus switching within systems in orange), and motor plan (different versus same responses on abscissa) affected initial accuracy ( $\hat{A}$  in row 1), learning asymptote ( $\hat{A} + \hat{C}$  in row 2), and learning rate ( $\hat{B}$  in row 3).

Figures 4A and 4B show a significant main effect of motor plan [ $F(1.0, 142.0) = 29.15, p < 0.001, \eta_p^2 = 0.17$ ], indicating that initial accuracy was lower when there were 4 response alternatives compared to only 2 [ $t(124.54) = 5.48, p < 0.001, d = 0.86$ ]. Figures 4C and 4D show a significant interaction between motor plan and attention [ $F(1.0, 142.0) = 5.07, p < 0.05, \eta_p^2 = 0.03$ ], indicating that unique motor plans led to higher learning asymptotes in the 2D attention conditions [ $t(78.83) = -2.61, p < 0.01, d = 0.54$ ] than in the 1D conditions [ $t(31.87) = 0.67, p = 0.5, d = 0.2$ ]. Figures 3C and 3D appear to indicate that under 2D attention conditions, unique motor plans benefited within-system switching more than between-system switching (i.e., by raising learning asymptotes). However, Figures 4C and 4D show that neither the motor plan  $\times$  memory system interaction [ $F(1.0, 142.0) = 0.26, p = 0.61, \eta_p^2 = 0.0$ ], nor the three-way motor plan  $\times$  attention  $\times$  memory system interaction [ $F(1.0, 142.0) = 0.45, p = 0.51, \eta_p^2 = 0.0$ ] were significant. A direct comparison via posthoc t-test was also non-significant [ $t(16.95) = -1.43, p = 0.17, d = 0.56$ ]. Even so, the effect size  $d = 0.56$  of this effect is fairly large. The combination of failure to reject the null with fairly large effect size may indicate that our study was under-powered with respect to this effect.

Figures 3E and 3F show that unique motor plans led to different learning effects when switching between memory systems versus when switching within memory systems [motor plan  $\times$  memory system interaction:  $F(1.0, 142.0) = 4.31, p < 0.05, \eta_p^2 = 0.03$ ]. Lower learning rates applied to between-system switching under unique motor plans than under same motor plans [ $t(49.08) = 2.23, p < 0.05, d = 0.59$ ]. When switching within memory systems, learning rates did not significantly change as a function of motor plan [ $t(76.2) = -0.05, p = 0.96, d = 0.01$ ].

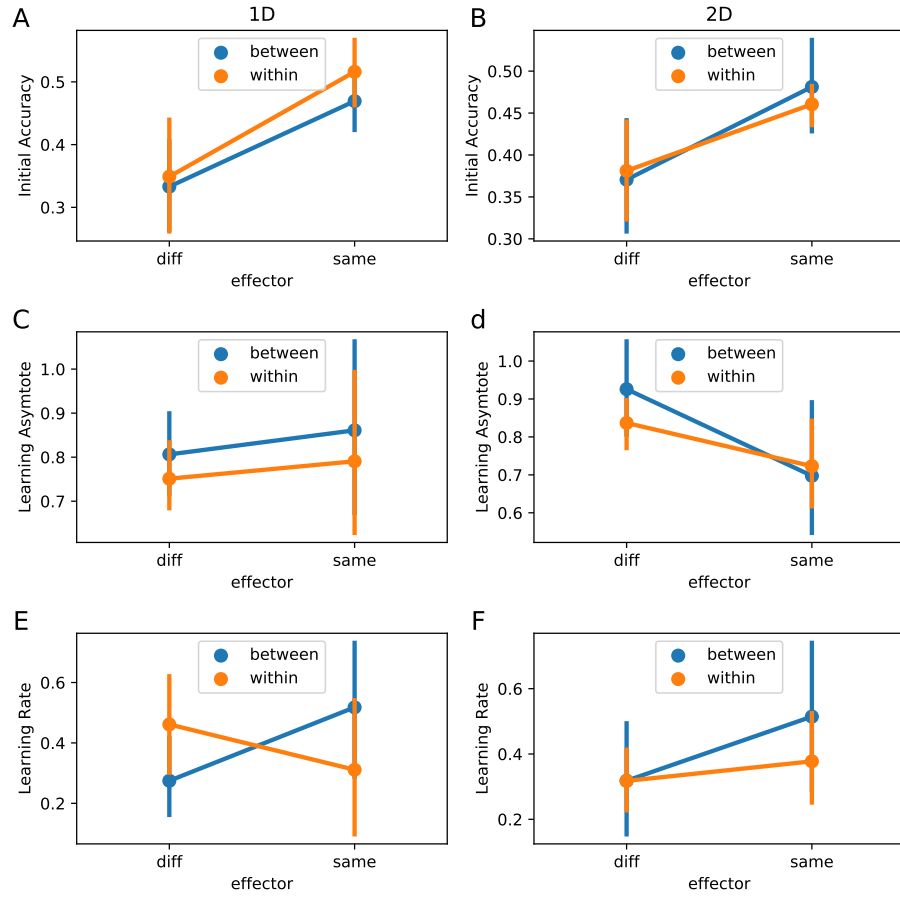
### Task Switching

Figure 5 shows mean accuracy per trial type for stay and switch trials. The left panel shows results when both subtasks are RB tasks and the right panel shows results when one

**Figure 3**

**A:** Mean accuracy per block in conditions that required switching between a 1D and 2D task. **B:** Mean accuracy predicted by the best-fitting tanh model per block in conditions matched to those in panel A. **C:** Mean accuracy per block in conditions that required switching between two 2D tasks. **D:** Mean accuracy predicted by the best-fitting tanh model per block in conditions matched to those in panel C. Error bands in all panels are 95% confidence intervals.

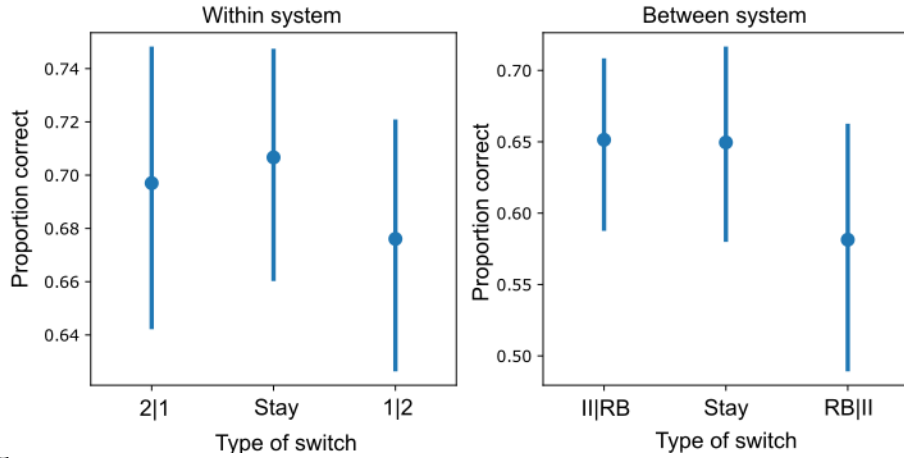
subtask is RB and the other is II. Mean accuracy in the within-system conditions on trials in which a switch occurred from subtask 1 to subtask 2 (denoted 1|2) was not significantly different from the reverse switch type (denoted 2|1) [ $t(31.0) = 1.642, p = 0.332, d = 0.14$ ], nor were 2|1-type switch trials different from stay trials [ $t(31.0) = -1.139, p = 0.791, d =$

**Figure 4**

**A:** Mean initial accuracy (i.e.,  $\hat{A}$ ) across participants in conditions that require switching between 1D and 2D tasks. **B:** Mean initial accuracy (i.e.,  $\hat{A}$ ) across participants in conditions that require switching between two 2D tasks. **C:** Mean learning asymptote (i.e.,  $\hat{A} + \hat{C}$ ) across participants in conditions that require switching between a 1D and a 2D task. **D:** Mean learning asymptote (i.e.,  $\hat{A} + \hat{C}$ ) across participants in conditions that require switching between two 2D tasks. **E:** Mean learning rate (i.e.,  $\hat{B}$ ) across participants in conditions that require switching between a 1D and a 2D task. **F:** Mean learning rate (i.e.,  $\hat{B}$ ) across participants in conditions that require switching between two 2D tasks. Error bars in all panels are 95% confidence intervals.

$-0.066$ ]. However, accuracy was significantly worse on 1|2 switch trials than on stay trials [ $t(31.0) = 3.325, p < 0.01, d = 0.231$ ].

A similar trend was apparent in the between-system conditions. Mean accuracy on trials in which a switch occurred from the II categories to the RB categories (denoted RB|II) was significantly worse than when the switch was in the opposite direction (denoted II|RB) [ $t(15.0) = 3.693, p < 0.01, d = 0.452$ ], and was also significantly worse than on stay trials [ $t(15.0) = 4.082, p < 0.01, d = 0.425$ ]. Accuracy on stay trials was not significantly different from accuracy on II|RB trials [ $t(15.0) = 0.148, p = 1.0, d = 0.014$ ]. Taken together, *RB|II* trials – that is, performing an RB trial after having just performed an II trial – appears more difficult than any other trial type.



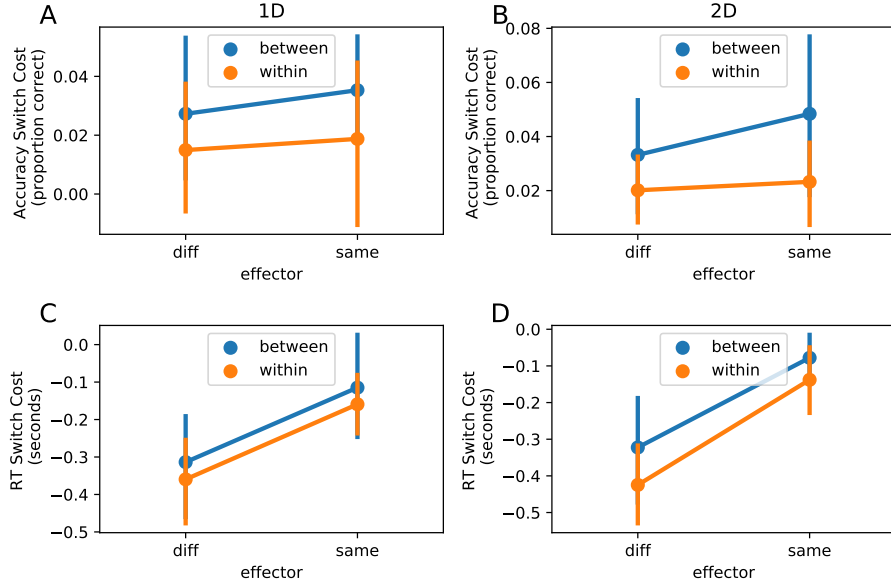
**Figure 5**

**Left:** Mean accuracy per trial type in the 2D within-system conditions. 2|1 indicates a switch from subtask 1 to subtask 2, whereas 1|2 indicates the opposite type of switch. **Right:** Mean accuracy per trial type in the 2D between-system conditions. II|RB indicates a switch from an RB subtask to an II subtask, whereas RB|II indicates the opposite type of switch. Error bars in both panels indicate 95% confidence intervals.

Taken together, these results indicate that much of the switch costs reported below may come from one switch direction more than another. Further research is required to ascertain if this effect is specifically tied to special difficulties with between-system switching or if it merely reflects participants willingness to perform poorly on an arbitrary task in order to perform slightly better on the other.

Figure 6 shows switch costs for levels of attention (1D tasks in left column versus 2D tasks in right column), memory system (between-system switching in blue versus within-system switching in orange), and motor plan (different versus same), separately for accuracy (top row) and RT (bottom row). An ANOVA indicated that none of the accuracy differences were significant, though the main effect of memory system was close [ $F(1.0, 142.0) = 3.29, p = 0.07, \eta_p^2 = 0.02$ ], with accuracy switch cost trending greater for between-system switching as compared to within-system switching [ $t(119.3) = 1.7, p = 0.09, d = 0.29$ ].

A separate ANOVA showed a significant main effect of motor plan on RT switch costs [ $F(1.0, 142.0) = 25.92, p < 0.001, \eta_p^2 = 0.15$ ], indicating that RT switch costs were sig-

**Figure 6**

**A:** Accuracy switch costs in conditions that require switching between a 1D and a 2D task. **B:** Accuracy switch costs in conditions that require switching between two 2D tasks. **C:** RT switch costs in conditions that require switching between a 1D and a 2D task. **D:** RT switch costs in conditions that require switching between two 2D tasks. Cost in all panels is defined by stay value minus switch value. Error bars in all panels are 95% confidence intervals.

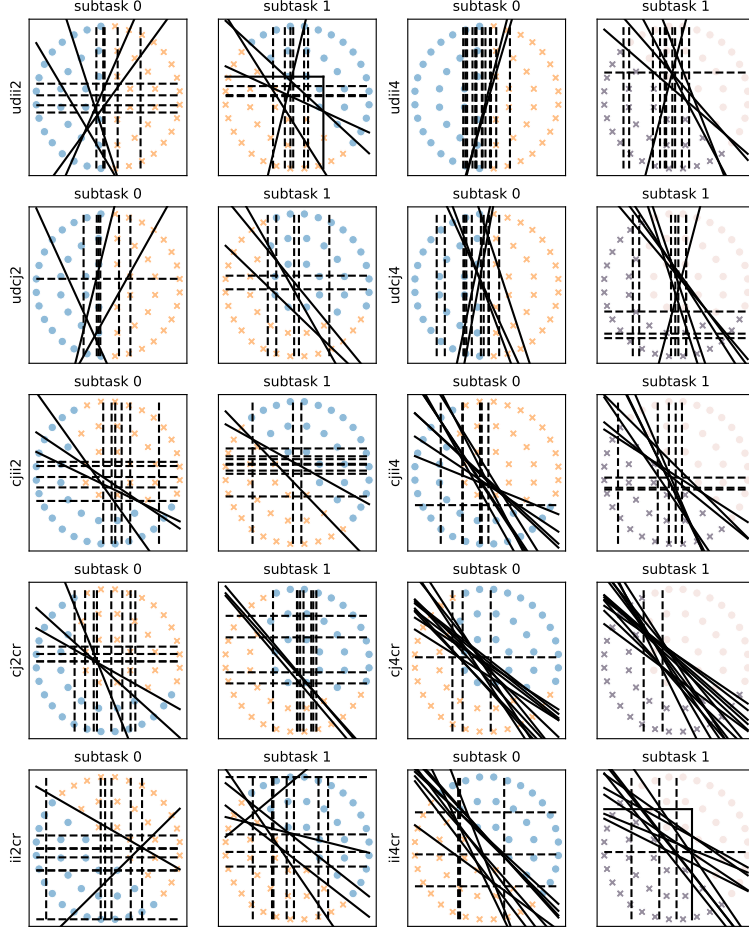
nificantly increased by unique motor plans [ $t(146.94) = 5.92, p < 0.001, d = 0.96$ ], regardless of demands on attention and memory systems.

### Decision Bound Models

Figure 7 shows the decision bounds from the best-fitting decision-bound models in all conditions overlaid on the underlying category distribution for each subtask. The first major takeaway from Figure 7 is that in all same-motor-plan conditions (udii2, udcj2, cjii2, cj2cr, ii2cr; all shown in the left two columns of Figure 7), both subtasks were predominantly best fit by 1D RB models regardless of the underlying category structure. This resonates with our accuracy-based finding that performance in these conditions was generally poor.

The second major takeaway from Figure 7 is that in all conditions where at least one subtask was 1D RB (udii2, udii4, udcj2, udcj4; all shown in the top two rows of Figure 7), both subtasks were predominantly best fit by 1D RB models regardless of whether same motor plans or different motor plans were used. This finding suggests that even though learning occurred in all of these conditions (see Figure 4), it came mostly from participants using a sub-optimal 1D RB strategy to achieve above-chance, but suboptimal performance on the 2D subtask.

The third major takeaway from Figure 7 is that in all conditions where both subtasks

**Figure 7**

*Category structures and decision bounds from the best-fitting decision bound model for each participant and subtask in each condition during the final 112 trials of training.*

were 2D (cjii2, cjii4, cj2cr, cj4cr, ii2cr, ii4cr), the decision bounds from the best-fitting models were predominantly 2D (e.g., GLC). This shows that participants were indeed able to learn a 2D strategy while switching tasks, and suggests that the deficit observed in conditions that contain a 1D RB task reflects the influence of the 1D RB task itself, not the complexity or difficulty of the 2D task.

Table 2 reports the results of binomial tests of whether the number of participants whose responses were best fit by a decision-bound model that assumed a decision bound of the optimal type was significantly greater than chance for each condition and subtask. This table shows that only the cj4cr condition had a significantly greater than chance number of participants who were best fit by an optimal model on both subtasks. Many participants in the ii4cr condition were also best fit by a model of the optimal type, but this was only greater than would be expected by chance on a single subtask. Together,

these results (1) underscore the importance of unique motor plans in learning during task switching, (2) suggest that switching between two categories of the same type is easier than switching between two categories of different types, and (3) suggest that switching between RB categories of the same type may be the easiest of all.

**Table 2**

*Binomial tests of whether the number of participants whose responses were best fit by a decision-bound model of the optimal type is significantly greater than chance for each condition and subtask.*

Condition	subtask	
udii4	1D RB	$\mathcal{B}(16, 0.5) = 14, p = 0.0^{**}$
	II	$\mathcal{B}(16, 0.5) = 5, p = 0.96$
udcj4	1D RB	$\mathcal{B}(16, 0.5) = 11, p = 0.11$
	2D RB	$\mathcal{B}(16, 0.5) = 6, p = 0.89$
udii2	1D RB	$\mathcal{B}(14, 0.5) = 9, p = 0.21$
	II	$\mathcal{B}(14, 0.5) = 5, p = 0.91$
udcj2	1D RB	$\mathcal{B}(10, 0.5) = 7, p = 0.17$
	2D RB	$\mathcal{B}(10, 0.5) = 3, p = 0.95$
cjii4	2D RB	$\mathcal{B}(16, 0.5) = 9, p = 0.4$
	II	$\mathcal{B}(16, 0.5) = 8, p = 0.6$
cj4cr	2D RB	$\mathcal{B}(16, 0.5) = 12, p = 0.04^*$
	2D RB	$\mathcal{B}(16, 0.5) = 14, p = 0.0^{**}$
ii4cr	II	$\mathcal{B}(16, 0.5) = 10, p = 0.23$
	II	$\mathcal{B}(16, 0.5) = 12, p = 0.04^*$
cjii2	2D RB	$\mathcal{B}(14, 0.5) = 3, p = 0.99$
	II	$\mathcal{B}(14, 0.5) = 3, p = 0.99$
cj2cr	2D RB	$\mathcal{B}(16, 0.5) = 4, p = 0.99$
	2D RB	$\mathcal{B}(16, 0.5) = 4, p = 0.99$
ii2cr	II	$\mathcal{B}(16, 0.5) = 2, p = 1.0$
	II	$\mathcal{B}(16, 0.5) = 5, p = 0.96$

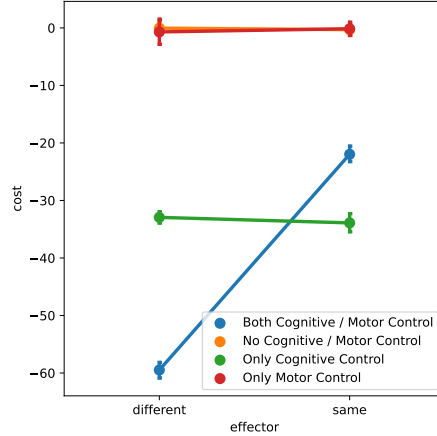
### Cognitive Control Computational Model

We used the cognitive control model described earlier to simulate the RT on 500 trials from 15 hypothetical participants. The simulations used a time step of 1 ms, with a maximum RT of 1,000 ms, and the category label and cue value were selected randomly on each trial. Figure 8 shows the results of these simulations for four different versions of the model.

The first version was a control model that included no inhibition of any kind – that is, no cognitive control and no lateral inhibition among motor output units. This was implemented by setting  $\beta$  and  $\gamma$  to 0 in Equation 2. The results of this model are shown in orange. The second version assumed motor-level control but not cue-level control (i.e.,  $\beta = 0$  and  $\gamma > 0$ ). The predictions of this model are shown in red. The third version, which is similar to existing models in the literature, assumed cue-level control but not motor-level control (i.e.,  $\beta > 0$  and  $\gamma = 0$ ). The predictions of this model are shown in green. Finally, the full model included both cue-level and motor-level cognitive control terms (i.e.,  $\beta > 0$  and  $\gamma > 0$ ). This is the model described by Equation 2 and its predictions are shown in blue.

The predictions in Figure 8 should be compared to the empirical RT switch costs



**Figure 8**

*Computational model results when cue-level and motor-level control terms equal zero (orange), when only the motor-level control term is non-zero (red), when only the cue-level control term is non-zero (green), and when both cue-level and motor-level control terms are non-zero (blue).*

shown in Figures 6C and 6D. As expected, the no-cognitive-or-motor-control model predicts no switch costs of any kind. The model that assumes motor-level control but not cue-level control also predicts no switch costs of any kind. Recall that, as described above where we define the model, lateral inhibition between motor units merely amplifies the switch cost originating with cue-level control. The model that assumes cue-level cognitive control, but not motor-level control, which is similar to existing models in the literature, predicts a switch cost of equal magnitude for both motor plan conditions (same versus different). Therefore, this model fails to account for the larger switch costs in the four-response conditions compared to the two-response conditions that are evident in Figure 6. Finally, note that the full model – that is, the new model proposed here that includes both cue-level and motor-level control – correctly predicts that both conditions will incur a switch cost, and that switch costs should be greater when the two tasks have different motor plans than when they share the same motor plan.

## Discussion

We investigated how simultaneous category learning and task switching are influenced by attention, motor planning, and memory systems. We found that providing participants with unique response keys in the two subtasks facilitated task learning during task switching in attention-demanding environments. Specifically, unique motor plans led to a higher learning asymptote in 2D attention conditions but not in 1D attention conditions. In every condition, the learning benefit imbued by unique motor plans came at the price of increased RT switch costs. Moreover, unique motor plans led to lower initial accuracy regardless of attention and memory system demands, perhaps reflecting difficulty in remembering which motor plans are correctly associated with each context.

### Relation to earlier work on task switching

Our data can be situated within the broader task-switching and cognitive control literature, though in every case heavy caveats must be applied, given that our study is the first to examine control / switching processes during initial task learning.

Non-overlapping response sets (i.e., unique motor plans) have been shown to reduce interference (often in the context of a Stroop task; Klein, 1964; Mayr, 2001; Redding and Gerjets, 1977; Yeung and Monsell, 2003), which is the reverse of what we observed here (i.e., greater switch cost for unique motor plans than for redundant motor plans). On the other hand, studies of task switching that manipulated response sets and response rules – both of which vary in the categorization tasks that we used – have reported that each is costly on its own (Philipp & Koch, 2011; Philipp et al., 2013), and that if both factors are changed during a task switch, then the largest costs are observed (Philipp & Koch, 2010). Consistent with this, we found larger switch costs for unique motor plans than for redundant motor plans. What sets our findings apart from these earlier results, is the observation that the increased cost of unique motor plans was a necessary price to pay in order for task learning to occur in the first place.

It has also been observed that it is harder to switch to the better-practiced of two tasks (Allport et al., 1994; Yeung & Monsell, 2003). This may in some ways anticipate our finding of increased switch costs for unique motor plans. In particular, since unique motor plans facilitate task learning, they would also usher in increased switch costs as the tasks become – in a sense – better practiced.

Finally, an old idea in the task switching literature is that task similarity – for example, of stimuli, responses, attention, etc – dictates the ease of switching (Arrington et al., 2003). Here, highly similar tasks are thought to be the easiest to switch between, since they involve switching between fewer cognitive processes. This idea would seem to predict: 1) that switch costs will be greater for unique motor plans than for redundant; 2) that switching between two tasks mediated within the same memory system ought to cost less than switching between tasks mediated by different systems; and 3) that switching between two 2D categorization tasks ought to cost less than switching between a 1D and a 2D task. Our data are inconsistent with all of these predictions, so it would seem that switching during task learning may entail an interaction of processes that behave differently than when switching between well-learned tasks.

### Relation to earlier work on category learning

Previous category-learning research indicates that switching between systems may be especially difficult – as compared to switching between explicit rules – though memory system effects can be difficult to dissociate from attention effects (but see Ashby et al., 2020; Smith, Crossley, et al., 2012). This is because 1D RB categories are often used as a means of accessing declarative systems, whereas 2D II categories are often used to access procedural systems. Here, we used both 1D and 2D RB categories, and are thereby well positioned to avoid this potential confound.

Our results are broadly consistent with the idea that within-system switching is easier than between-system switching, and that switching between RB tasks is easier than switching between II tasks. First, we found that the accuracy-based learning benefit imbued

by unique motor plans appeared greater for within-system switching than for between-system switching, although this difference was not statistically significant. Even so, the effect size was moderate, indicating that we may simply have lacked statistical power to detect this effect. Second, we found lower learning rates for between-system switching under unique motor plans than under redundant motor plans. In contrast, when participants switched within memory systems, learning rates did not significantly change as a function of motor plan. Third, there was a trend for greater accuracy switch costs during between-system switching as compared to within-system switching. Though this trend was non-significant, it also carried a moderate effect size, perhaps reflecting insufficient power to detect this particular effect. Finally, the number of participants whose responses were best fit by a model of the optimal type was only significantly greater than chance on both subtasks when participants were switching between two 2D RB tasks.

Our data also showed that participants failed to adopt a 2D categorization strategy in conditions that included a 1D RB subtask. This effect mostly trumps the benefit of having unique motor plans in each subtask – that is, even in conditions that required unique motor plans in each subtask, a 1D subtask structure reliably predicted that participants would fail to adopt a 2D response strategy on the second subtask. The primacy of simple 1D RB strategies and their ability to block the formation of more accurate 2D response strategies converges with earlier observations from our labs Gregory Ashby and Crossley, 2010. In essence, it seems that if reasonably good task performance can be obtained with simple 1D RB strategies, then participants will rarely if ever abandon them for more complicated 2D strategies.

### Relation to existing models

To our knowledge, our study is the first to investigate task switching during initial task learning. As such, our results represent an important empirical anchor for the development of models that fuse cognitive control and learning into one cohesive framework. Extant cognitive control models assume the tasks being controlled (e.g., switched between) are already well learned (Abrahamse et al., 2016; Blais et al., 2007; Botvinick et al., 2001; Verguts & Notebaert, 2008), so these models should only be applied to our data with caution. Several category-learning theories, however, make predictions not just about learning, but also about switching. These models assume that trial-by-trial switching routinely occurs both between competing rule-sets and between competing memory systems, and further assume that switching does not impair initial learning in any way (Ashby et al., 1998; Erickson & Kruschke, 1998). Switching is assumed to occur with ease simply on the basis of whatever system or strategy is most confident on a given trial. Learning can occur robustly in the presence of frequent switching in these models because they assume each system and strategy gets independent feedback to use for learning.

On the other hand, it is also important to note that extant category-learning theories were not designed to account for task-switching data. For example, when the first multiple-systems models were developed (Ashby et al., 1998; Erickson & Kruschke, 1998) there were virtually no relevant task-switching data available that could be used to guide the assumptions these models made about system switching. Although both models assume that switching occurs with ease, they are not explicitly equipped with algorithms to interpret and process contextual cues such as the stimulus colors that were used in the present study

to cue the relevant response policy. Thus, the present data may be seen as outside the scope of these models. To the degree that predictions can be derived from them, they would seem to predict that switching occurs often, and without cost to accuracy or RTs attributable to the switching mechanism itself. The first prediction is at odds with the observation here and in previous reports that task switching in category learning is very difficult, even after substantial practice (Crossley et al., 2018; Erickson, 2008; Turner et al., 2017), and the second prediction is at odds with an enormous literature reliably indicating switch costs are due to task switching (Kiesel et al., 2010; Monsell, 2003). Therefore, our results seem to suggest that the trial-by-trial switching assumptions made by current category-learning theories might need some revision.

## Conclusions

This article makes several novel contributions. First, to our knowledge, this is the first article to investigate task switching when the two tasks are novel and must be learned. Second, our results document a number of novel effects that motor planning can have on task switching. In particular, we found that in attention-demanding tasks (i.e., our 2D conditions) learning is better when each task includes unique motor responses than when the same motor responses are required in both tasks (see Figure 4D). This is especially impressive given that initial accuracy was lower in unique motor response conditions (see Figures 4A and 4B). On the other hand, this improved performance came at the cost of higher RT switch costs (see Figures 6C and 6D). Third, we proposed a novel computational model of task switching that accounts for these results. In addition to including a cognitive (i.e., cue level) control mechanism, which appears in virtually all existing models, our new model also includes a unique motor control mechanism in which there is competition among competing response options. It seems clear that the cognitive control processes that drive task switching interface with the processes that drive task learning in a non-trivial way. We hope that this article is the first of many to fuse these two important and interconnected facets of cognition and action.

## Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors did not use any generative AI or AI-assisted technologies. All content is the author's own. The authors take full responsibility for the content of the publication.

## References

- Abrahamse, E., Braem, S., Notebaert, W., & Verguts, T. (2016). Grounding cognitive control in associative learning. *Psychological Bulletin*, 142(7), 693–728. <https://doi.org/10.1037/bul0000047>
- Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In *Attention and performance 15: Conscious and nonconscious information processing*. (pp. 421–452). The MIT Press.

- Arrington, C. M., Altmann, E. M., & Carr, T. H. (2003). Tasks of a feather flock together: Similarity effects in task switching. *Memory & Cognition*, 31(5), 781–789. <https://doi.org/10.3758/BF03196116>
- Ashby, F. G., & Maddox, W. T. (2010). Human category learning 2.0. *Annals of the New York Academy of Sciences*, 1224, 147–161.
- Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. [Publisher: American Psychological Association]. *Psychological Review*, 105(3), 442–481.
- Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of experimental psychology. Learning, memory, and cognition*, 14(1), 33–53. <https://doi.org/10.1037/0278-7393.14.1.33>
- Ashby, F. G., Smith, J. D., & Rosedahl, L. A. (2020). Dissociations between rule-based and information-integration categorization are not caused by differences in task difficulty. *Memory & Cognition*, 48(4), 541–552. <https://doi.org/10.3758/s13421-019-00988-4>
- Ashby, F. G., & Valentin, V. V. (2017). Multiple systems of perceptual category learning: Theory and cognitive tests. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science, Second Edition* (pp. 157–188). New York: Elsevier.
- Blais, C., Robidoux, S., Risko, E. F., & Besner, D. (2007). Item-specific adaptation and the conflict-monitoring hypothesis: A computational model. *Psychological Review*, 114(4), 1076–1086. <https://doi.org/10.1037/0033-295X.114.4.1076>
- Botvinick, M. M., Carter, C. S., Braver, T. S., Barch, D. M., & Cohen, J. D. (2001). Conflict Monitoring and Cognitive Control, 29.
- Branch, M. A., Coleman, T. F., & Li, Y. (1999). A Subspace, Interior, and Conjugate Gradient Method for Large-Scale Bound-Constrained Minimization Problems. *SIAM Journal on Scientific Computing*, 21(1), 1–23. <https://doi.org/10.1137/S1064827595289108>
- Chiu, Y.-C., & Yantis, S. (2009). A Domain-Independent Source of Cognitive Control for Task Sets: Shifting Spatial Attention and Switching Categorization Rules. *Journal of Neuroscience*, 29(12), 3930–3938. <https://doi.org/10.1523/JNEUROSCI.5737-08.2009>
- Crossley, M. J., Roeder, J. L., Helie, S., & Ashby, F. G. (2018). Trial-by-trial switching between procedural and declarative categorization systems. *Psychological Research*, 82(2), 371–384. <https://doi.org/10.1007/s00426-016-0828-4>
- Erickson, M. A. (2008). Executive attention and task switching in category learning: Evidence for stimulus-dependent representation [Publisher: Springer]. *Memory & Cognition*, 36(4), 749–761.
- Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning [Publisher: American Psychological Association]. *Journal of Experimental Psychology: General*, 127(2), 107–140.
- Gregory Ashby, F., & Crossley, M. J. (2010). Interactions between declarative and procedural-learning categorization systems. *Neurobiology of Learning and Memory*, 94(1), 1–12. <https://doi.org/10.1016/j.nlm.2010.03.001>

- Hélie, S., & Fansher, M. (2018). Categorization system-switching deficits in typical aging and Parkinson's disease. *Neuropsychology*, 32(6), 724–734. <https://doi.org/10.1037/neu0000459>
- 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>
- Klein, G. S. (1964). Semantic Power Measured through the Interference of Words with Color-Naming. *The American Journal of Psychology*, 77(4), 576. <https://doi.org/10.2307/1420768>
- Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics*, 53(1), 49–70. <https://doi.org/10.3758/BF03211715>
- Mayr, U. (2001). Age differences in the selection of mental sets: The role of inhibition, stimulus ambiguity, and response-set overlap. *Psychology and Aging*, 16(1), 96–109. <https://doi.org/10.1037/0882-7974.16.1.96>
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134–140. [https://doi.org/10.1016/S1364-6613\(03\)00028-7](https://doi.org/10.1016/S1364-6613(03)00028-7)
- Nagahama, Y., Okada, T., Katsumi, Y., Hayashi, T., Yamauchi, H., Oyanagi, C., Konishi, J., Fukuyama, H., & Shibasaki, H. (2001). Dissociable Mechanisms of Attentional Control within the Human Prefrontal Cortex. *Cerebral Cortex*, 11(1), 85–92. <https://doi.org/10.1093/cercor/11.1.85>
- Philipp, A. M., & Koch, I. (2010). The Integration of Task-set Components Into Cognitive Task Representations. *Psychologica Belgica*, 50(3-4), 383. <https://doi.org/10.5334/pb-50-3-4-383>
- Philipp, A. M., & Koch, I. (2011). The role of response modalities in cognitive task representations. *Advances in Cognitive Psychology*, 7(-1), 31–38. <https://doi.org/10.2478/v10053-008-0085-1>
- Philipp, A. M., Weidner, R., Koch, I., & Fink, G. R. (2013). Differential roles of inferior frontal and inferior parietal cortex in task switching: Evidence from stimulus-categorization switching and response-modality switching. *Human Brain Mapping*, 34(8), 1910–1920. <https://doi.org/10.1002/hbm.22036>
- Ravizza, S. M., & Carter, C. S. (2008). Shifting set about task switching: Behavioral and neural evidence for distinct forms of cognitive flexibility. *Neuropsychologia*, 46(12), 2924–2935. <https://doi.org/10.1016/j.neuropsychologia.2008.06.006>
- Redding, G. M., & Gerjets, D. A. (1977). Stroop Effect: Interference and Facilitation with Verbal and Manual Responses. *Perceptual and Motor Skills*, 45(1), 11–17. <https://doi.org/10.2466/pms.1977.45.1.11>
- 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/10.1162/089892902760807159>
- Smith, J. D., Berg, M. E., Cook, R. G., Murphy, M. S., Crossley, M. J., Boomer, J., Spiering, B., Beran, M. J., Church, B. A., Ashby, F. G., & Grace, R. C. (2012). Implicit and explicit categorization: A tale of four species. *Neuroscience & Biobehavioral Reviews*, 36(10), 2355–2369. <https://doi.org/10.1016/j.neubiorev.2012.09.003>

- Smith, J. D., Crossley, M. J., Boomer, J., Church, B. A., Beran, M. J., & Ashby, F. G. (2012). Implicit and explicit category learning by capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology*, 126(3), 294–304. <https://doi.org/10.1037/a0026031>
- Stelzel, C., Basten, U., & Fiebach, C. J. (2011). Functional Connectivity Separates Switching Operations in the Posterior Lateral Frontal Cortex. *Journal of Cognitive Neuroscience*, 23(11), 3529–3539. [https://doi.org/10.1162/jocn\\_a\\_00062](https://doi.org/10.1162/jocn_a_00062)
- Turner, B. O., Crossley, M. J., & Ashby, F. G. (2017). Hierarchical control of procedural and declarative category-learning systems. *NeuroImage*, 150, 150–161. <https://doi.org/10.1016/j.neuroimage.2017.02.039>
- Verguts, T., & Notebaert, W. (2008). Hebbian learning of cognitive control: Dealing with specific and nonspecific adaptation. *Psychological Review*, 115(2), 518–525. <https://doi.org/10.1037/0033-295X.115.2.518>
- Witt, S. T., & Stevens, M. C. (2013). fMRI task parameters influence hemodynamic activity in regions implicated in mental set switching. *NeuroImage*, 65, 139–151. <https://doi.org/10.1016/j.neuroimage.2012.09.072>
- Yeung, N., & Monsell, S. (2003). Switching between tasks of unequal familiarity: The role of stimulus-attribute and response-set selection. *Journal of Experimental Psychology: Human Perception and Performance*, 29(2), 455–469. <https://doi.org/10.1037/0096-1523.29.2.455>