



Ye, W., & Damian, M. F. (2022). Exploring task switch costs in a color-shape decision task via a mouse tracking paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, 48(1), 8-20. <https://doi.org/10.1037/xhp0000975>

Peer reviewed version

Link to published version (if available):
[10.1037/xhp0000975](https://doi.org/10.1037/xhp0000975)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via American Psychological Association at <https://psycnet.apa.org/doi/10.1037/xhp0000975> . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: <http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Running head: TASK SWITCHING AND MOUSE TRACKING

Exploring Task Switch Costs in a Colour-Shape Decision Task

via a Mouse Tracking Paradigm

Wenting Ye and Markus F. Damian

University of Bristol, United Kingdom

Word count: 8,236 (main text excl. Abstract, Figure legends and References)

Author's note:

Data and code will be made available here: <https://osf.io/u6gan/>

This work was supported by China Scholarship Council (PhD scholarship, 201808060111) to the first author, and research grant RPG-2019-054 by the Leverhulme Trust to the second author.

We would like to thank Dr. Felice van't Wout for invaluable advice on setting up the experiment.

Correspondence concerning this article should be addressed to Wenting Ye, School of Psychological Science, University of Bristol, 12a Priory Road, Bristol, BS8 1TU, United Kingdom. Email: wy14940@bristol.ac.uk

© 2021, American Psychological Association. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without authors' permission. The final article will be available, upon publication, via its DOI:

10.1037/xhp0000975

Abstract

Several decades of cognitive research have explored the processes and mechanisms which underlie task switching. Here we report an experiment in which young adult participants were presented with coloured shapes, and were randomly cued to categorise them according to colour, or to shape. Responses were made via dynamic movements of the computer mouse (“mouse tracking”), which allows insight into how decision making unfolds. The results showed that a range of classic findings (mix cost; switch cost; task congruency effects, etc.) emerged strongly in movement trajectories. Initiation of movements was sensitive to task/cue driven but not to stimulus/response driven variables, as well as to task switching. This constellation suggests that task switch costs emerge as a combination of task set reconfiguration, and stimulus-driven sensory-motor mappings.

Keywords: cognitive control; task switching; decision making; embodied cognition

Public significance statement: Switching between different mental activities is a key human skill which has been explored by a vast research literature. Here we explored task switching in an experimental study in which participants randomly oscillated between categorising coloured shapes by colour or by shape, and generated responses via computer mouse movements (“mouse tracking”). The results suggest that mental flexibility involves both a wilful reconfiguration between different tasks, and interference from previously learned associations between presented stimuli and responses.

Exploring Task Switch Costs in a Colour-Shape Decision Task via a Mouse Tracking Paradigm

The ability to interrupt a current activity, to shift to something entirely new and to flexibly alternate between different (sometimes entirely unrelated) tasks is one of the core components of executive control. A voluminous literature on task switching in laboratory-based tasks explores this flexibility via experiments in which individuals are shown stimuli with two different dimensions and are asked to classify them by either one or the other dimension. For instance, in the 'colour-shape' switching task (e.g., Prior & MacWhinney, 2010) coloured shapes are presented (e.g., red or blue triangles or circles), and on each trial participants are instructed to either categorise them according to colour, or to shape. In "pure" experimental blocks, the response dimension remains constant; in "mixed" blocks, participants are cued at the beginning of each trial to focus either on colour or on shape, perhaps via presentation of a rainbow (for the colour task) or a set of geometrical figures (for the shape). Two key observations are *mix costs* (slower latencies and/or more errors in mixed compared to pure blocks) and *switch costs* (slower latencies and/or more errors when in a mixed block, the type of categorisation has switched from the preceding trial, relative to when the task stayed the same). Since these observations were first reported (mix cost: Jersild, 1927; switch cost: Rogers & Monsell, 1995), many additional aspects of experimental task switching have been explored (for comprehensive reviews, see e.g., Meiran, 2010; Monsell, 2003; Vandierendonck et al., 2010).

Two main theoretical accounts of task switching have been put forward: task-set reconfiguration (TSR) views and interference accounts. Both views assume that task-driven behaviour requires the formation of a task set, i.e., a transient bundle of settings and parameters which determine behaviour in a given context (Logan & Gordon, 2001), such as stimulus categorisation, response selection and response execution. Both accounts also assume that task switching requires the reconfiguration of a subset of settings and parameters from the previous task set. According to the TSR view, task switching costs reflect the duration of the endogenous control processes which are required to detach from the

previous setting, and attach to the current one (e.g., Mayr & Kliegl, 2000; Rogers & Monsell, 1995). By contrast, interference (sometimes called “proactive interference”; Goschke, 2000) accounts attribute task switching costs to exogenously (rather than endogenously) triggered processes: switch costs arise from competing task activations between previous and current trials and task sets. For instance, Allport et al. (1994) suggested that the task set which was activated on a previous trial persists and interferes with the current task set, as the previous activation is no longer relevant and needs to be overcome. More specifically, a stimulus-response association might have been acquired in the previous trial, but the association would become irrelevant in the current trial (Wylie & Allport, 2000). Persisting activation from previous task set interferes with the relevant stimulus-response association triggered by the current task set (Waszak & Hommel, 2007), hence slowing down responses and rendering them more error prone. However, TSR and interference views are not mutually exclusive, and a third view states that both endogenous and exogenous control processes contribute to task switching costs (e.g., Arrington, 2008; Meiran et al., 2000; Vandierendonck et al., 2010).

A range of empirical findings constrain theorising on task switching (see Vandierendonck et al., 2010, for a comprehensive list of benchmark findings) and a few will be highlighted below. First, with an increasing time interval between cue and stimulus presentation (*cue-stimulus interval*, or CSI), switch costs are reduced but not fully eliminated (‘residual switch costs’; e.g., Altmann, 2004; Logan & Bundesen, 2003; Mayr & Kliegl, 2003). This observation casts a strict TSR view into doubt according to which sufficiently long CSI ought to allow individuals to reconfigure the task set in advance of stimulus onset and switch costs should disappear. By contrast, the interference view poses that task switching costs are driven by proactively acquired stimulus-response mappings, which could be evoked even following substantial CSIs (this account is also known as ‘task set inertia’). A further important observation relates to so-called *task congruency effects*. Most task-switching experiments use bivalent stimuli (i.e., the same set of stimuli serves as targets for both tasks). For example, in a colour-shape task,

participants classify coloured shapes either according to colour, or to shape. With a given set of stimulus-response mappings, half of the stimuli will be congruent (both tasks/dimensions generate the same response) whereas the other will be incongruent (the two dimensions generate different responses). Congruency effects (faster responses on congruent than on incongruent trials) have been demonstrated (e.g., Sudevan & Taylor, 1987) and might emerge exclusively on “switch” but not on “non-switch” trials (Goschke, 2000). Congruency effects in switch tasks probably arise from activated overlearned response category codes in long-term memory that are currently irrelevant (Meiran & Kessler, 2008), hence reflecting difficulty in filtering out currently relevant from irrelevant information (Meiran, 2010). A further relevant observation is that in switch tasks, *repeated responses* on consecutive trials lead to faster responses on non-switch trials but to slower responses on switch trials (e.g., Rogers & Monsell, 1995). Mayr and Bryck (2005) argued that this pattern reflects the fact that stimulus and response become associated with the task on trial N-1. On trial N, the association is retrieved, and when the task has changed, the retrieved information is inconsistent, turning the benefit from repeated responses (on non-switch trials) into a cost (on switch trials). A final set of relevant results reports an interaction between *congruency on trial N-1* and *task switch costs*, with larger switch costs following incongruent than following congruent trials (Goschke, 2000). The presence of task congruency effects on a given trial (see above) presumably suggests a conflict between two incompatible responses. Because a successful response is nonetheless made in most cases, it is likely that the task-irrelevant dimension has been inhibited, or decoupled from its response. Inhibition hence predicts that on a trial following this incongruency, task switch costs should be increased, relative to a situation in which trial N-1 was congruent, because more efforts are made to recover the inhibition in the preceding incongruent trial. This is indeed the case (e.g., Brown et al., 2007; Goschke, 2000).

Overall, all four observations (long CSIs being associated with residual switch costs; congruency effects; interaction between response repetition and task switching; interaction between response

repetition and task switching) highlight the contribution of stimulus-driven, learned stimulus-response associations in task switching experiments.

Measuring decision making and response execution via “mouse tracking”

In conventional psychological experiments, responses are often made via key presses, rendering a single data point per trial. Over the last few years, a range of methods have been developed in which participants provide responses by dynamic movements carried out via a reaching response to a target, or via responses made on digital tablets or with a computer mouse (see Wirth et al., 2020, for a recent overview). Compared to the results from key press experiments, these methods generate a much richer, and potentially more informative, picture of the decision making process. In such tasks, various characteristics of the individual response can be explored, such as its initiation time, the movement duration, the response latency, as well as the curvature of the trajectory. The method used in the work reported below is referred to as ‘mouse tracking’ (e.g., Freeman et al., 2011; Spivey et al., 2009; Freeman & Ambady, 2010): responses are made by participants dragging the cursor from a start area (typically at the bottom of the screen) to one of two response fields (typically in the left and right upper corner of the screen). Numerous studies have reported experimental manipulations which emerge in the dynamic response made to the target, spanning a wide field of psychological domains such as social psychology (e.g., Faust et al., 2019), development (e.g., Krueger & Storkel, 2020), and psycholinguistics (e.g., Tomlinson et al., 2013). A particular point of interest is the curvature of trajectories which is characterised as the deviation relative to a straight line from start to response (Freeman & Ambady, 2010). Curvature is often characterised by the *area under the curve* (AUC, the geometric area between the response and the straight line), or by *maximum deviation* (MD, the largest perpendicular deviation between the trajectory and the line). Typically, experimental manipulations emerge powerfully in curvatures. For instance, in ‘conflict tasks’ such as Simon in which compatibility between stimuli and responses is manipulated, incompatibility generates an ‘attraction’ of the response trajectory toward

the incorrect response (Wirth et al., 2020). On a broad level, the inference that decision making and response execution are intertwined is in line with modern cognitive theories which conceive of processing as a continuous, graded state in which mind, body and environment dynamically interact (e.g., Spivey & Dale, 2004).

The relevance of mouse tracking to task switching is twofold. According to a strict reading of the TSR view, decision making and response execution involved in task switching could be characterised as staged. For instance, in Kieras et al.'s (2000) EPIC model, a rule-based "central cognitive processor" establishes and reconfigures task sets, and once completed, the output is set to "mechanical effectors". By contrast, the interference resolution notion of task switching is more in line with a "continuous" or "graded" relation between decision and action. For instance, in Gilbert and Shallice's (2002) interactive activation model of task switching in the colour-word Stroop task, separate colour-naming and word-reading pathways compete for output, and residual activation from previous trials successfully accounts for the various observations made in the task switch literature. Hence task switch effects emerging in the curvature of response trajectories would broadly support interference accounts over TSR views, or at minimum, speak against a strictly serial interpretation of TSR.

A second relevant aspect is more specific. A range of design decisions influence behaviour in tasks with dynamic responses (Schoemann et al., 2019; Wirth et al., 2020). In its simplest form, a participant in a mouse tracking experiment initiates a trial by clicking on a "start" button at the bottom of the screen, and the stimulus appears immediately thereafter or after a short delay. Such a *static* starting procedure contrasts with a *dynamic* one in which the stimulus only appears once they movement has begun (see Scherbaum & Kieslich, 2018, for a comparison between the characteristics between static and dynamic procedures). With a static starting procedure and a simple task, response movements are normally initiated relatively soon after the stimulus display appears, with typical average initiation times between 150 and 300 ms. Such swift movement initiation makes it unlikely that the participant had fully

completed their decision when the response began. This predicts that stimulus-driven experimental manipulations (such as 'conflict' manipulations in Flanker, Simon and similar tasks) which emerge in errors, response latencies, and curvatures of movements should not be present in movement initiation times. Relevant evidence comes from Experiment 1 reported by Wirth et al. (2020) in which in a Simon task, responses made with a computer mouse were initiated quickly (average of 174 ms) and initiation times were not affected by congruency (by contrast, when responses were made on a touchscreen of a tablet computer, initiation times were much slower, with an average of 380 ms, and a congruency effect emerged). The finding concerning responses carried out with a computer mouse converges with our own experimental findings using a range of 'conflict' tasks: we generally found fast average initiation times which were unaffected by congruency manipulations on a given trial (e.g., Ye & Damian, under revision). Critically, however, other experimental effects do not depend on stimulus processing but rather arise from other aspects of the task. These could for instance be carry-over effects from previous trials, such as in 'congruency sequence effects' where incongruency on a previous trial affects processing on a given trial (e.g., Braem et al., 2014). Effects of this type are likely to emerge in movement initiation times (in addition to the other dependent measures such as errors, latencies, and curvature). Movement initiation times in mouse tracking studies with a static starting procedure therefore potentially allow to disentangle stimulus-driven from task-driven experimental effects (note that with a dynamic starting procedure, initiation times are by definition not sensitive to stimulus-driven effects because the movement is initiated before the stimulus appears).

The relevance for theoretical accounts of task switching is as follows. As summarised above, a number of experimental variables have been highlighted in the literature on task switching, and a subset of these are stimulus/response dependent (i.e., congruency; response on trial N-1) whereas others are not (CSI; congruency on trial N-1). All variables and their interplay with task switch costs are expected to emerge in latencies, as well as potentially in measures of response movement such as curvatures.

However, we predict that stimulus/response dependent variables will not affect movement initiation times, whereas task/cue dependent variables will. The critical question is, does task switching affect initiation times? If so, then the inference would be that a theoretical account which explains task switch costs as purely stimulus/response driven cannot be entirely correct. Instead, switch costs would arise from classic TSR, or a combination of TSR and interference.

To our knowledge, only two previous studies have explored task switching with the mouse tracking technique, and they have reported somewhat contradictory results. Hindy and Spivey (2008) employed a colour-shape switch task (adopted from Diamond & Kirkham, 2005) in which participants were cued at the beginning of each trial to either attend to colour, or to shape. Upon initiating the trial, three coloured objects were presented simultaneously, with a stimulus object at the bottom of the screen, and two target objects at the upper left and right corners. Participants were instructed to move the cursor toward, and click on, the target object which matched the stimulus object according to the cued dimension. The results showed a complex pattern, with responses to the colour dimension overall faster and less curved than responses to the shape dimension, and switch costs emerging in initiation times but not in curvatures for colour decisions, but the opposite pattern for shape decisions. Overall, these results were interpreted as being incompatible with a strictly top-down model of task switching according to which movement characteristics should not be affected at all by task switching, and more in line with task switching arising from competition between multiple sensory and response pathways. Regarding our specific predictions regarding task switching and movement initiation times outlined in the previous passage, Hindy and Spivey's result showed relatively fast initiation times (approximately 200 ms) but inconclusive results regarding whether these are affected by switching.

Weaver and Arrington (2013) conducted a switch task in which participants judged either parity (odd/even) or numerical magnitude (large/small) of digits 1-9 (excluding 5), using two CSIs of 200 and 1,000 ms. They also explored possible congruency effects between the two decision dimensions. CSI,

switch and congruency effects appeared in movement curvatures. Initiation times were much slower than in Hindy and Spivey's (2008) experiment (481 ms) and these were affected by task switch and congruency, but surprisingly not by CSI. A second experiment used two sets of cues for each task dimension in order to address a confound between 'cue repetition' and 'task repetition' previously highlighted by Logan and Bundesen (2003). Results showed that it was indeed mainly the repetition of the cue, rather than of task, which affected response latencies and movement characteristics, as well as initiation times. For this experiment, the authors reported no descriptive statistics, so it is difficult to judge the relative speed of initiation times, as well as how they were affected by CSI, switching, and congruency.

In summary, the two previous studies in which task switching was explored via mouse tracking offer a complex and inconclusive pattern of results. On a broad level, it is likely that movement trajectories are affected by task switching, although Weaver and Arrington's (2013) results certainly raise the possibility that at least part of this was due to cue rather than task repetition. This is an issue which could potentially also have affected Hindy and Spivey's (2008) results because they used just a single set of cues. Concerning movement initiation times, results again are complex. Hindy and Spivey reported fast initiation times which numerically exhibited task switch effects, but significantly so only for colour but not for shape decisions. Weaver and Arrington reported much slower initiation times which were affected by task switch, but this arose from cue rather than task repetition (they also reported, contrary to our predictions, congruency effects in initiation times which in our view mainly arose from the fact that initiation times were atypically slow).

The present study

In the study below, we designed a switch task with responses provided via computer mouse in which we attempted to address most of the shortcomings of the previous studies with this technique. We used a colour-shape switch task in a form which we take to be canonical in the literature (e.g., Prior

& MacWhinney, 2010). Following the task cue, a single coloured shape was presented as the target, and participants classified it according to either colour or shape (dependent on the cue), a procedure which is in line with numerous previous key-press experiments on task switching (e.g., Prior & MacWhinney, 2010). A response was carried out by clicking on one of two unlabelled response fields at the upper left and right corner of the screen. To avoid the potential confound between cue and task repetition outlined above, we used two alternating sets of cues, hence cues were never repeated on consecutive trials and the emerging switch cost is genuinely attributable to task switching. We used two relatively short CSIs (200 and 500 ms) which were chosen to maximise our chances of capturing task switching in mouse responses. With such short CSIs, task switching should still be strongly present even with the longer interval, whereas only with longer CSIs switch cost would be reduced to a residual degree.

In our analysis, we focused on the following dimensions of theoretical interest: a) mix costs (difference between pure and mixed blocks), b) switch costs and their dependence on CSI, c) interplay of switch costs and congruency effects, d) effects of response repetition from trial N-1 on switch costs on trial N, and d) effects of congruency on trial N-1 on task switch costs. To reiterate the main predictions: some experimental manipulations arise from stimulus aspects (i.e., congruency effects) or even response dimensions (i.e., response repetition between trial N-1 and N) and here it is likely that they should affect response latencies (and potentially errors) and possibly also the curvature of response movements. In line with previous mouse tracking studies which had used a 'static' starting procedure (e.g., Experiment 1 reported in Wirth et al., 2020) and also in parallel to the results of Hindy and Spivey (2008) but diverging from the results of Weaver and Arrington (2013), we predicted relatively fast initiation times. If so, movements are probably initiated too early to be affected by stimulus properties, and so we predict that stimulus-driven effects should not emerge in initiation times. By contrast, other effects do not depend on the stimulus display of a current trial (i.e., the effect of CSI) and here we expected them to affect initiation times as well. The critical issue was how switch costs manifest

themselves. The TSR view predicts that switch costs arise from task set reconfiguration by the cue on the current trial, and that they are likely to be already present at the point in time when a response movement begins. By contrast, the interference view characterises switch costs as stimulus-driven, and hence the prediction is that like congruency effects, they should not be present in movement initiation times.

Experiments with dynamic responses such as in mouse tracking generate data sets which are considerably more complex than those resulting from key-press experiments. Responses can be analysed along a plethora of potential dependent variables¹ (errors, initiation times, response latencies, various measures of curvature, velocity, acceleration, etc.; see Wirth et al., 2020, for a comprehensive overview), and presenting these results and analyses in an accessible format poses a considerable challenge. In the results reported below, we opted to report most of the key findings in figures which present on the one hand average trajectories, and on the other hand the key dependent variables of errors; initiation times; response latencies; plus MD as a measure of curvature (analyses using AUC rather than MD generally showed statistically equivalent results). The dependent variables are presented as inset plots with corresponding statistical information underneath. We expected the results to form a complex mix of positive and null findings and the latter are problematic to interpret within a frequentist framework, hence we based our analysis on Bayesian rather than frequentist statistics.

Method

Participants

¹ Note that in mouse tracking and related techniques, the various dependent measures are not necessarily independent from one another. For instance, Grage et al. (2019) recently showed a negative correlation between initiation times and curvature, as well as a negative relationship between initiation time and response latency.

40 undergraduate students (mean age: 23.8; female N=25, male N=13, prefer not to say N=2) from the University of Bristol were recruited for this experiment and received course credit. Participants gave informed consent for this study. All participants confirmed that they have normal or corrected to normal vision, were comfortable with operating a computer mouse with their right hand, and were not colour-blind. Participants were tested in small groups. Ethical approval for this study was granted by the Faculty of Science Research Ethics Committee at University of Bristol (no. 75221).

Materials and Procedure

In this experiment, participants judged coloured shapes by colour (red/blue) or shape (triangle: 169 x 144 pixels; 4.5 x 3.8cm; or circle: 145 x 154 pixels; 4.0 x 4.0 cm). The trial structure is displayed in Figure 1. Participants were instructed to initiate a trial by clicking on a “Start” button at the bottom of screen (hence, the *response-cue interval*, or RCI, was variable). A task cue appeared in the lower half of the screen (180 pixels below the centre of the screen). Two cue sets alternated for each task: colour was cued either by a stylised rainbow (179 x 69 pixels; 4.7 x 1.8 cm) or a colour wheel (129 x 129 pixels; 3.4 x 3.4 cm); shape was cued either by a horizontal display of a circle, square, and triangle (187 x 52 pixels; 5.0 x 1.4 cm), or a circle embedded within a triangle, itself embedded within a square (126 x 126 pixels; 3.3 x 3.3 cm). Cues remained on the screen according to the CSI (either 200 or 500ms), and were

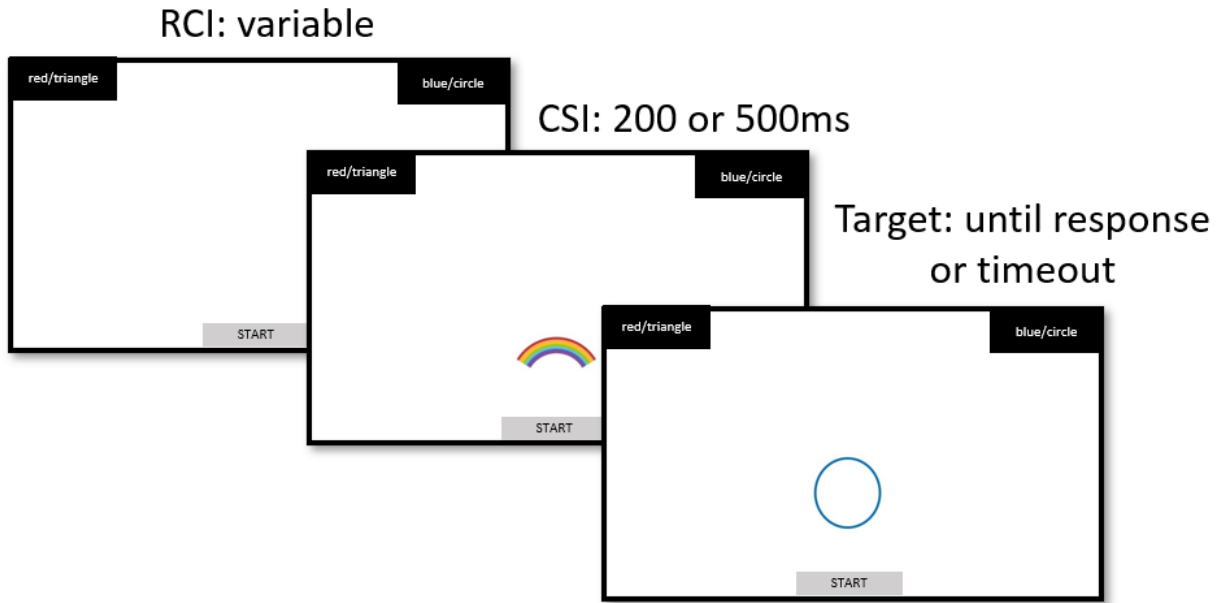


Figure 1. Trial structure. RCI = Response-cue interval; variable as each trial was initiated by the participants clicking on the “START” field. CSI = Cue-stimulus interval (either 200 or 500ms).

replaced with the target display (e.g., a blue circle) presented in the same location. Participants made their response by moving the mouse from the “Start” field to the appropriate response field and clicking on it. They were instructed to make the response as quickly and accurately as possible. Targets disappeared either when participants had completed a response, or after 2,000ms. We used the same stimulus-response mapping throughout (left response: red/triangle; right response: blue/circle).

The experimental session began with four pure blocks of 18 trials with two buffer trials at the beginning. The first two blocks involved only colour decisions, with Block 1 a practice block and Block 2 the experimental block. During the practice trials, the word *red* appeared in the top-left corner and *blue* in the top-right corner. Blocks 3 and 4 involved only shape decisions, again with the first a practice block and the second an experimental block. During the practice trials, the words *triangle* appeared in the top-left corner and *circle* in the top-right corner. Next, participants were instructed that from now on they would be cued as to whether to perform a colour or a shape decision on a given trial. There was one practice block of 32 trials, followed by three experimental blocks of 98 trials each, with two buffer trials

at the beginning of each block. Trials were pre-scrambled between colour and shape decisions and there was a 50% probability of a task change, resulting in the same number of switch and non-switch trials. No feedback on accuracy or latency was provided to participants during the experimental session.

Participants were seated approximately 60 cm from the computer screen (23 inch Dell P2319H flat screen monitor with screen resolution 1920×1080 ; Dell MS116 computer mouse). MouseTracker software (Freeman & Ambady, 2010) was used for presentation of the stimuli. Participants were instructed to click on a grey box (192×72 pixels; 5.1×2.9 cm) in the bottom centre of the screen to initiate a trial, and following the display of the target stimulus, they responded by moving their mouse to, and clicking on, one of two response fields (288×144 pixels; 7.6×3.8 cm) on either the top left or the top right corner of the screen. MouseTracker collected the raw data of each mouse trajectory, recording x and y coordinates of the trajectory of the mouse movement every 16 ms. The 'cursor speed' parameter which overrides the Windows settings and ranges from 1 (slowest) to 20 (fastest) was set to 12.

Data preprocessing and statistical analysis. Data were processed in *R* (R Core team, 2020) using the package *mousetrap* (Kieslich et al., 2019). For each trial, the response accuracy, initiation time (the time at which a participant initiated the mouse movement, measured relative to the onset of the target display), response latency (the time interval between onset of the target display, and clicking on the response field) and MD (measured in cm) were computed. Because mouse clicks on the "start" button oftentimes are associated with miniscule movements which are then erroneously recorded as extremely short initiation times, we instead computed initiation time as the first time sample relative to target onset in which the mouse cursor left a specific starting region. With a rectangular start region, initiation times are potentially confounded with the starting angle of a movement (Wirth et al., 2020) and hence a circular starting region is preferable. We followed this recommendation and computed initiation times as the first time sample at which movements left a virtual circular region with a diameter equal to the

height of the response box (72 pixels). Measures were then aggregated within and across participants. Trials on which participants had made an error were excluded from the analysis of the other dependent variables. We further excluded data from trials on which participants had made no response, with latencies faster than 500ms, as well as from trials with initiation times longer than 1 sec (3.9%).

In the following we report the outcomes of Bayesian statistical analysis, based on the packages *BayesFactor* (Morey & Rouder, 2018) and *bayestestR* (Makowski et al., 2019). Because the outcome of Bayesian tests varies slightly across repeated analyses, we conducted each reported analysis ten times, and report average Bayes factors (the range of the obtained Bayes factors can be found in the supplementary materials).² Raw data and R analysis files are available on the [Open Science Framework](#).

Results

Pure vs. Mixed blocks. In a first analysis, measures were computed for the pure blocks (either colour decisions only, or shape decisions only) compared to the mixed block (colour and shape decisions randomly intermixed; for this comparison only non-switch trials were used, in line with the previous literature, e. g., Paap et al., 2017; Rubin & Meiran, 2005). The difference in performance between pure and mixed blocks constitutes the *mix cost*. Compared to responses in the pure condition, responses on mixed trials were more error prone (by 4.2%; $BF_{10} = 322.55$), took longer to initiate (by 101ms; $BF_{10} > 1,000$), as well as slower response latencies (by 354ms; $BF_{10} > 1,000$), and had more curved trajectories, as indicated by MD (by 3.1cm; $BF_{10} > 1,000$). Hence, randomly intermixing two tasks incurred the predicted mix cost on all dependent measures. See Supplementary material A for descriptive statistics.

² Throughout, we used Lee and Wagenmakers' (2013) heuristic that a Bayes Factor of 1-3 constitutes "anecdotal" evidence, 3-10 is "moderate", 10-30 is "strong", 30-100 is "very strong" evidence, and > 100 is "extreme" evidence for H_1 . Conversely, a BF of 1/3-1 constitutes "anecdotal" evidence, 1/10-1/3 is "moderate", and $< 1/10$ is "strong" evidence, etc. for H_0 . Clearly these should only be interpreted with caution, and not be used to dichotomise results into present vs. absent.

Switch costs and CSI. From here on, analysis was restricted to the mixed block. Next, we considered the impact of task switching on performance in mouse tracking, in conjunction with exploring the effects of CSI. Switch costs tend to be more pronounced with shorter than with longer CSIs (e.g., Rubin & Meiran, 2005). Our own study used two different CSIs. Trials were coded as to whether a task on trial N (colour or shape decision) was the same as the task on trial N-1 (“non-switch”), or different (“switch”), as well as whether the CSI was 200 or 500 ms. Figure 2 shows the results. Rightward trajectories were flipped to appear as pointing toward the left response. Average trajectories³ on the left show clear signs of a switch cost (orange/switch average trajectories are more curved than blue/non-switch trajectories) but only a very minor effect of CSI. The inset panels show strong Bayesian evidence for an effect of task switch on all dependent variables. This includes the particularly theory-relevant variable of initiation times: these are 17 ms slower on switch than on non-switch trials. CSI clearly affected initiation times and latencies, but the effect on curvature was inconclusive; the visual impression from the figure is certainly that if there is an effect on curvature, it was relatively minor. Finally, the evidence regarding an interaction between CSI and task switch costs was not conclusive, but again visual inspection of the figure suggests that an interaction, if genuine, must be quite subtle. We conclude that switch costs emerge clearly in properties of response movements, whereas compared to the short CSI, the longer CSI accelerates the movement but does not substantially affect its shape. See Supplementary material B for descriptive statistics.

Switch costs and Congruency. Switch costs are potentially affected by congruency between the two task instruction sets, with larger congruency effects on switch trials, and less or no congruency effects on

³ One of the reviewers alerted us to the possibility that with average movement duration times of 973 ms, it is possible that participants responded in a discontinuous manner (e.g., participants might stop or slow down after they initiated a movement, work out the response, and then re-initiate the movement). To explore this possibility, we plotted average X and Y movements, velocity and cumulative distance from start position over the first 1 sec of each trial. From this figure (see Supplementary material C) we found no evidence for a discontinuity in the movements.

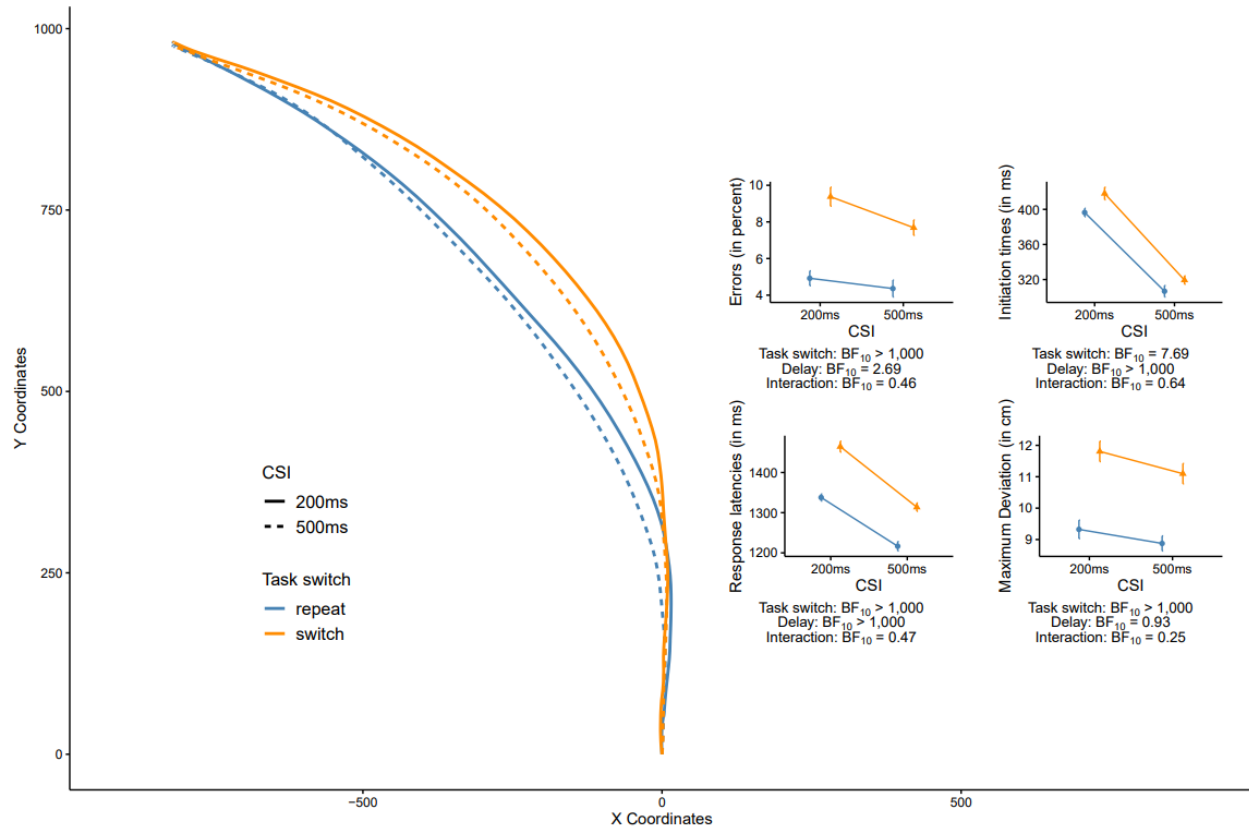


Figure 2. Task switch (repeat vs. switch) and CSI (200 vs 500 ms). Left side: average time-standardised movement trajectories. Right side: dependent variables (errors, initiation times, response latencies, and MD). Error bars in inset panels show within-subjects standard errors using the Cousineau-Morey-O'Brien (Cousineau & O'Brien, 2014) method.

non-switch trials (e.g., Goschke, 2000; Sudevan & Taylor, 1987). For this analysis, trials were additionally coded regarding whether on a given trial N , the irrelevant task generated the same response as the relevant one (congruent) or not (incongruent). Figure 3 shows the results. Average trajectories on the left show clear signs of a switch cost and a congruency effect (dashed/incongruent lines are more curved than solid/congruent lines), as well as an interaction between switch costs and congruency, with the most curved trajectory in the switch/incongruent combination. The inset panels show Bayesian evidence for an effect of task switch, of congruency, as well as of a switch x congruency interaction, on all measures except for initiation times. In summary, the results show the interaction between switch costs and congruency previously reported in the literature which manifests itself in errors, response latencies,

and movement curvature. The exception is in initiation times: task switch clearly affects initiation times, whereas congruency does not. See Supplementary material D for descriptive statistics.

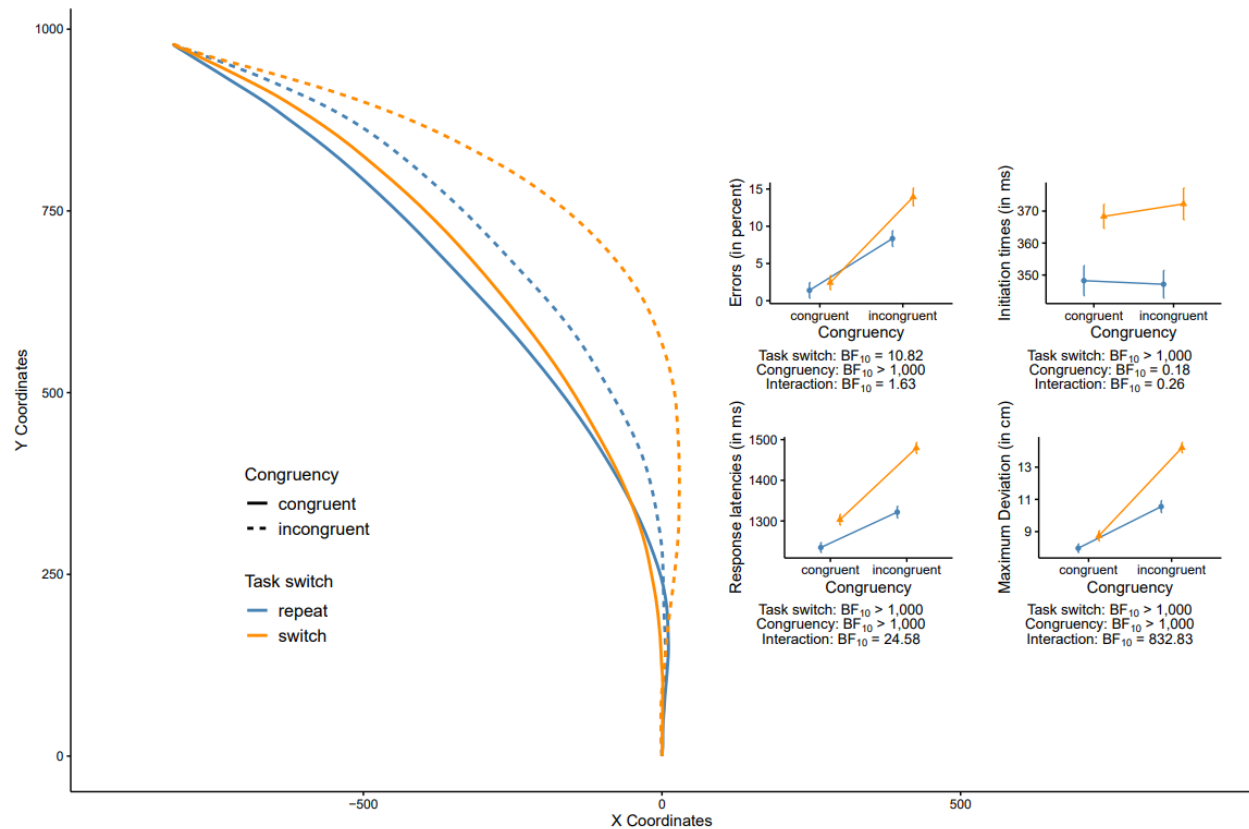


Figure 3. Task switch (repeat vs. switch) and Congruency (congruent vs. incongruent) effects. Error bars in inset panels show within-subjects standard errors using the Cousineau-Morey-O'Brien (Cousineau & O'Brien, 2014) method.

Switch costs and response repetition. In switch tasks, repeated responses on consecutive trials lead to faster responses on non-switch trials but to slower responses on switch trials (e.g., Rogers & Monsell, 1995). To explore this pattern, trials were coded into whether on a given trial N, the response (left or right) was the same as on the preceding trial. Figure 4 shows the results. Trajectories showed the predicted response repetition benefit on non-switch trials but a repetition cost on switch trials. An interaction between task switch and response repetition emerged strongly in response latencies and movement curvature (but not in errors). Initiation times were not affected by response repetition,

whereas the interaction with task switch was inconclusive. See Supplementary material E for descriptive statistics.

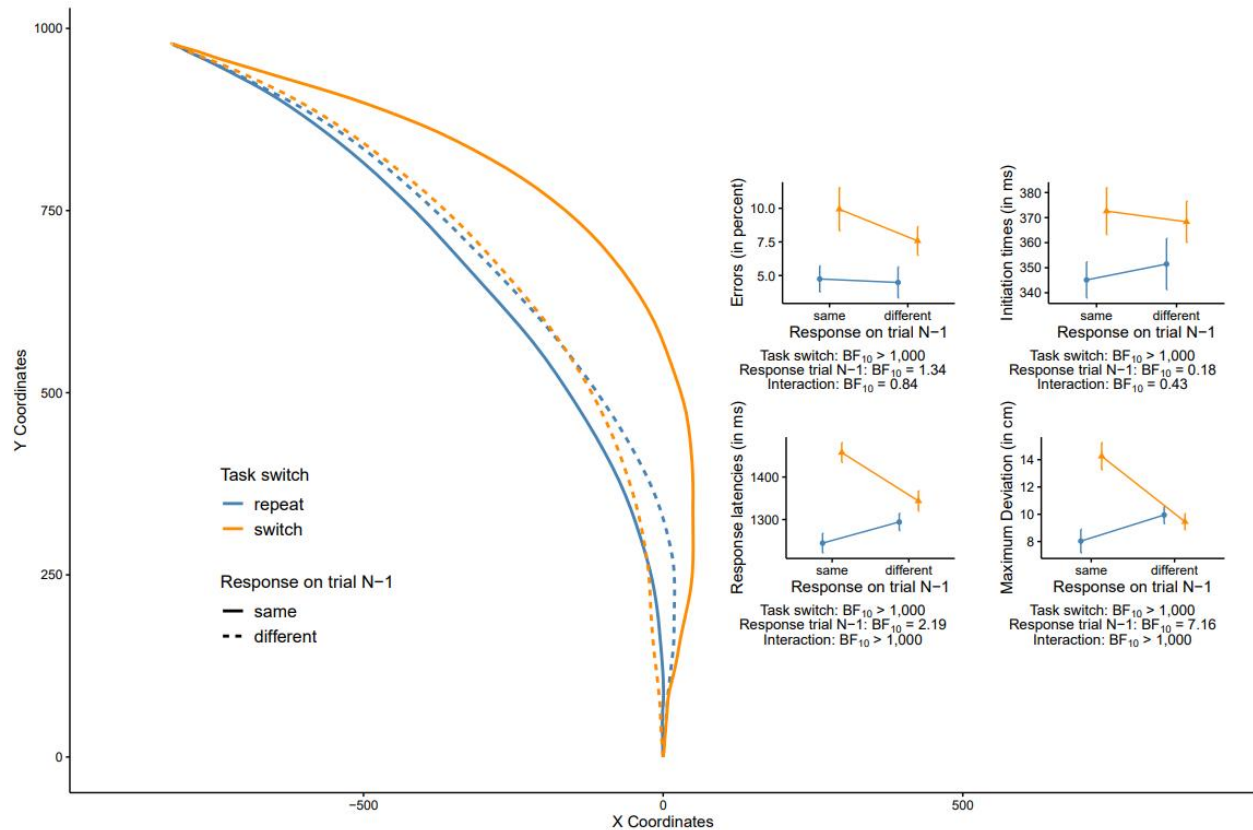


Figure 4. Task switch effects (repeat vs. switch) dependent on Response on trial N-1 (responses on trial N-1 and N are “same” or “different”). Error bars in inset panels show within-subjects standard errors using the Cousineau-Morey-O'Brien (Cousineau & O'Brien, 2014) method.

Switch costs dependent on congruency on trial N-1. Switch costs are potentially more pronounced following incongruent than congruent trials (Goschke, 2000). To explore this pattern, trials were coded into whether trial N-1 was congruent or incongruent. Figure 5 shows the results. The predicted interaction was found in errors, response latencies, and movement curvature. The impact of congruency on trial N-1 on initiation times, as well as its interaction with task switch, was less conclusive. See Supplementary material F for descriptive statistics.

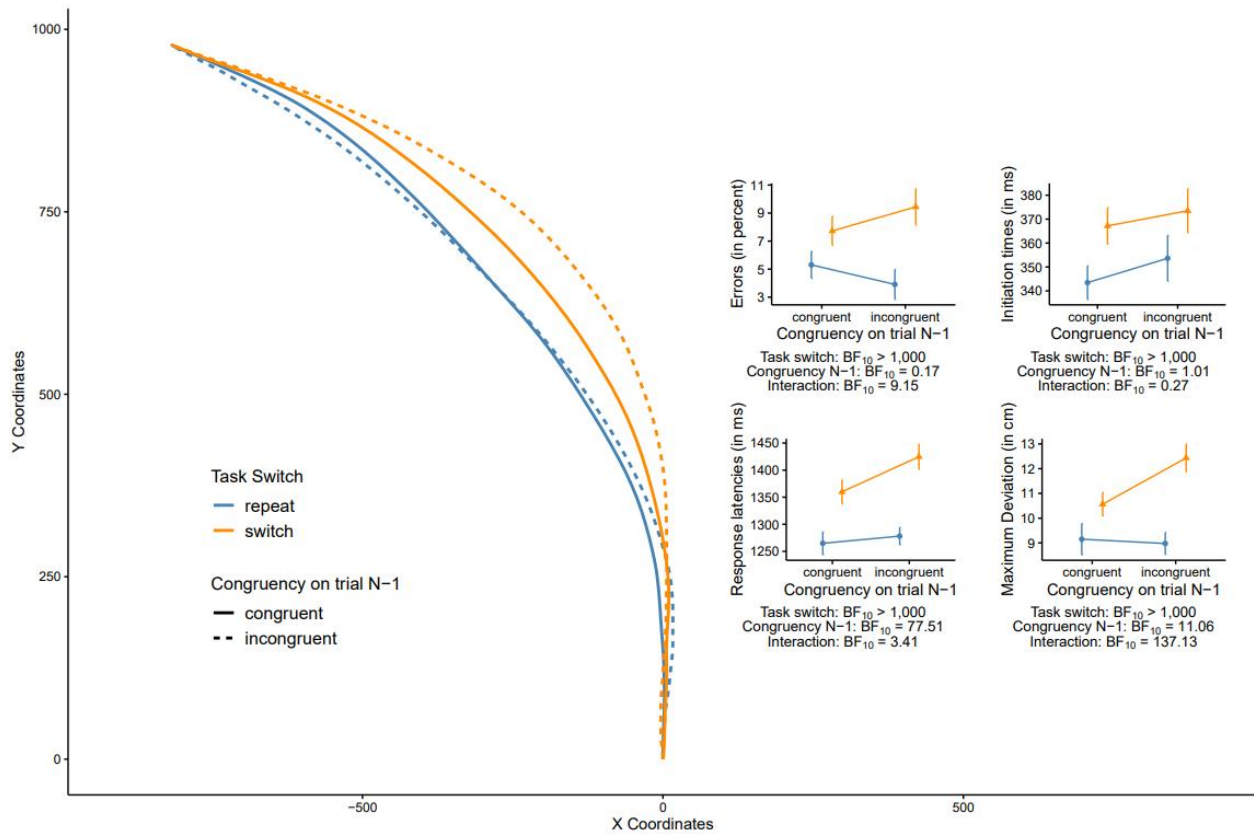


Figure 5. Task switch effects (repeat vs. switch) dependent on Congruency on trial N-1 (congruent vs. incongruent). Error bars in inset panels show within-subjects standard errors using the Cousineau-Morey-O'Brien (Cousineau & O'Brien, 2014) method.

Time course analysis across all relevant variables. All theory-relevant variables were combined in a rudimentary time course analysis. The logic is as follows: some of the variables are not in any obvious way stimulus/target-dependent, and so should exert a relatively early effect on response trajectories. For instance, effects of CSI do not depend on target characteristics, and so should affect responses relatively early on. By contrast, congruency clearly depends on the specific stimulus presented on trial N, and response repetition from trial N-1 depends on the specific response to the stimulus on trial N and its interplay with the response on trial N-1. These variables should affect response trajectories at a relatively late time point. The important issue regards the onset of the task switch variable: does it cluster with the stimulus-dependent or the stimulus-independent variables?

We performed a time course analysis for each variable in which x -coordinates at each sampled time step were aggregated for the first 1,000ms of responses, and the deviations from the zero midpoint, dependent on the condition, were tested via Bayesian t -tests.⁴ Figure 6 shows the results, with the y -axis showing the logarithm of BF_{10} . Four of the five variables rise to strong Bayesian evidence in the time window of 350-500 ms post stimulus onset, with the remaining variable (response N-1) showing a much later rise. CSI shows the earliest rise, with task switch and congruency N-1 next, followed by congruency, and response N-1 last. A possible inference from this finding (unfolded in more detail in the discussion) is that both cue- and stimulus-driven aspects of the task combine to generate the classic switch cost.

⁴ We performed this analysis on raw rather than time-normalised trajectories because we felt that raw trajectories would be better suited to capture the relative onsets of experimental variables. However, raw trajectories have variable end points. In the analysis reported here, 98% of trajectories had end points larger than 700 ms, 91% had end points > 800 ms, and 82% had end points > 900 ms. Hence there is a certain degree of data loss due to 'inactive' data points toward the right side of the figure. At the same time, the critical findings regarding the onset of the various effects come from virtually all trajectories included in the data set.

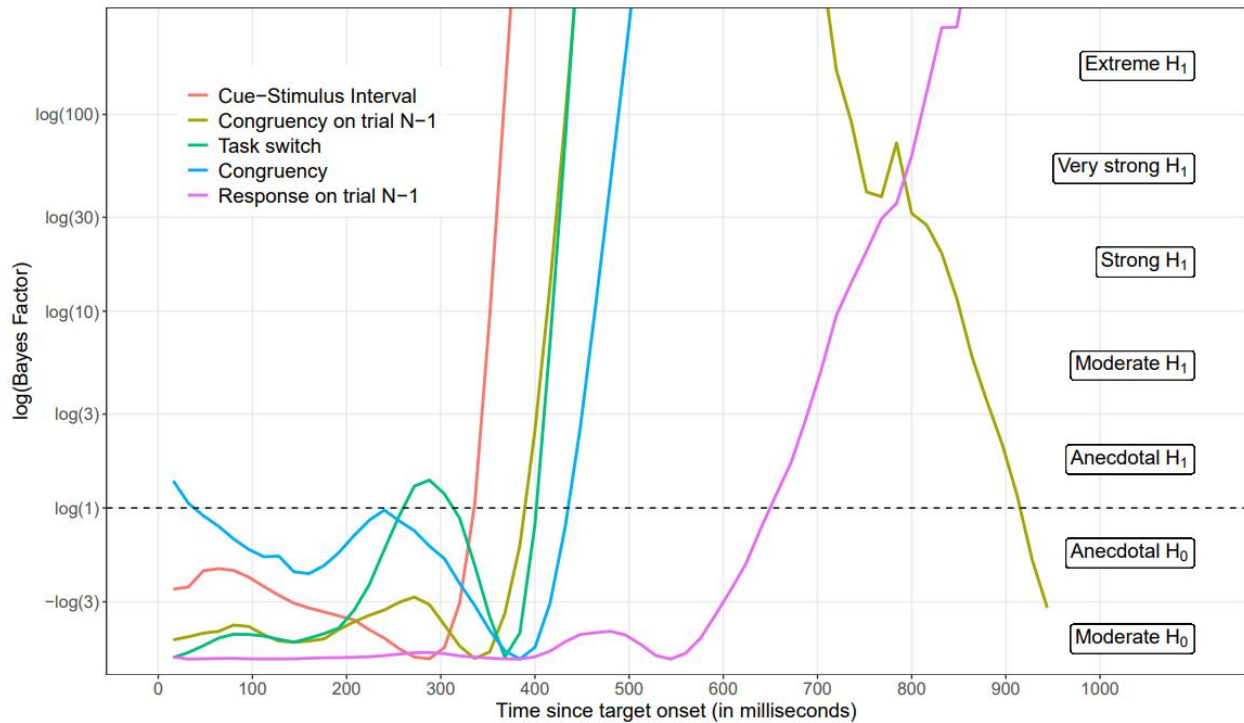


Figure 6. Time course analysis (first 1,000ms of a trial post stimulus onset) of five variables of interest. See text for details.

Discussion

The current study explored task switching in a classic colour-shape decision experiment: participants were randomly cued on each trial to classify coloured shapes (red/blue circles/triangles) either according to colour or to shape. Contrary to numerous prior studies in which a similar manipulation had been used, in our study participants gave their responses via dynamic computer mouse movements, which allowed us to explore four dependent measures: errors, the time at which a movement was begun (initiation times), response latencies (the time it took to click on the response field), and a measure of curvature of the response trajectory (MD). The findings can be briefly summarised as follows.

Both a mix cost (difference between blocks in which participants randomly switched between two tasks, and blocks in which they only carried out a single task) and a switch cost (difference between

trials on which the task from the previous trial was repeated or switched) emerged in all dependent measures. CSI mainly manifested itself as a shift of the entire response in time (with slower initiation times and latencies for the short than for the long interval) but with only minor effects on the shape of the trajectory. A task congruency effect (on a given trial, temporarily irrelevant task instructions interfered with the relevant ones) was found in all measures other than initiation times, and congruency interacted strongly with switch costs. Furthermore, we found response repetition effects (i.e., whether or not on a given trial N , the response was the same or different than on the preceding trial $N-1$) and these interacted with the task switch effect. Finally, we found an effect of congruency on trial $N-1$: for incongruent $N-1$ trials, greater switch costs were found on errors, latencies, and trajectory curvatures than for congruent $N-1$ trials.

A subset of these findings had been reported in previous studies. Hindy and Spivey (2008) used a colour-shape switch task (although in a slightly different form from ours; see below), and reported statistically significant switch costs in initiation times, response latencies, and trajectory curvatures as assessed by MD. Weaver and Arrington (2013) conducted a number switch task and showed both switch and congruency effects on accuracy, initiation times, and response latencies, but a significant switch \times congruency interaction emerging only in errors. Both switch and congruency effects appeared in movement curvatures, with a marginally significant switch \times congruency interaction. To our knowledge, potential effects of response repetition, and of congruency on trial $N-1$, have not previously been explored in task switch studies with the mouse tracking methodology.

We consider movement initiation times to be the most informative dependent variable considering theoretical explanations of task switching. As highlighted in the introduction, our rationale was that with a static starting procedure, response movements should be initiated relatively soon following stimulus presentation. This was also the case here, with average initiation times of 360 ms. These are slower than those in Hindy and Spivey (2008; approx. 200 ms) but faster than in Weaver and Arrington (2013; 481

ms). However, we note that average initiation times in our study are still longer than in previous studies using 'conflict' tasks, e.g. Wirth et al., (2020, Experiment 1) reported average initiation times of 174 ms in a Simon task. We suspect that this arises from the fact that task switching is cognitively more challenging than simple judgements such as those in standard conflict tasks.

Our prediction was that stimulus-driven experimental effects such as those of congruency or of response repetition should not affect initiation times, presumably because the movement is initiated before the decision has been completed. This was indeed the case, with little evidence that congruency or response N-1 affected initiation times. By contrast, task- or cue-driven effects, such as those arising from CSI, or from incongruency on the previous trial, were predicted to affect initiation times. This was clearly the case for CSI, but the evidence regarding congruency N-1 was inconclusive. Our critical variable of interest was task switch, and here it was clearly the case that initiation times were affected. This finding suggests that switch costs cannot be fully accounted for by 'interference accounts' but that part of it must arise from task set reconfiguration which takes place prior to stimulus/response processing. The powerful presence of stimulus/response-driven effects such as congruency of course underscores the claim that performance in switch tasks is substantially affected by sensitivity to learned associations. Hence, we suggest that the most plausible account of task switching rests on a contribution of both endogenous and exogenous control processes (e.g., Vandierendonck et al., 2010).

The time course analysis of all combined variables presented in Figure 6 is admittedly rudimentary and more sophisticated methods have recently been developed to analyse temporal aspects of mouse tracking data (e.g., Scherbaum & Dshemuchadse, 2020). Nonetheless, our findings regarding the relative temporal onset of effects are certainly suggestive. The results show that CSI is the first to affect response movements, followed by congruency N-1 and task switch, then congruency, and finally response N-1. Hence, cue/task driven effects emerge first (CSI), and stimulus/response driven effects later (congruency and response N-1 respectively), with task switch emerging in between. In accordance

with our interpretation of the initiation time patterns, this suggests that switch cost is neither entirely cue- nor stimulus-dependent, but probably arises as a combination of endogenous (top-down) and exogenous (bottom-up) processes.

As summarised in the introduction, two previous studies have combined task switching with the measurement of computer mouse trajectories, and it is important to examine to what extent the current results converge with the previous findings, and where they differ. Hindy and Spivey (2008) used a colour-shape switch task (although in a slightly different form from ours; see below), and reported statistically significant switch costs in initiation times, response latencies, and trajectory curvatures. They also reported that the task switch effect depended on the task (colour vs shape): for colour decisions, initiation times were significantly affected by task switch but response curvature was not; for shape decisions, initiation times were unaffected by task switch but curvatures were. In a subsidiary analysis of our own data, we also explored potential effects of task (colour vs. shape) but found only negligible differences. It is worth noting that Hindy and Spivey possibly overinterpreted a constellation of significant and non-significant results: for colour trials, the effect of tasks switch on initiation times was just-significant; $p = .049$, whereas the corresponding comparison for shape trials failed to reach conventional significance; $p = .23$ (no test for a statistical interaction was reported). Corresponding Bayes Factors are 0.90 and 0.37 respectively, so neither offers clear evidence for the presence or absence of an effect on initiation times. By contrast, in our own study task switch powerfully affected initiation times ($BF_{10} > 1,000$).

Hindy and Spivey's (2008) experiment differed from ours in a number of other experimental properties. First, their study adopted a procedure from Diamond and Kirkham (2005) in which participants were cued as to whether to focus on colour or on shape, then three coloured shapes were presented in an inverted triangle shape, and participants clicked on the shape in the top row which matched the bottom one. By contrast, in our task following the task cue, a single bivalent stimulus (e.g.,

a blue triangle) was presented and participants clicked on one of the two response fields at the top of the screen. Perhaps these procedural details affected potential task differences between colour and shape decisions. Second, in their study, the delay between presentation of the cue (the words *colour* or *shape*) and of the stimulus (the coloured shape) was determined by when participants chose to click on a “start” button at the bottom of the screen. A large body of literature exists on the effects of CSI in task switching (see introduction) and using a self-determined and hence variable CSI renders interpretation of the results more difficult (e.g., perhaps participants chose different CSIs for colours and for shapes). Third, only a single set of cues was used, which implies that for the non-switch trials, task repetition is conflated with cue repetition (Logan & Bundesen, 2003; Mayr & Kliegl, 2003). To avoid this problem, in our experiment we used two sets of cues and hence cues were never repeated on consecutive trials. Overall, further research is needed to evaluate Hindy and Spivey’s suggestion that for some tasks, responses are ‘graded’ and hence task switch costs are absent in initiation times but emerge in curvatures, whereas for other task, responses are ‘staged’ and hence emerge in initiation times but not in curvatures. In our own study, Bayesian analysis suggested that task switch powerfully affected both initiation times and curvatures of responses.

Our findings also did not fully replicate those of Weaver and Arrington (2013; see introduction) in which mouse tracking was combined with a numerical size-parity task switching experiment. Most importantly, they reported relatively slow initiation times which were affected by task switch but also by congruency (the effect of CSI was less consistent). The positive finding regarding congruency is particularly surprising to us, but it is compatible with the observation that responses were initiated slowly. It is, however, not clear why this is the case. Response latencies were certainly in line with those found in our own study. We suspect that some as yet unidentified aspect of the procedure or instructions induced more ‘conservative’ responses in participants than in our own study. Further worth highlighting is our finding that we found clear task switch effects on movement characteristics despite

us avoiding the potential confound with of task and cue repetition; by contrast, Weaver and Arrington attributed the bulk of their effects to cue rather than task switching. Despite these differences between the earlier and our own findings, we broadly concur with Weaver and Arrington's inference that their results likely reflect a combination of TSR and stimulus-response interference. It is worth noting that Weaver and Arrington used a different task from ours (size-parity vs. colour-shape), and task switching performance can to some extent be task-specific (e.g. Yehene & Meiran, 2007).

In our study, we used a static starting procedure (cue and target were shown after participants initiated a trial by clicking on a Start region at the bottom of the screen, and participants chose when to initiate the response toward the target). This contrasts with a dynamic procedure in which a stimulus is only shown once a response movement is already in motion. For the current research topic, we argued that use of the static procedure is essential because with a dynamic procedure by definition, movements are initiated before a stimulus is presented, hence making it impossible to disentangle stimulus/response from cue/task driven effects. However, it is acknowledged that only with a dynamic procedure do cognitive effects take place in reasonably close temporal proximity to the measured movement trajectories and for this reason, the dynamic procedure is normally preferable. Hence, our time course analysis reported in Figure 6 potentially allows us to determine the sequence in which a range of effects affect movements, but these provide little time about the underlying time course of cognitive effects. If the aim is to obtain true 'online' measures of the temporal neurodynamics of cognitive processing, then electrophysiological methods are much preferable (e.g., Cutini et al., 2021).

Mouse tracking as a window into cognitive control

Mouse tracking, alongside similar methods such as reach tracking, provides a powerful tool for investigating cognitive control, and task switching in particular: all important findings reported in the literature (e.g., mix costs, switch costs, congruency effects etc.) came out very clearly in trajectories, perhaps more so than in RTs. Mouse tracking experiments typically involve some sort of spatial stimulus-

response (S-R) compatibility manipulation. For instance, Freeman and Ambady (2009) presented participants with either female or male human faces which were either sex-typical or -atypical, and for the atypical faces they found a continuous 'spatial attraction' toward the response associated with the opposite-gender stereotype. Dale and colleagues (2007) asked participants to categorise pictures or words into two semantic categories (e.g., mammal or fish), and found that responses to atypical exemplars (e.g., *whale*) were curved toward the incorrect response more so than responses to typical exemplars (e.g. *cat*). Barca and Pezzulo (2012) reported a lexical decision task with computer mouse responses in which low frequency words were curved more toward a nonword response than high frequency words. In studies such as these, the manipulation is generally such that the curved stimuli induce a tendency to respond with the incorrect response (a tendency which is typically suppressed, as participants eventually click on the correct answer). In our own experiment, the effect of congruency is clearly of this nature as well: the temporarily irrelevant task instructions interfere with the currently relevant ones, and if they are incongruent, then responses are attracted toward the incorrect answer, resulting in the signature curvature in average trajectories.

Interestingly, however, the classic switch cost which emerged in our response trajectories is of a different kind: responses on switch trials are more curved than responses on non-switch trials, and in fact, congruency and task switch effects emerge visibly in a quite comparable fashion in the trajectories (cf Figure 3). But for the switch/non-switch contrast, the difference in trajectories cannot be due to a spatial incongruency (as is the case for the congruency manipulation) and hence a mechanism different from spatial S-R compatibility must underlie the effect. We suggest that the switch cost emerging in response trajectories further highlights the fact that decision making and response execution are strongly intertwined. Task switching incurs a cognitive cost which not only delays responses in conventional key press experiments, but it also negatively affects the efficiency of a response when dynamically carried out via computer mouse movements. Hence, mouse trajectories are not only

sensitive to spatial attraction toward the incorrect response, but they also exhibit effects which are centrally based and then cascade into response execution.

References

- Allport, D. A., Styles, E. A. & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà, M. Moscovitch, et al. (Eds.), *Attention and Performance XV: Conscious and Nonconscious Information Processing*. MIT Press.
- Altmann, E. M. (2004). Advance preparation in task switching: What work is being done? *Psychological Science*, *15*(9), 616-622.
- Arrington, C. M. (2008). The effect of stimulus availability on task choice in voluntary task switching. *Memory & Cognition*, *36*(5), 991-997.
- Barca, L., & Pezzulo, G. (2012). Unfolding visual lexical decision in time. *PLoS One*, *7*(4), e35932.
- Braem, S., Abrahamse, E. L., Duthoo, W., & Notebaert, W. (2014). What determines the specificity of conflict adaptation? A review, critical analysis, and proposed synthesis. *Frontiers in Psychology*, *5*, 1134.
- Brown, J. W., Reynolds, J. R., & Braver, T. S. (2007). A computational model of fractionated conflict-control mechanisms in task-switching. *Cognitive Psychology*, *55*(1), 37-85.
- Cousineau, D., & O'Brien, F. (2014). Error bars in within-subject designs: a comment on Baguley (2012). *Behavior Research Methods*, *46*(4), 1149-1151.
- Cutini, S., Duma, G. M., & Mento, G. (2021). How time shapes cognitive control: A high-density EEG study of task-switching. *Biological Psychology*, *160*, 108030.
- Dale, R., Kehoe, C., & Spivey, M. J. (2007). Graded motor responses in the time course of categorizing atypical exemplars. *Memory & Cognition*, *35*(1), 15-28.
- Diamond, A., & Kirkham, N. (2005). Not quite as grown-up as we like to think: Parallels between cognition in childhood and adulthood. *Psychological Science*, *16*(4), 291-297.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143-149.

- Faust, N. T., Chatterjee, A., & Christopoulos, G. I. (2019). Beauty in the eyes and the hand of the beholder: Eye and hand movements' differential responses to facial attractiveness. *Journal of Experimental Social Psychology, 85*, 103884.
- Freeman, J. B., & Ambady, N. (2009). Motions of the hand expose the partial and parallel activation of stereotypes. *Psychological Science, 20*(10), 1183-1188.
- Freeman, J. B., & Ambady, N. (2010). MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods, 42*(1), 226-241.
- Freeman, J., Dale, R., & Farmer, T. (2011). Hand in motion reveals mind in motion. *Frontiers in Psychology, 2*, 59.
- Gilbert, S., & Shallice, T. (2002). Task switching: A PDP model. *Cognitive Psychology, 44*, 297-337.
- Goschke, T. (2000). Intentional reconfiguration and involuntary persistence in task set switching. In S. Monsell & J. S. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 331–355). Cambridge, MA: MIT Press.
- Grage, T., Schoemann, M., Kieslich, P. J., & Scherbaum, S. (2019). Lost to translation: How design factors of the mouse-tracking procedure impact the inference from action to cognition. *Attention, Perception, & Psychophysics, 81*(7), 2538-2557.
- Hindy, N. C., & Spivey, M. J. (2008). Motor dynamics of task switching. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, Washington, DC, July 23–26, 2008 (pp. 2474–2479). Hoboken, NJ: Wiley-Blackwell.
- Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology, 14*, 5-81.
- Kieras, D., Meyer, D., Ballas, J., & Lauber, E. (2000). Modern computational perspectives on executive mental processes and cognitive control: Where to from here? In S. Monsell & J. Driver (Eds.), *Attention and Performance XVIII: Control of Cognitive Processes*. Cambridge, MA: MIT Press.

- Kieslich, P. J., Henninger, F., Wulff, D. U., Haslbeck, J. M. B., & Schulte-Mecklenbeck, M. (2019). Mouse-tracking: A practical guide to implementation and analysis. In M. Schulte-Mecklenbeck, A. Kühberger, & J. G. Johnson (Eds.), *A handbook of process tracing methods* (pp. 111-130). New York, NY: Routledge.
- Krueger, B. I., & Storkel, H. L. (2020). Children's response bias and identification of misarticulated words. *Journal of Speech, Language, and Hearing Research, 63*, 259-273.
- Lee, M D., & Wagenmakers, E. J. (2013). *Bayesian data analysis for cognitive science: A practical course*. New York, NY: Cambridge University Press.
- Logan, G. D., & Bundesen, C. (2003). Clever homunculus: Is there an endogenous act of control in the explicit task-cuing procedure? *Journal of Experimental Psychology: Human Perception and Performance, 29*, 575-599.
- Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review, 108*, 393-434.
- Makowski, D., Ben-Shachar, M. S., & Lüdtke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian Framework. *Journal of Open Source Software, 4*(40), 1541.
- Mayr, U., & Bryck, R. L. (2005). Sticky rules: integration between abstract rules and specific actions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(2), 337.
- Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1124–1140.
- 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.

- Meiran, N. (2010). Task switching: Mechanisms underlying rigid vs. flexible self control. In R. R. Hassin, K. Ochsner, & Y. Trope (Eds.), *Self control in society, mind and brain* (pp. 202–220). New York, NY: Oxford University Press.
- Meiran, N., & Kessler, Y. (2008). The task rule congruency effect in task switching reflects activated long-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(1), 137.
- Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. *Cognitive Psychology*, *41*(3), 211-253.
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*, 134–140.
- Morey, R. D., & Rouder, J. N. (2018). *BayesFactor: Computation of Bayes Factors for common designs*. R package version 0.9.12-4.2. <https://CRAN.R-project.org/package=BayesFactor>
- Paap, K. R., Myuz, H. A., Anders, R. T., Bockelman, M. F., Mikulinsky, R., & Sawi, O. M. (2017). No compelling evidence for a bilingual advantage in switching or that frequent language switching reduces switch cost. *Journal of Cognitive Psychology*, *29*(2), 89-112.
- Prior, A., & MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and cognition*, *13*(2), 253-262.
- R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, *124*(2), 207.
- Rubin, O., & Meiran, N. (2005). On the origins of the task mixing cost in the cuing task-switching paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(6), 1477.
- Scherbaum, S., & Dshemuchadse, M. (2020). Psychometrics of the continuous mind: Measuring cognitive sub-processes via mouse tracking. *Memory & Cognition*, *48*(3), 436-454.

- Scherbaum, S., & Kieslich, P. J. (2018). Stuck at the starting line: How the starting procedure influences mouse-tracking data. *Behavior Research Methods*, *50*(5), 2097-2110.
- Schoemann, M., Lüken, M., Grage, T., Kieslich, P. J., & Scherbaum, S. (2019). Validating mouse-tracking: How design factors influence action dynamics in intertemporal decision making. *Behavior Research Methods*, *51*(5), 2356-2377.
- Spivey, M. J., & Dale, R. (2004). On the continuity of mind: Toward a dynamical account of cognition. *Psychology of Learning and Motivation*, *45*, 87–142.
- Spivey, M. J., Richardson, D., & Dale, R. (2009). The movement of eye and hand as a window into language and cognition. In E. Morsella, J. A. Bargh, & P. M. Gollwitzer (Eds.), *Oxford handbook of human action* (pp. 225–249). New York: Oxford University Press.
- Sudevan, P., & Taylor, D. A. (1987). The cuing and priming of cognitive operations. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(1), 89.
- Tomlinson Jr, J. M., Bailey, T. M., & Bott, L. (2013). Possibly all of that and then some: Scalar implicatures are understood in two steps. *Journal of Memory and Language*, *69*(1), 18-35.
- Vandierendonck, A., Liefoghe, B., & Verbruggen, F. (2010). Task switching: Interplay of reconfiguration and interference control. *Psychological Bulletin*, *136*, 601–626.
- Waszak, F., & Hommel, B. (2007). The costs and benefits of cross-task priming. *Memory & Cognition*, *35*(5), 1175-1186.
- Weaver, S. M., & Arrington, C. M. (2013). Tracking the multitasking mind. *Zeitschrift für Psychologie*, *221*, 51-60.
- Wirth, R., Foerster, A., Kunde, W., & Pfister, R. (2020). Design choices: Empirical recommendations for designing two-dimensional finger-tracking experiments. *Behavior Research Methods*, *52*(6), 2394-2416.

Wylie, G., & Allport, A. (2000). Task switching and the measurement of "switch costs". *Psychological Research*, 63(3-4), 212-233.

Yehene, E., & Meiran, N. (2007). Is there a general task switching ability? *Acta Psychologica*, 126(3), 169-195.