



SOFTWARE TOOL ARTICLE

REVISED Introducing the Task Switching Game: a paradigm for neuroimaging and online studies [version 2; peer review: 1 not approved]

Danielle L. Kurtin ¹, Dr. Matthew Jaquiere ², Dr. Tibor Auer ¹, Prof. Adam Hampshire³, Dr. Ines R. Violante¹

¹Neuromodulation Laboratory, School of Psychology, Faculty of Health and Medical Science, University of Surrey, Guildford, GU2 7XH, UK

²Yeung Lab, Department of Experimental Psychology, Oxford University, Oxford, UK

³Imperial College London, UKDRI: Care Research and Technology Centre, Department of Brain Sciences,, Imperial College London, London, W12 0BZ, UK

V2 First published: 31 Mar 2022, 11:377
<https://doi.org/10.12688/f1000research.109729.1>
Latest published: 27 Jul 2022, 11:377
<https://doi.org/10.12688/f1000research.109729.2>

Abstract

While writing this abstract I received an email, which I promptly answered. When I returned my attention to the abstract, I struggled to regain my flow of writing. In order to understand this deficit in performance associated with switching from one task to another, or "switch cost", cognitive neuroscientists use task switching paradigms to recreate similar experiences. However, many researchers may be familiar with the difficulties that accompany modifying an established paradigm to suit their experimental design, or even the challenge of creating a new, unvalidated paradigm to perturb a particular aspect of cognitive function. This software tool article introduces a novel task switching paradigm for use and adaptation in online and neuroimaging task switching studies. The paradigm was constructed with a flexible, easily-adapted framework that can accommodate a variety of designs. This paradigm utilizes three psychometrically opposed but visually similar tasks- the Digit Span, the Spatial Span, and the Spatial Rotation. In two Use Cases we demonstrate the reliable nature of overall task performance and the dependence of switch costs on certain task parameters. This task framework can be adapted for use across different experimental designs and environment, and we encourage researchers to modify the task switching game for their experiments.

Keywords

task switching, experimental design, software, cognitive task, switch cost

Open Peer Review**Approval Status** 

1

version 2


(revision)
27 Jul 2022

version 1

31 Mar 2022



view

1. **Frini Karayanidis** , The University of Newcastle, Callaghan, Australia

Any reports and responses or comments on the article can be found at the end of the article.



This article is included in the **Mini Hacks** collection.

Corresponding author: Danielle L. Kurtin (d.kurtin@surrey.ac.uk)

Author roles: **Kurtin DL:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Resources, Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Jaquierey DM:** Data Curation, Methodology, Software, Writing – Original Draft Preparation, Writing – Review & Editing; **Auer DT:** Conceptualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Hampshire PA:** Conceptualization, Supervision, Writing – Review & Editing; **Violante DIR:** Conceptualization, Formal Analysis, Methodology, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: IRV and TA were funded by the BBSRC (Ref: BB/S008314/1). DK was funded by the Vice Chancellor Studentship at the University of Surrey. MJ was funded by the Medical Research Council (Ref: 17/18_MSD_661552).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2022 Kurtin DL *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Kurtin DL, Jaquierey DM, Auer DT *et al.* **Introducing the Task Switching Game: a paradigm for neuroimaging and online studies [version 2; peer review: 1 not approved]** F1000Research 2022, 11:377 <https://doi.org/10.12688/f1000research.109729.2>

First published: 31 Mar 2022, 11:377 <https://doi.org/10.12688/f1000research.109729.1>

REVISED Amendments from Version 1

In response to feedback received during peer review, we have updated the text's Methods section and further contextualized our work within existing task switching literature. The first addition to the Methods section details the different implications of including varying inter trial intervals in neuroimaging or online versions of the task. Our second addition to the Methods section clarifies how we defined and computed switch costs. To contextualize this work within wider task switching literature, we added sentences in the Introduction and Limitations to ensure readers are aware that results from our experiment may not generalize to task switching studies with more standard designs. In the Introduction and Limitations sections, we state that our paradigm does not consistently introduce a switch cost, and we clarified that our task design does not permit comparison of switch vs restart costs, nor mixing costs, which may have influenced the ability to induce switch costs. Further additions to the introduction emphasized how and why our paradigm is intentionally different from most task switching paradigms. These additions state our paradigm enables investigators to study research questions traditional task switching paradigms are not well-suited to investigating, such as evaluating the neural correlates of large set shifts.

Any further responses from the reviewers can be found at the end of the article

Introduction

Switch costs are defined as the deficit in task performance incurred when switching between one task and another.^{1,2} Behavioral switch costs are observed when comparing successive trials in which participants switch between tasks to those where the same task was repeated. This switch cost can be viewed as a result of the increasing demand on executive function incurred by restructuring one mental “task set” (the goals, rules, and attentional focus unique to one task) to a different one.^{3,4} Put simply, switch costs may be a result of interference in cognitive restructuring processes. Neuroimaging studies classify the reconfiguration process as a result of the changes in brain network activity and functional connectivity that reflect the changes in task set.^{5,6} The neural correlates of each task set may be considered the brain state unique to each task set.

Many studies have investigated how unique, overlapping, dissociable, and predicative these brain states are.⁷⁻¹⁰ To determine how unique the brain states for each task set are, Soreq *et al* built a classifier that identifies which working memory task a participant was completing based on their brain states.¹¹ They showed that behaviorally distinct aspects of working memory mapped to distinct but densely overlapping patterns of activity and connectivity within the brain, known as the multiple demand cortex.^{12,13} The differing mental processing characteristics that underpin participant's task performance are known as psychometric characteristics,¹⁴ and three tasks used in Soreq *et al*'s study- the Digit Span, Spatial Span, and Spatial Rotation- maximized the psychometric distance across orthogonal factors- visuospatial reasoning and verbal reasoning.¹⁵ Though all tasks recruited the multiple demand cortex, brain states could be separated according to the working memory processes recruited by each task, showing a high correspondence among behavioral constructs and resulting working memory subprocesses.^{11,16,17}

This invites the follow-up question “How does the brain reconfigure between these different brain states?” The literature here is sparser, with a lack of studies that model how neural networks reconfigure when transitioning from one discrete task to another (referred to as “set switching” or “context switching”¹⁸). In future studies we hope to characterize the trajectory neural networks take to effectively switch between tasks. Therefore, we created a cued task switching paradigm that aims to generate a behavioral and physiological switch cost for use in experiments that will characterize, model, and modulate the switch cost. We chose three psychometrically opposed tasks from Soreq *et al*¹¹ to force distinct reconfiguration from one brain state to the next. Rather than switching between stimulus response mappings or rules, our task switches between entire task sets, similar to.¹⁹ Yet our task differs from to Allport's set-shifting task by shifting among different, psychometrically opposed working memory tasks, rather than rules or stimuli within a task. These differences were introduced with the aim of inducing large set shifts observable by fMRI, where future studies may explore how neural networks reconfigure to meet the demands of different working memory tasks.

Two versions of the task exist- one is written in JavaScript to collect behavioral data online, and the other is written in Python for use in neuroimaging studies. These versions are designed to be highly similar to one another. We describe both in detail below, then present the results of two pilot studies. The pilot studies observe that, though the two versions of the task do not consistently induce a switch cost, the paradigm operates within an optimal difficulty range, and participants do not exhibit learning effects. Though our task does not produce traditional switch costs, we believe this paradigm is useful given its highly adaptable, multi-modal, open-source nature.

Methods

Implementation

In this section, we first provide a description of each of the three tasks' features. Then, we detail how the overall task is compiled, and what sections may be modified to suit different experimental designs.

Task Descriptions

The cued task switching paradigm switches between three tasks- a Spatial Rotation, Spatial Span, and Digit Span task. Variants of these tasks have been created and implemented over the years.^{10,20} The essential components of each task are clarified below.

All tasks use a similar stimuli presentation and response framework to reduce visual and motor confounds. Stimuli are created using normed pixel units, and the screen angle is standardized to reduce color variation across devices. Stimuli are presented on a 6x6 grid in the middle of the screen. To ensure the tasks are visually similar, each task's stimulus grid flashes cells and contain numbers, even if not strictly necessary for the task. After presentation is complete, the stimulus grid disappears, and three answer grids appear in a row across the screen. One of the answer grids contains the correct answer, and the other two have one of the cells from the correct answered shifted, meaning two of the three grids' answers are incorrect by one cell (Figure 1).

Digit Span measures verbal working memory capacity. In this task, a sequence of 6 numbers within a shaded box on the stimulus grid appears one after another. One of the three answer grids will contain the correct sequence of numbers, and the other two will be correct except for one digit. This is a variant of the WAIS-R intelligence test that evaluates working memory.²⁰ **Spatial Span** tests visuospatial working memory capacity. 6 squares flash digits appearing in a random sequence, one after another, in the stimulus grid. The correct answer grid will display the same sequence of numbers that flashed in the stimulus grid, while the other two display a sequence that is incorrect by one cell. Finally, **Spatial Rotation** measures the ability of the participant to mentally rotate objects in memory. Similar to the Spatial Span, shaded cells appear one after another, though in this task the previous cells continue to flash with each new addition. The resulting end stimulus is a flashing grid of 6 cells. The answer grids contain a 90, 180, or 270-degree rotation of the final grid, with two of the three answer grids incorrect by one cell. Gifs showing trials of each task can be found here: https://github.com/daniellekurtin/task_switching_paradigm/tree/master/TaskGifs.

Compiling the paradigm

We describe the implementation of the task switching paradigm as created for our experimental use, rather than the software package as a whole. We provide the software package as an example of how it may be implemented and used for an experiment.

Running the *main.py* script initiates an implementation of the task switching paradigm. The paradigm consists of blocks composed of a sequence of tasks. Each task is composed of a run of trials. Trials consist of stimuli and answer grids (Figure 1). Runs are set up so that the last run on one block continues as the first run on another block. For example, if the last run within a block consists of 9 trials of Digit Span, then the break could occur on trial 7, and after the break, the remaining two trials would be the first two trials of the next block. This approach maximizes the number of task switches in each block while keeping the number of runs balanced across each of the three task types. The *main.py* implementation begins with a popup to record participant and session information (*taskSwitching.participant_gui.py*). After the popup is dismissed, a scanner sync process is initiated, creating a Pythonic interface for neuroimaging experiments. Then *main.py* constructs a demo that participants may play multiple times to ensure their familiarity with how to play the tasks. The demo's parameters are set by *taskSwitching.ExperimentTaskSwitch* class (and its parent *taskSwitching.Experiment* class), with trials determined by *main.py*. Then, *main.py* constructs a new task blueprint using the default parameters set by *taskSwitching.ExperimentTaskSwitch* class. If desired, implementations may specify the types of tasks the paradigm will switch between, the length of the cue cards, the number of trials per task, the duration of each stimuli, and more, as demonstrated in the tutorial construction. A pseudorandomized list of trials, runs, and blocks are constructed based on the provided specifications, ensuring there are an equal number of switches for each task type. Each task's trials are instances of classes unique to each task type: *taskSwitching.TrialDigitSpan*, *taskSwitching.TrialSpatialSpan*, and *taskSwitching.TrialSpatialRotation*. Parameters may be set at the Experiment, Component, Trial, or specific trial task level, with the later parameters overriding earlier ones where there are conflicts. Values that can be set in this way include how stimuli and answers are created and displayed, and for how long.

Trials are instances of Components, and cue cards, instructions, and breaks are also components. *taskSwitching.Components* include the following: *taskSwitching.ComponentRest* determines the rest screen; *taskSwitching.ComponentStart* is the screen before participants begin the task switching game; *taskSwitching.ComponentInfoCard* creates the cue cards that prompt a task switch; and *taskSwitching.ComponentTrialGap* fixes the screen that appears between trials.

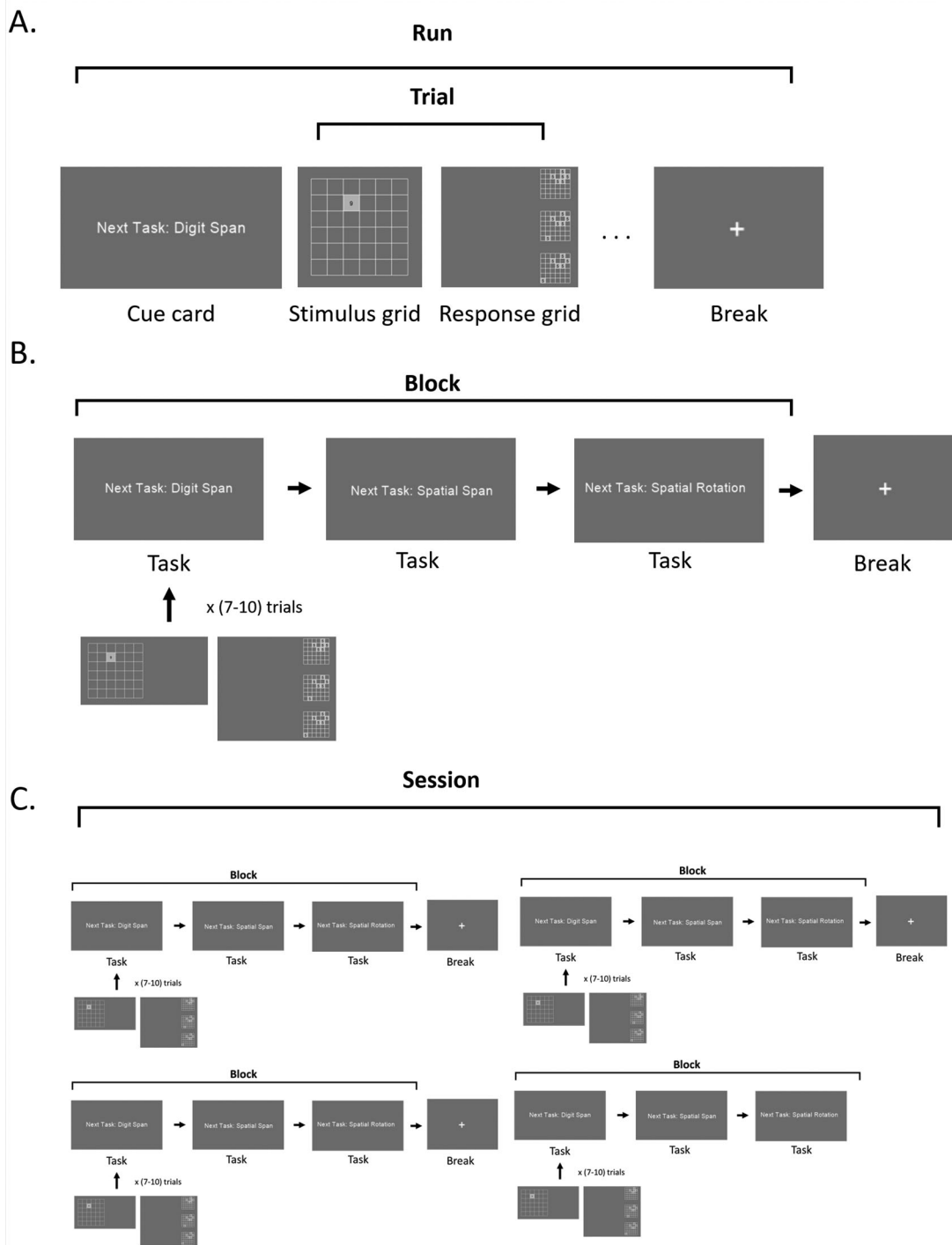


Figure 1. Depiction of how A. trials build into a run of a task. These tasks form B. blocks, which together, compose C. a session. This depiction represents the task described in the manuscript, but all components shown can be changed to suit the needs of other experiments.

As aforementioned, trials consist of stimuli and their answer grids, which are constructed according to the *taskSwitching.Grid* class. Finally, as the task is played, information is saved to a.csv file. What is saved and the file format is set by the *taskSwitching.Experiment* class.

Operation

This task is executable in a Pythonic environment. Touch events (i.e., participant's responses) can be collected via button box, through keyboard strokes, and by mouse clicks. Responses minimize motor confounds through requiring a single button press or click to select an answer grid for all three tasks. We will now describe the workflow and design features for both the neuroimaging and online versions.

Neuroimaging studies

The paradigm begins with a participant GUI that requires entering the participant ID, age, gender, and session ID (fields can be adjusted depending on specific study needs). Participants then play a demo that includes each of the tasks at least once. The demo includes performance feedback: after participants select an answer grid, a green box will highlight either the correct answer grid, or the space it would occupy. This gives participants a better understanding of how well they comprehend the task's rules. After a loading screen the participant is presented with a cue card stating they may press any button on the button box to begin. Once a button is pressed, the first cue card is presented, followed by the first trial of that task. After the first stimulus is finished, there is a variable delay before participant responses are enabled. Participants then have a window to respond. Once the participant selects their answer grid, the other two disappear, and the first answer grid is held on the screen for the remainder of the response window. This serves two purposes. First, by eliminating the other answer grids, we provide feedback that their answer has been recorded, preventing repetitive button presses. Second, we eliminate the potential for participants to compare their answer to the other answer grids. Once the trial is over, there is a variable intertrial interval (e.g. 100 to 1100 msec) to introduce a jitter. The jitter is used to improve reliability of fMRI signals, and increase the spatio-temporal resolutions.²¹ For online studies, we recommend researchers include a jitter-induced delay related regressor in models of BOLD activity, and remove jitter during offline studies as recommended by Ref. 22 for response-cue trial intervals. The number of trials per run range is modifiable (5-10 might be a reasonable number for fMRI experiments). Once the task is complete, a cue card stating "Next Task: [Digit Span/Spatial Span/Spatial Rotation]" is displayed to indicate which task is next. There are no task repeats (i.e., if the previous task was the Digit Span, the next would be either the Spatial Span or Spatial Rotation). The duration of the cue card is a random choice of either 0.5 or 4.0 seconds, though the number of cue cards and their length can be varied. This enables future investigation into the effects of short vs long cue presentation on neural network dynamics and task performance. After the trials within the block are complete, a break occurs (though the presence and/or duration of a break can be modified). The break screen contains a centered fixation cross, and a countdown until the task restarts. Once all blocks are complete, the task quits, and data is saved in a.csv file. Participant reaction time is computed as the difference between when they submit an answer and when a response was enabled for that trials. Reaction time is measured in seconds with hundred-millionths period precision.

Online studies

Our version of the task switching paradigm is hosted on the University of Surrey's web servers. The university servers serve three main functions: enabling participants to access and play the task, recording their performance, and storing "task blueprints" (Figure 2). These "task blueprints" are pre-compiled sessions (the order of tasks, the number of trials per task, etc), and are the same as the tasks generated locally. Uploading the task blueprints is simple, and reduces the burden on the server. These blueprints are created using *serve - trial - sequences.py*, with dependencies and scripts used to communicate among servers located in the *www* folder of the paradigm's repository. Participants access the task from the link <http://www.task-switching-game.surrey.ac.uk>. They are walked through a tutorial with written instructions and accompanying animations. Participants then play the same demo described in the above section. After the demo, participants are invited to either play it again, or continue on to the main task. Once they continue, they read an ethics statement and fill out consent checklists and their participant information (participant ID, age, sex). Next, they receive the instruction to "Press next to begin." At this stage the online version of the task is as the neuroimaging version, except it consists of one, 20-minute long block, and answer selection is done using the mouse (the task is configured to work using a keyboard, mouse, or button box).

Use Case 1: Online study

The study was advertised on SONA, a participant recruitment and experiment management system that connects participants to ongoing studies. Participants could sign up and play the task switching game, and were awarded course credit for completion. All participants gave informed consent. This study was conducted with ethical approval by the University of Surrey Ethics Committee.

Parameters for this use case are as follows:

- Delay from stimulus end to participant response window: 0.15 s

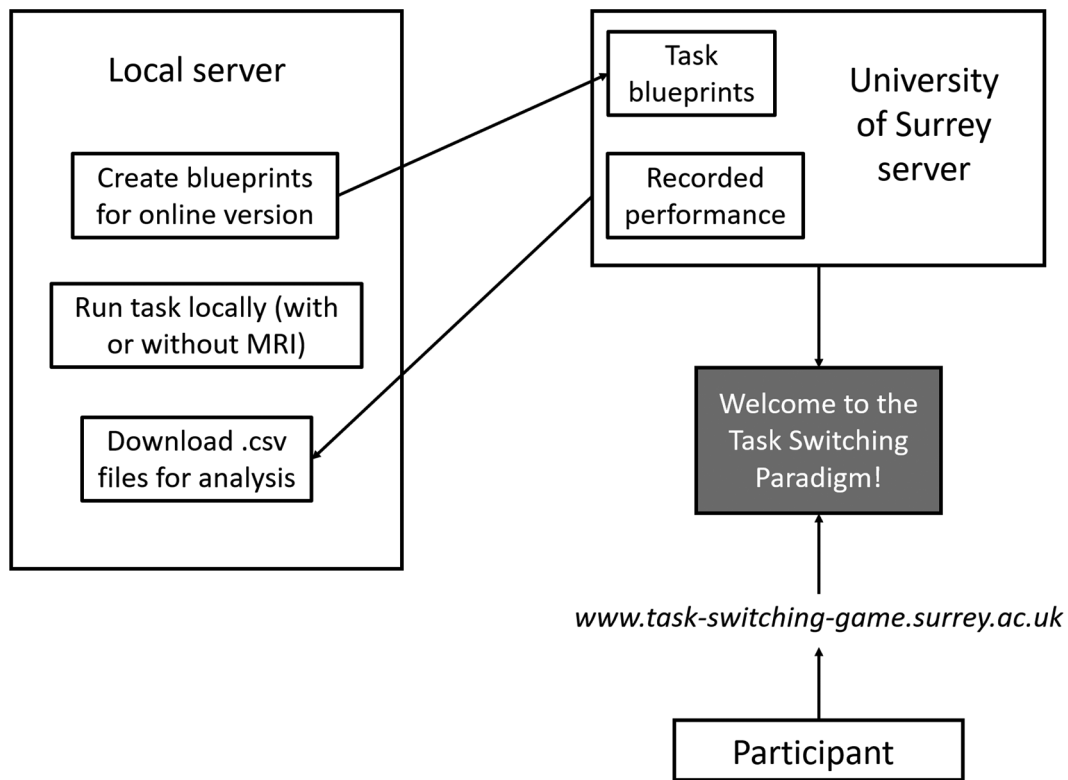


Figure 2. Schematic of various servers used to execute the online version of the paradigm and their broad functions. Participants accessed the task using a link that can be opened via web browsers using either a mobile device or computer (though we specified we prefer that participants used a computer).

- Participant response window: 3.0 s
- Block break: 120.0 s
- Response mode: mouse
- Inter-trial interval: 100-1100 msec
- Cue card length: 0.5 or 4.0 s
- Number of trials per occurrence: 6-9 trials
- Range of trials: 176-187

Data analysis was conducted using the .csv file output by the online task in a MATLAB environment. Non-normally distributed performance data was normalized by computing the z-score with a center of zero and one standard deviation. Switch task types are defined as the first trial after a switch between tasks, and stay trial types are all other trials. Occurrence refers to the number of times a participant has played a task. For example, if the task starts with a Digit Span, then switches to the Spatial Rotation, then back to the Digit Span, the occurrences would be 1, 1, 2. Linear mixed effects models are used to assess the effect of task type, occurrence, and trial type on behavioural performance, with subjects included as random effects. Post-hocs are evaluated using T-tests.

Results

Data cleaning The total number of online participants was $n=87$, with a mean age of 19.68, and all participants were university students. We removed any sessions with less than 100 trials ($n=7$). There were no participants that had $>20\%$

Table 1. Behavioral performance for the first online pilot.

| | Overall mean \pm sd | Digit Span mean \pm sd | Spatial Span mean \pm sd | Spatial Rotation mean \pm sd |
|----------|-----------------------|--------------------------|----------------------------|--------------------------------|
| MRT (ms) | 1826.5 \pm 501 | 1837.7 \pm 505 | 1826.2 \pm 489 | 1813.1 \pm 511 |
| Accuracy | 60.57% \pm 22 | 66.31% \pm 23 | 53.68% \pm 21 | 61.51% \pm 21 |

omissions in any task. We removed participants that performed below chance level for any task ($n=19$), leaving us with a final cohort of 61 participants ($n=52$ females). Each participant had an average of 178.9 trials ($sd=12.37$) overall.

Overall performance-reaction time and accuracy Performance was evaluated using accuracy and mean reaction time (MRT) (Table 1). Kolgov-Smirnov tests show both accuracy ($D(1509)=0.55$, $p\leq 0.0001$) and MRT ($D(10909)=1$, $p\leq 0.0001$) are non-normally distributed, and were thus normalized.

There was a significant effect of task type on accuracy ($F[2,1503]=10.8$, $p=2.1e-05$). Post-hoc T-tests show each task's accuracy is significantly different than the others, with Digit Span performing 12.6% and 4.8% better than Spatial Span ($t(502)=-9.6$, $p=3.1e-20$) and Spatial Rotation ($t(502)=-3.6$, $p=3.7e-04$) respectively, (Figure 3A). Spatial Rotation performance was 7.8% higher than Spatial Span performance ($t(502)=-6.1$, $p=5.5e-11$). There was no significant effect of occurrence on accuracy ($F[1,1503]=2.6$, $p=0.10$) (Figure 3B).

There was a significant effect of occurrence on MRT ($F[1,1481]=8.4$, $p=0.004$) (Figure 3D). A Bland-Altman plot was created to investigate whether the effect of occurrence was a result of unreliable RT recording or learning effects. There are no remarkable effects on the data, as shown in the (Figure 4A). There was no significant effect of task type on MRT ($F[2,1481]=0.1$, $p=0.92$) (Figure 3C).

Switch Cost There is a significant overall switch cost in accuracy ($t(1447)=3.0$, $p=0.003$) (Figure 5A) but not MRT ($t(811)=0.85$, $p=0.39$) (Figure 5B). Because there is an effect of task type on accuracy, we looked at whether there was an effect of switch on each task's performance. After Bonferroni correction for multiple comparisons, there is a significant switch cost in Spatial Span ($F[1,120]=7.4$, $p=0.008$) and Spatial Rotation ($F[1,120]=7.4$, $p=0.007$) accuracy, but not Digit Span ($F[1,120]=0.25$, $p=0.62$) (Figure 5C-E).

We also sought to determine whether switch cost was influenced by switch type- the six possible combinations of how one task may switch to another. For example, a switch from Digit Span to Spatial Span is one switch type, and vice versa, another. We found a significant effect of switch type ($F[1,425]=6.8$, $p=0.009$) on accuracy but not MRT ($F[1,405]=1.3$, $p=0.26$). Post-hoc ttests found no significant differences in accuracy per switch type after Bonferroni corrections for multiple comparisons.

Discussion

In this first online pilot of the task switching paradigm we found task type influenced accuracy. The better performance in the Digit Span compared to the Spatial Span is not surprising. A study of 44,600 participants playing a range of cognitive tasks online found that, when playing the Digit Span, the average number of stimuli remembered by participants is 7, whereas the average number of stimuli remembered for the Spatial Span is 6.²⁰ This means that, using our 6x6 grid, the number of stimuli to retain for the Digit Span is well within the abilities of our population. The discrepancy in performance between the Digit Span and Spatial Rotation is less clear. However, in a previous study comparing performance between working memory tasks that greatly resemble the Digit Span and Spatial Rotation, performance on the Digit Span analog was significantly better than their analog visuospatial task.²³ Finally, the difference in performance between the Spatial Span and the Spatial Rotation may be a result of participant's ability to form effective strategies for each task. A study by Gardony *et al* investigated mental rotation tasks and found that, as difficulty increased, cognitive strategies shifted in order to meet the demands of the task.²⁴ Participants playing the Spatial Span can more easily rely on recognition strategies than in the Spatial Rotation, where participants not only need to recall patterns, but perform a mental rotation of the patterns as well. The additional demands of the Spatial Rotation task may have resulted in the discrepancy between Spatial Span and Spatial Rotation performance.

We found an influence of occurrence on MRT. The greatest difference in MRT per occurrence is between occurrence 1 and 7. MRT in occurrence 7 is 6% faster than during occurrence 1; a marginal improvement over the duration of the experiment.

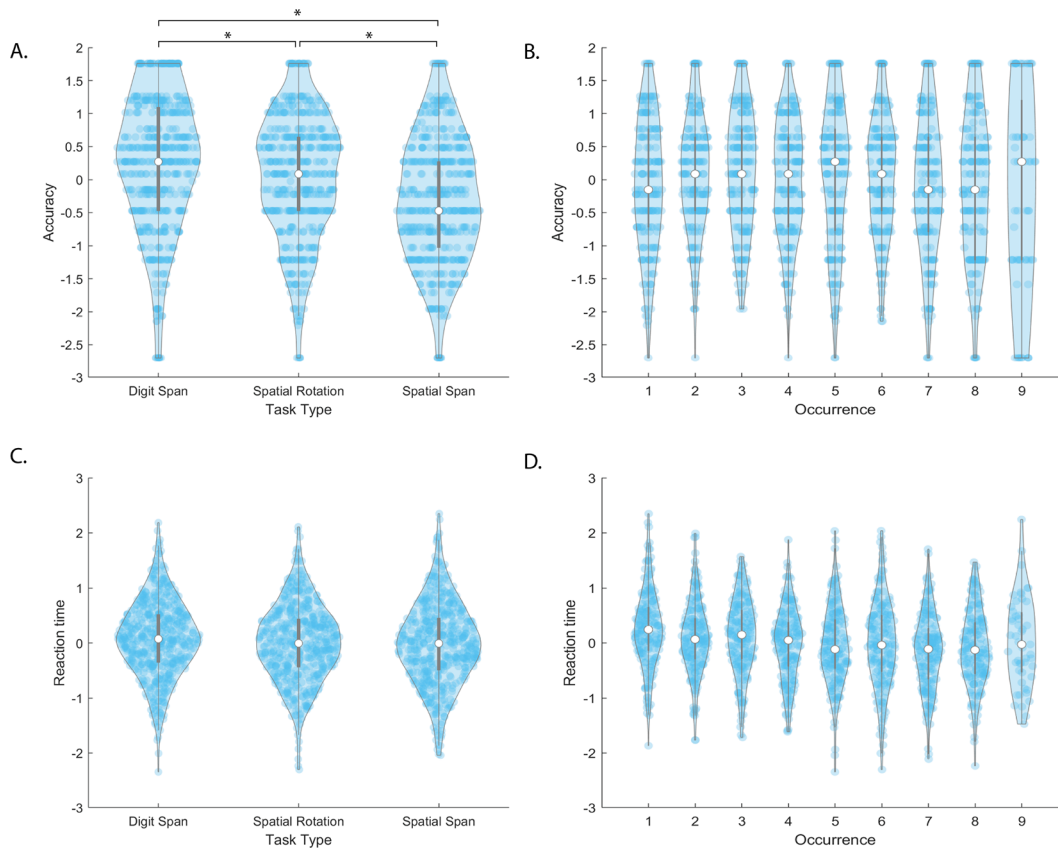


Figure 3. Violin plots show accuracy A. per task type and B. occurrence, as well as MRT A. per task type and D. occurrence. The top and bottom edges of the grey boxes represent the 25th and 75th percentiles, with the mean being the white dot. Extension of the whiskers limits outliers. A kernel density estimate of the data provides the edges to the violin plot, and individual data points are dark blue. * denotes significance at $p \leq 0.05$ after Bonferroni correction for multiple comparisons.

The switch cost in accuracy, but not reaction time, demonstrates a partial success of our aim to create a task switching paradigm that forces a behavioral switch cost. The presence of switch costs in accuracy, but not reaction time, may be a function of the response window imposed on participants. A study by Hughes *et al* found that switch accuracy fell 29% by introducing a response time window,²⁵ but this switch cost did not extend to reaction time. Our response window of three seconds is likely enough to induce a time pressure on participants, as well as the added switch cost in accuracy, but not reaction time. The explicit cues informing participants a switch is about to occur may have reduced the behavioral switch cost. A study by Merian also using a random, cued, task switching paradigm reported a smaller switch cost than the switch cost observed in a study by Monsell without random task switches.^{1,26} Tornay and Milan compared the two studies and hypothesized that the cue for a change in task gives participants time to suppress the current task set. This initiates the cognitive restructuring process of task switching, thus increasing participants' ability to quickly reconfigure to the demands of the new task.⁴ Our cue card intervals of 0.5 and 4.0 seconds were likely long enough to allow participants to suppress the currently active task set when the cue card is shown. This preparatory process may decrease the switch cost, but not the cognitive restructuring process of switching. We did not collect neuroimaging data for this study, but we plan to in the future, and will investigate this arm of research.

Due to an error in the data collection process, we were unable to assess how the variation in cue card presentation length effected participant's performance. Because of the potentially significant influence this variability may have had on participant's performance, we standardized the cue card length from 4.0 seconds to 0.5 seconds, and conducted a second round of data collection. The results from this second use case are detailed below.

Use Case 2: Online study with standardized cue card length

As a result of our inability to calculate the impact of cue card length on MRT and accuracy, we standardized the cue card length to be 0.5 seconds. Our data collection and analysis were conducted using the same methods as above.

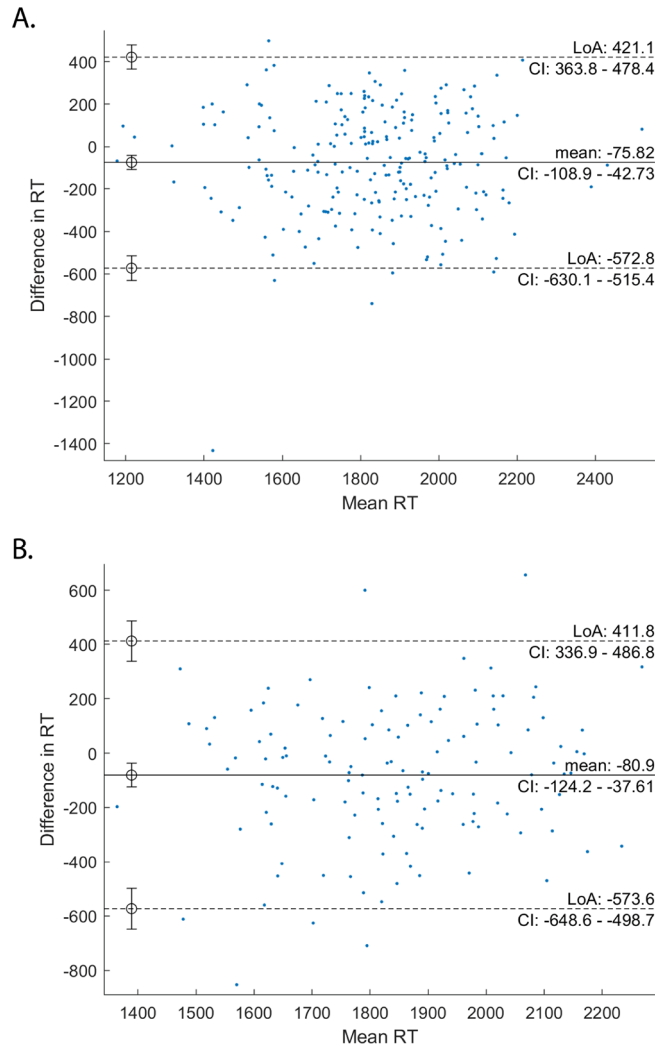


Figure 4. Bland-Altman plots of the difference in mean reaction time between the first and second half of the session per participant (y-axis) against the mean reaction time (x-axis) for the (A.) first and (B.) second Use Cases. Plots show the mean (solid line), 95% confidence limits (dashed lines), and the data points (blue points).

Results

Data cleaning The total number of online participants was n=40, and all participants were university students. We removed any sessions with less than 100 trials (n=4). We also removed participants that had >20% omissions in any task (n=0), participants that performed below chance level for any task (n=3), leaving us with a final n=33 (n=31 females) Each participant had an average of 182 trials (sd = 3.35) overall.

Overall performance-reaction time and accuracy Accuracy and MRT were used to evaluate performance (Table 2). Kolgov-Smirnov tests show both accuracy (D(825)=0.55, p=4.4e-215) and MRT (D(825)=1, p≤0.0001) are non-normally distributed.

There was a significant effect of task type (F[1,819]=3.1, p=0.047) (Figure 5A), but not occurrence (F[1,819]=0.01, p=0.94) (Figure 6B), on accuracy. Post-hoc T-tests show each task’s accuracy is significantly different than one another, with Digit Span performing 14.9% and 5.5% better than Spatial Span (t(274)=-5.6, p=4.9e-08) and Spatial Rotation (t(274)=-3.4, p=7.8e-04), respectively. Spatial Rotation performance was 9.34% higher than Spatial Span performance (t(274)=-9.1, p=2.3e-17).

There was a significant effect of occurrence on MRT (F[1,813]=12.0, p=0.0006) (Figure 6D). A Bland-Altman plot was created to investigate whether the effect of occurrence was a result of unreliable RT recording or learning effects. There

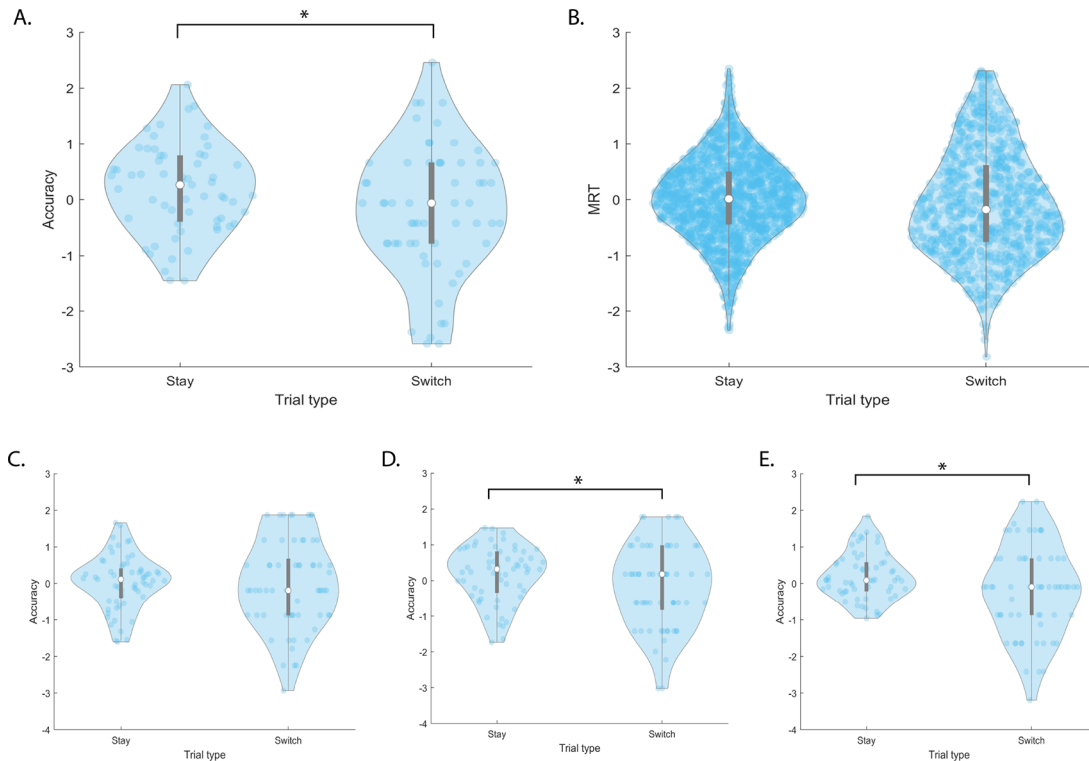


Figure 5. Violin plots show switch costs in A. accuracy, B. MRT, and C. accuracy per task type. The top and bottom edges of the grey boxes represent the 25th and 75th percentiles, with the mean being the white dot. Extension of the whiskers limits outliers. A kernel density estimate of the data provides the edges to the violin plot, and individual data points are dark blue. * denotes significance at $p \leq 0.05$ after Bonferroni correction for multiple comparisons.

are no remarkable effects on the data, as shown in the (Figure 4B). There was no significant effect of task type on MRT ($F[2,813]=0.39$, $p=0.67$) (Figure 6C).

Switch cost There is no significant switch cost in accuracy ($t(791)=-0.69$, $p=0.49$) (Figure 7A) or MRT ($t(469)=-0.34$, $p=0.73$) (Figure 7B). There was no significant effect of switch type on accuracy ($F[1,229]=0.01$, $p=0.92$).

Discussion

This online pilot sought to evaluate performance on the task switching paradigm and see how the standardization of cue cards influenced performance.

There was high similarity between the first and second pilot, with both pilots showing an effect of occurrence on MRT and task type on accuracy. However, the switch cost in the Spatial Span and Spatial Rotation observed in the first pilot was not present in the second. The loss of switch cost is surprising, especially given that the cue card length was standardized to 0.5 seconds as opposed to 4.0 second, suggesting that switch costs may be driven by a longer cue card.

This is supported by a task switching study by Periáñez and Barcelo²⁷ that studied the role of exogenous (cues) and endogenous (task-set activation) in the behavior and EEG markers of switch costs. Their experimental paradigm randomly varied the cue-trial interval (CTI) between participants as either 800 or 2000 ms. They found that the shorter CTI did not consistently lead to a greater switch cost, and in fact, influenced a cue-switch *benefit*. The results from our study are similar- Use Case 1, which had CTIs of either 500 ms or 4000 ms, exhibited a greater switch cost than in Use Case 2, which solely had CTIs of 500 ms. Their EEG results suggest this phenomena may be a result of reduced P3 activity that arises from an interplay between time-dependent endogenous (anticipatory task set reconfiguration) and exogenous (cue) factors. We suggest future studies utilize the neuroimaging compatibility of our task switching paradigm to replicate this finding.

Table 2. Behavioral performance for the second online pilot.

| | Overall mean \pm sd | Digit Span mean \pm sd | Spatial Span mean \pm sd | Spatial Rotation mean \pm sd |
|----------|-----------------------|--------------------------|----------------------------|--------------------------------|
| MRT (ms) | 1844.0 \pm 496 | 1872.5 \pm 481 | 1821.6 \pm 519 | 1831.5 \pm 492 |
| Accuracy | 60.34% \pm 23 | 68.02% \pm 22 | 53.16% \pm 21 | 62.50% \pm 21 |

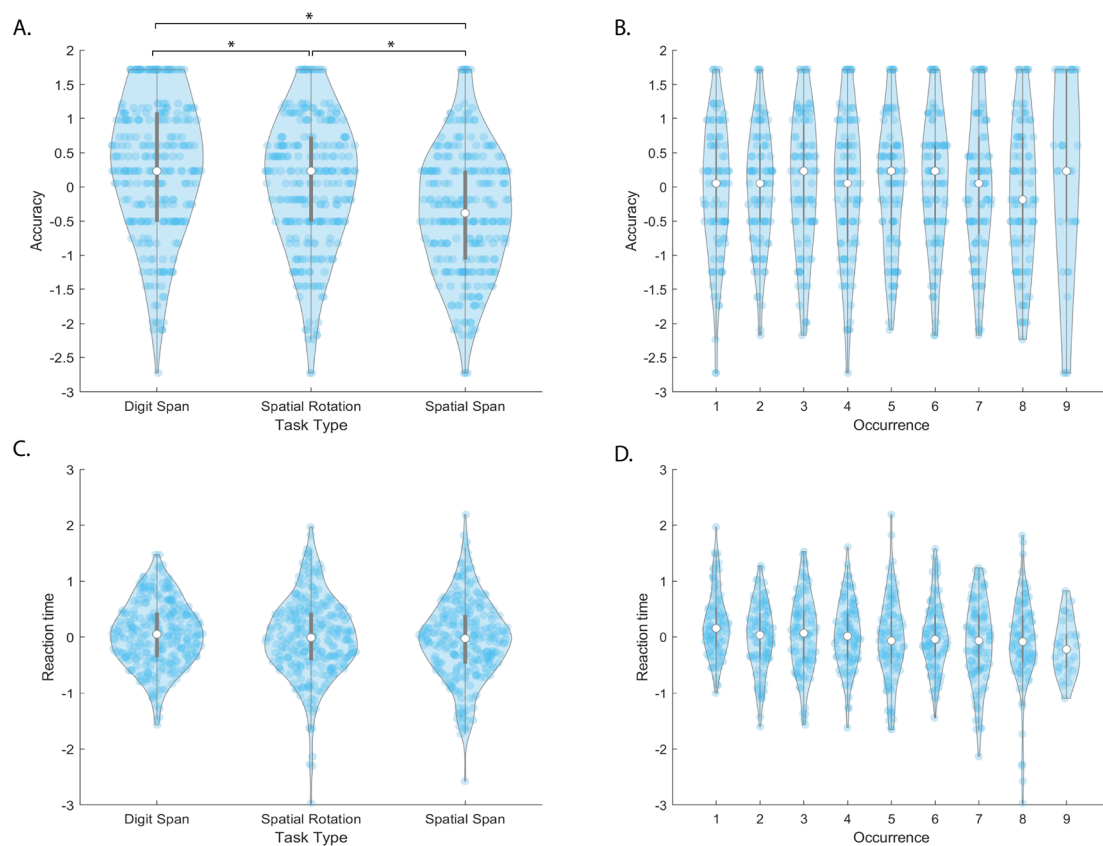


Figure 6. Violin plots show accuracy A. per task type and B. occurrence, as well as MRT A. per task type and D. occurrence. The top and bottom edges of the grey boxes represent the 25th and 75th percentiles, with the mean being the white dot. Extension of the whiskers limits outliers. A kernel density estimate of the data provides the edges to the violin plot, and individual data points are dark blue. * denotes significance at $p \leq 0.05$ after Bonferroni correction for multiple comparisons.

Limitations

This task contains two visual confounds. First, in the Spatial Span and Digit Span the boxes disappear after the initial presentation; in the Spatial Rotation, they build upon one another. The resulting end image is a visually more complex image. This confound is unavoidable due to the nature of the Spatial Rotation task. Second, the stimuli in the Digit Span present in half the time as the stimuli for the Spatial Span and Spatial Rotation, 0.25 seconds as opposed to 0.50 seconds, respectively. This difference was implemented after rounds of piloting the task, where it was noticed the Digit Span was markedly easier than the other two tasks. By reducing the stimulus presentation time we increase the difficulty of the Digit Span, making it more comparable to the other two tasks. This is important, as cognitive load influences the brain network activity and connectivity within a task.²⁸ Though the piloting of this task was performed with healthy control participants, future researchers may want to assess differences between healthy control and patient populations. Mixing costs may be more sensitive to between-group variability,²⁹ and one limitation of our task structure is that it does not permit the exploration of mixing costs. Moreover, the current design did not allow repeats blocks of the same task (for example, this order would not occur: Digit Span, Spatial Span, Spatial Span, Spatial Rotation), and therefore cannot investigate the difference in switch vs restart costs.¹⁹ Future researchers are invited to adapt the paradigm's design to allow repeat task

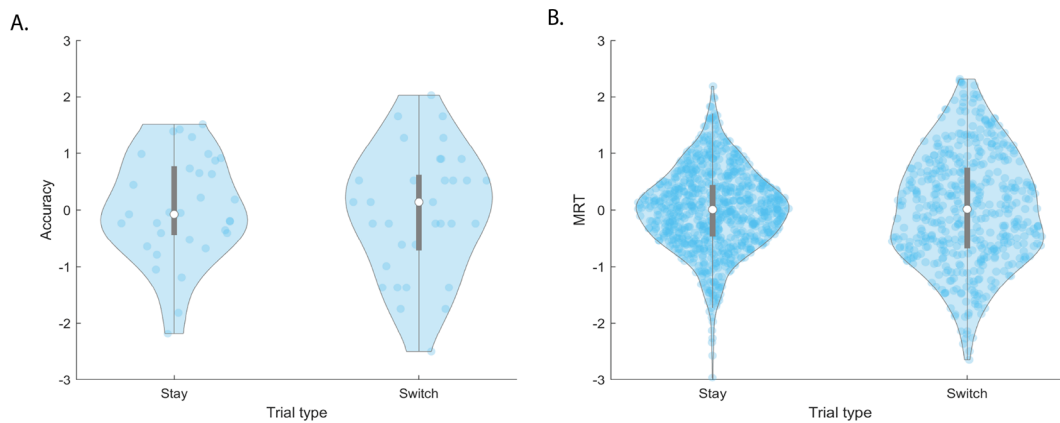


Figure 7. Violin plots show switch costs in **A. accuracy** and **B. MRT**. The top and bottom edges of the grey boxes represent the 25th and 75th percentiles, with the mean being the white dot. Extension of the whiskers limits outliers. A kernel density estimate of the data provides the edges to the violin plot, and individual data points are dark blue.

blocks, to investigate switch vs restart costs, and explore whether introducing mixed-task blocks induces switch costs not seen in this version of the paradigm.

Finally, our task differs from most variants of task switching paradigms, and this should be taken into consideration when comparing results from this task to literature using different paradigms.

Future researchers are encouraged to modify these parameters as it suits their task. The task switching paradigm was built with flexibility in mind, so it may be easily adapted to various experimental designs.

Conclusions

Searching for “task switching paradigms” reveals a staggering amount of task designs, theories, and neuroimaging data. The quantity and heterogeneity of experimental designs address specific facets of switch costs, and by proxy, cognitive function. The authors are not aware of an existing framework that can be adapted easily to suit the demands of different experiments, leaving researchers to either re-use old tasks or create entirely new ones to suit their experimental designs. We needed to construct a novel task switching paradigm, and chose to create one within a framework that can be adapted to suit the needs of different experiments. Here we introduce a flexible software package to create task switching paradigms. It can accommodate nuanced designs within a stable and robust framework for on or offline studies, and is compatible with neuroimaging methods. The task switching paradigm does induce minimal switch costs, but efforts are underway to improve the switch cost.

Data availability

Repository: Task Switching Paradigm. https://github.com/daniellekurtin/task_switching_paradigm with an MIT license.

This project contains the following underlying data:

- rawdata_pilot4.csv. (Data downloaded from the task server for Use Case 1.)
- rawdata_pilot5.csv. (Data downloaded from the task server for Use Case 2.)

Data are available under the terms of the repository’s MIT license.

Software availability

Source code available from: https://github.com/daniellekurtin/task_switching_paradigm with an MIT license.

Acknowledgements

We would like to thank Henry Hebron for his guidance in creating the figures in this manuscript.

References

- Rogers RD, Monsell S: **Costs of a predictable switch between simple cognitive tasks.** *J. Exp. Psychol. Gen.* 1995; **124**(2): 207–231.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Jersild AT: **Mental set and shift.** *Arch. Psychol.* 1927.
- Vandierendonck A, Liefvooghe B, Verbruggen F: **Task switching: interplay of reconfiguration and interference control.** *Psychol. Bull.* 2010; **136**(4): 601–626.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Tornay FJ, Milán EG: **A more complete task-set reconfiguration in random than in predictable task switch.** *The Quarterly Journal of Experimental Psychology Section A.* 2001; **54**(3): 785–803.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Yeung N, Nystrom LE, Aronson JA, et al.: **Between-task competition and cognitive control in task switching.** *J. Neurosci.* 2006; **26**(5): 1429–1438.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Dosenbach NUF, Visscher KM, Palmer ED, et al.: **A core system for the implementation of task sets.** *Neuron.* 2006; **50**(5): 799–812.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Lorenz R, Violante IR, Monti RP, et al.: **Dissociating frontoparietal brain networks with neuroadaptive bayesian optimization.** *Nat. Commun.* 2018; **9**(1): 1–14.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Shi Y, Wolfensteller U, Schubert T, et al.: **When global rule reversal meets local task switching: The neural mechanisms of coordinated behavioral adaptation to instructed multi-level demand changes.** *Hum. Brain Mapp.* 2018; **39**(2): 735–746.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Hampshire A, Daws RE, Neves ID, et al.: **Probing cortical and sub-cortical contributions to instruction-based learning: Regional specialisation and global network dynamics.** *NeuroImage.* 2019; **192**: 88–100.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Soreq E, Leech R, Hampshire A: **Dynamic network coding of working-memory domains and working-memory processes.** *Nat. Commun.* 2019; **10**(1): 1–14.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Soreq E, Violante IR, Daws RE, et al.: **Neuroimaging evidence for a network sampling theory of individual differences in human intelligence test performance.** *Nat. Commun.* 2021; **12**(1): 1–13.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Duncan J, Owen AM: **Common regions of the human frontal lobe recruited by diverse cognitive demands.** *Trends Neurosci.* 2000; **23**(10): 475–483.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Fedorenko E, Duncan J, Kanwisher N: **Broad domain generality in focal regions of frontal and parietal cortex.** *Proc. Natl. Acad. Sci.* 2013; **110**(41): 16616–16621.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Deary IJ: **Psychometric intelligence differences and brain function.** *Novartis Foundation Symposium.* Wiley Online Library; 2000; pages 58–72.
- Hitch GJ, Allen RJ, Baddeley AD: **Attention and binding in visual working memory: Two forms of attention and two kinds of buffer storage.** *Atten. Percept. Psychophys.* 2020; **82**(1): 280–293.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Baddeley A: **Exploring the central executive.** *The Quarterly Journal of Experimental Psychology Section A.* 1996; **49**(1): 5–28.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Baddeley AD, Hitch GJ, Allen R: **A multicomponent model of working memory.** *Working memory: State of the science.* 2020: 10–43.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Kim C, Cilles SE, Johnson NF, et al.: **Domain general and domain preferential brain regions associated with different types of task switching: A meta-analysis.** *Hum. Brain Mapp.* 2012; **33**(1): 130–142.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Allport DA, Wylie GR: **Task-switching, stimulus-response bindings, and negative priming.** In Monsell S, Driver J, (Eds.), *Control of cognitive processes: Attention and performance XVIII.* Cambridge, MA: MIT Press. 2000. pp. 35–70.
- Hampshire A, Highfield RR, Parkin BL, et al.: **Fractionating human intelligence.** *Neuron.* 2012; **76**(6): 1225–1237.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Watanabe M, Bartels A, Macke JH, et al.: **Temporal jitter of the bold signal reveals a reliable initial dip and improved spatial resolution.** *Curr. Biol.* 2013; **23**(21): 2146–2150.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Ruge H, Jamadar S, Zimmermann U, et al.: **The many faces of preparatory control in task switching: reviewing a decade of fmri research.** *Hum. Brain Mapp.* 2013; **34**(1): 12–35.
- Hubber PJ, Gilmore C, Cragg L: **Mathematics students demonstrate superior visuo-spatial working memory to humanities students under conditions of low central executive processing load.** 2019.
- Gardony AL, Eddy MD, Brunyé TT, et al.: **Cognitive strategies in the mental rotation task revealed by eeg spectral power.** *Brain Cogn.* 2017; **118**(0): 1–18.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Hughes MM, Linck JA, Bowles AR, et al.: **Alternatives to switch-cost scoring in the task-switching paradigm: Their reliability and increased validity.** *Behav. Res. Methods.* 2014; **46**(3): 702–721.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Meiran N: **Reconfiguration of processing mode prior to task performance.** *J. Exp. Psychol. Learn. Mem. Cogn.* 1996; **22**(6): 1423–1442.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Periáñez JA, Barceló F: **Updating sensory versus task representations during task-switching: Insights from cognitive brain potentials in humans.** *Neuropsychologia.* 2009; **47**(4): 1160–1172.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Cocchi L, Halford GS, Zalesky A, et al.: **Complexity in relational processing predicts changes in functional brain network dynamics.** *Cereb. Cortex.* 2014; **24**(9): 2283–2296.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Karayanidis F, Whitson LR, Heathcote A, et al.: **Variability in proactive and reactive cognitive control processes across the adult lifespan.** *Front. Psychol.* 2011; **2**: 318.

Open Peer Review

Current Peer Review Status: 

Version 1

Reviewer Report 19 May 2022

<https://doi.org/10.5256/f1000research.121272.r129896>

© 2022 Karayanidis F. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

 **Frini Karayanidis** 

Functional Neuroimaging Laboratory, School of Psychological Sciences, College of Engineering, Science and Environment, The University of Newcastle, Callaghan, NSW, Australia

This paper presents a well-motivated objective – to design and make available a well-controlled task-switching paradigm that can be used in multiple contexts (eg., imaging and online studies) and with different groups. This will facilitate the comparison of findings across studies and participant groups.

The authors put substantial effort into making the three tasks similar in visual presentation and response requirements, so that differences between them can be attributed to different cognitive processes rather than sensory or motor processes. They select three visual tasks that have previously established psychometric properties and focus on digit span, spatial span, and spatial rotation, with a multiple-choice response format (i.e., given 3 options and have to select the correct response). The design includes short blocks of trials (7-10 trials) on the same task. Each block begins with a cue card that indicates which task will be performed. Task order is pseudorandomised, as the same task is not repeated across two blocks. The duration of the cue card is variable, the duration of each target card is not specified, but the response grid appears after variable delay. Once the participant selects one of the three options, the others disappear and after another variable interval, the next trial appears. The first data set shows significant accuracy and RT differences between tasks, and a significant 'switch cost' on accuracy but not RT. The second data set is run with a fixed cue duration (0.5 secs) and the effect of the task is found again, but no 'switch cost' in either accuracy or RT.

The paper is generally very well-written, and the data are clearly presented with good quality, analytical figures. However, I found that there are many weakly justified decisions in the paradigm development that question whether the paradigm can produce data comparable to the vast volume of task-switching literature available and therefore weaken the potential impact of the paper. My concerns are mainly related to the fact that the design of the task-switching paradigm, the choice of tasks, the timing parameters, the response options, and the conditions are not consistent with any of the multiple paradigm structures available. Moreover, the paradigm itself

did not produce robust switch effects, which questions its usefulness as a task-switching paradigm.

There are many well-established variants of the task-switching paradigm (e.g., Grange and Houghton¹, Jamadar *et al.*², Karayanidis and McKewen³). These produce somewhat different ways to measure switch costs and additional measures of interest. The paradigm used here is probably most similar to that used by Allport *et al.*⁴, but with significant differences to make comparisons difficult. The paper does not actually specify how 'switch cost' is measured. An example trial sequence is: Task A: 1 2 3 4 5 6 7, Task B: 1 2 3 4 5 6 7 8 Task C: 1 2 3 4 5 6 7, etc. I can only assume that trial 1 in each block is considered the 'switch' trial. However, it's not clear which of the other trials are used as the 'repeat' trial, e.g., is it trial 2, the average of all other trials, etc. This will impact the value of the repeat trial, and therefore switch cost estimation.

Allport and Wylie⁵ showed that these paradigms produce both switch and restart costs, as the first trial of each block produces poorer performance whether the task switches or not. Without occasional repeat blocks, e.g., Task A A B C C B A, it is impossible to differentiate between the switch and restart costs. In addition, this paradigm does not allow estimation of mixing cost (the difference between repeat trials in a single-task vs a mixed-task block), a measure that has been found to be more sensitive to variability in some groups (e.g., ageing, see Karayanidis and McKewen³, Karayanidis *et al.*⁶).

Another major concern is the timing parameters. The paradigm includes very slow trials – the exact inter-trial interval is not specified but appears to be in the order of seconds – and substantial jitter. Monsell and colleagues^{7,8} have shown that timing, jitter, etc can result in measurable changes in outcomes. Ruge *et al.*⁹ reviewed the fMRI task-switching literature and concluded that timing variations used to adapt the paradigm to fMRI timing can significantly impact the cognitive control processes being activated.

Finally, the tasks themselves are very different from the tasks typically used in task-switching paradigms. In previous studies, tasks tend to involve a simple visual stimulus (e.g., letter, number, shape) that requires a 2-choice decision with a discrete response associated with that decision on each trial (e.g., is the letter a vowel or a consonant, press left for vowel, right for consonant). Here, the stimulus is a matrix and different types of exemplars appear over time. So, the stimulus involves a sequence of processes that evolve over time (how long?). The response is not the result of a discrete decision associated with that target. It involves the outcome of a process of comparing three different matrices against the representation of the stimulus held in working memory and selecting which is the closest match. The set of cognitive operations involved is very different to those in typical task-switching paradigms, and the memory-related operations are likely to drown out any effects of task-switching that only apply to the first trial of each block. This may account for the weak switch effects reported. Moreover, any such task-switching processes are not tightly timed to an event (e.g., cue) so will not be readily targeted by event-related fMRI or EEG measures. In addition, this paradigm is likely to produce a lot of motion, especially eye movements, which may create artefacts for fMRI as well as EEG.

Overall, the aim of the paper is sound, the approach well-executed, and the concept of designing paradigms that can be used across platforms and labs is highly commendable. The paradigm appears suited to investigating the cognitive and neural processes engaged by three different visual attention tasks – spatial span, digit span and visual rotation. However, in my opinion, it is

not suited to measuring the cognitive and neural processes involved in task-switching.

References

1. Grange J, Houghton G: Task Switching and Cognitive Control. *Oxford University Press*. 2014.
2. Jamadar S, Thienel R, Karayanidis F: Task switching processes In Toga A. W. (Ed.), Brain mapping: An encyclopedic reference. *Cambridge, MA: Academic Press*. 2015. 327-355
3. Karayanidis F, McKewen M: More than “just a test”—Task-switching paradigms offer an early warning system for cognitive decline. In K. D. Federmeier (Ed.), The psychology of learning and motivation. *Elsevier Academic Press*. 2021. 141-193
4. Allport A, Styles EA, Hsieh S: Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà and M. Moscovitch (Eds.), Vol. XV. Attention and performance. *Cambridge: MIT Press*. 1994. 421-452
5. Allport DA, Wylie GR: Task-switching, stimulus-response bindings, and negative priming. In S. Monsell and J. Driver (Eds.), Control of cognitive processes: Attention and performance XVIII. *Cambridge, MA: MIT Press*. 2000. 35-70
6. Karayanidis F, Whitson LR, Heathcote A, Michie PT: Variability in proactive and reactive cognitive control processes across the adult lifespan. *Front Psychol*. 2011; **2**: 318 [PubMed Abstract](#) | [Publisher Full Text](#)
7. Monsell S, Mizon GA: Can the task-cuing paradigm measure an endogenous task-set reconfiguration process?. *J Exp Psychol Hum Percept Perform*. 2006; **32** (3): 493-516 [PubMed Abstract](#) | [Publisher Full Text](#)
8. Rogers R, Monsell S: Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*. 1995; **124** (2): 207-231 [Publisher Full Text](#)
9. Ruge H, Jamadar S, Zimmermann U, Karayanidis F: The many faces of preparatory control in task switching: reviewing a decade of fMRI research. *Hum Brain Mapp*. 2013; **34** (1): 12-35 [PubMed Abstract](#) | [Publisher Full Text](#)

Is the rationale for developing the new software tool clearly explained?

Yes

Is the description of the software tool technically sound?

Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?

Partly

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?

No

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?

No

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Cognitive control processes, task-switching paradigm, cognitive ageing, behavioural and EEG measures. I do not have sufficient expertise to evaluate the software code.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Author Response 19 Jul 2022

Danielle Kurtin

Response to Prof Frini Karyanadis's Peer Review Report

Prof Karyanadis provided a thorough summary of our manuscript and its intent to introduce a novel task switching paradigm for use in online or neuroimaging studies. She captured the careful design we implemented to reduce visual and motor confounds, the discrete nature of the visual working memory tasks the paradigm switches among, as well as the results of our two Use Cases.

We thank her for the careful reading of the manuscript and for the valuable comments. We appreciate her main concerns centre around our task design and how it diverges from existing literature, the task's timing, the study's results, and considerations on how we define and evaluate switch costs. We have addressed Prof Karyanadis's concerns in our new version of the manuscript and have responded to each concern below.

Our paradigm's design is different from the existing literature

Our task's different experimental design draws three concerns from Prof Karyanadis - that its novelty and difference make comparison to existing literature difficult, that its design cultivates memory-related processes rather than switching processes, and that switch events are not tightly timed to a cue. In this section, we address each concern in turn.

We understand Prof Karyanadis's concern that it may be difficult to compare our paradigm to existing task switching literature that utilizes standard task switching frameworks. This paradigm is intentionally different from most task switching paradigms so one can study research questions traditional task switching paradigms are not well-suited to investigating - in our case, evaluating the neural correlates of large set shifts. In the likely event there are other researchers who also need a new paradigm to meet the needs of their research question, we created our paradigm in an easily modifiable framework so that fewer entirely new paradigms need to be created. We acknowledge we may not have made this paradigm's differences and their impact clear enough. To ensure readers are aware that results from our experiment may not generalize to task switching studies with more standard designs, we have added the following sentences (in **bold**) to the Introduction: *"This invites the follow-up question "How does the brain reconfigure between these different brain states?" The literature here is sparser, with a lack of studies that model how neural networks reconfigure when transitioning from one discrete task to another (referred to as "set switching" or "context switching" (Kim et al 2012). In future studies we hope to characterize the trajectory*

neural networks take to effectively switch between tasks. Therefore, we created a cued task switching paradigm that aims to generate a behavioral and physiological switch cost for use in experiments that will characterize, model, and modulate the switch cost. We chose three psychometrically opposed tasks from Soreq et al (2021) to force distinct reconfiguration from one brain state to the next. **Rather than switching between stimulus response mappings or rules, our task switches between entire task sets, similar to Allport et al (1994). Yet our task differs from Allport's set-shifting task by shifting among different, psychometrically opposed working memory tasks, rather than rules or stimuli within a task. These differences were introduced with the aim of inducing large set shifts observable by fMRI, where future studies may explore how neural networks reconfigure to meet the demands of different working memory tasks.**"

We have also added another sentence in the Limitations section, to ensure readers consider the difference in our paradigms to most others in their interpretation of results and application of this paradigm to other studies: *"Finally, our task differs from most variants of task switching paradigms, and this should be taken into consideration when comparing results from this task to literature using different paradigms."*

Prof Karyanadis's second point centres around a concern that the paradigm probes memory, rather than switch-related processes. We agree that these shifts in working memory processes are different than the common 2-choice decision tasks, both in terms of the number of cognitive operations required and the three, rather than two, response options. Nevertheless, we intentionally switch between three psychometrically opposed working memory tasks because we wanted to investigate the neural correlate of large set shifts, which could not be possible using more standard task switches, such as stimulus-response or rule switches.

Regarding Prof Karyanadis's third point that there is not an event to which we can identify the task switch events, we would like to highlight that our Cue Cards precede each switch, and are timed cues by which we can localize task switching events. Figure 1 provides a visual depiction of how a Cue Card precedes a set of trials that build into a run of a task before another Cue Card announces a switch to the next run of trials of a different task.

Task timing

We understand Prof Karyanadis has two concerns regarding task timing, with the first being that we do not specify a precise intertrial interval (ITI). We do not specify a specific ITI because, as described in the Neuroimaging studies section, the ITI varies randomly between 100 to 1100 msec. Should future researchers wish to change this, the task was created to be easily customizable, and we remark in several places that all task parameters (such as the ITI) can be changed to suit the researcher.

Second, we understand Prof Karyanadis's concern about the influence the variable ITI exerts on cognitive control processes or fMRI analysis. To ensure readers are aware of the impact of adding a variable ITI, we have included the following sentence: *"For online studies, we recommend researchers include a jitter-induced delay related regressor in models of BOLD activity, and remove jitter during offline studies as recommended by [1] for response-cue trial*

intervals." in the Neuroimaging study section.

Study results

We acknowledge Prof Karyanadis is unsure whether a task switching paradigm qualifies as such if it does not induce switch effects. As with any study, the results are the one aspect of an experiment that cannot be controlled. We argue that the absence of positive results does not invalidate the design of the paradigm.

We also would like to highlight that we did pilot our work to identify suitable task parameters that induce a switch cost. Before conducting any Use Cases, we determined which parameters would ensure above-chance and below-ceiling performance in each task. For example, literature has shown most healthy young adults can remember 7 digits in a digit span [2], but we use 6x6 grids for the working memory tasks, meaning that the number of stimuli participants are asked to remember for the digit span should be within their ability. Early piloting showed that, when 6 stimuli were presented at the standard 0.5 seconds per stimuli, participants exhibited ceiling effects. We, therefore, shortened the stimuli presentation time for digit span stimuli to 0.25 seconds, with the intent to standardize performance across tasks.

With these parameters, we began conducting Use Case 1. In Use Case 1, the cue cards preceding each task were either 4.0 or 0.5 seconds. We observed a switch cost in accuracy but not reaction time, and, based on literature showing short cue cards induce switch costs [3][4], we chose to standardize our cue card duration to 0.5 seconds in an attempt to induce a switch cost in both accuracy and reaction time. Though we did not observe a switch cost after standardizing our cue card duration, in our discussion for the second Use Case we suggest that our negative results may be due to an interplay between endogenous, preparatory processes and exogenous, cue-driven processes [5]. Because of the customizable nature of the paradigm, we encourage future researchers to pilot the task using parameters (cue cards, length of time stimuli are presented, the number of tasks, etc) suitable for their experiment and assess the presence of a switch cost.

Switch costs

We thank Prof Karyanadis for identifying that we did not fully specify how we measure switch cost. We have clarified this in the manuscript in the Use Case 1 section as follows: *"Switch task types are defined as the first trial after a switch between tasks, and stay trial types are all other trials."*

We appreciate that our task design does not permit a comparison of switch vs restart costs, nor mixing costs, and how this may have influenced our ability to induce switch costs. We have added the following sentences to our Limitations section to make this clearer to readers: *"Though the piloting of this task was performed with healthy control participants, future researchers may want to assess differences between healthy control and patient populations. Mixing costs may be more sensitive to between-group variability [6], and one limitation of our task structure is that it does not permit the exploration of mixing costs. Moreover, the current design did not allow repeats blocks of the same task (for example, this order would not occur:*

Digit Span, Spatial Span, Spatial Span, Spatial Rotation), and therefore cannot investigate the difference in switch vs restart costs [7]. Future researchers are invited to adapt the paradigm's design to allow repeat task blocks, to investigate switch vs restart costs, and explore whether introducing mixed-task blocks induces switch costs not seen in this version of the paradigm."

Finally, to make it abundantly clear to readers that our paradigm does not consistently introduce a switch cost, we have included the following sentence at the end of the Software Tool Article's Introduction: *"The pilot studies observe that, though the two versions of the task do not consistently induce a switch cost, the paradigm operates within an optimal difficulty range, and participants do not exhibit learning effects. Though our task does not produce traditional switch costs, we believe this paradigm is useful given its highly adaptable, multi-modal, open-source nature."*

We hope this response addresses Prof Karyanadis's concerns surrounding our Software Tool Article Manuscript. We appreciate her insight, and the changes we have made as a result of her review have improved and strengthened the manuscript.

References:

- [1] Ruge H, Jamadar S, Zimmermann U, Karayanidis F: The many faces of preparatory control in task switching: reviewing a decade of fmri research. *Human brain mapping*. 2013; 34(1):12–35
- [2] Hampshire A, Highfield RR, Parkin BL, et al.: Fractionating human intelligence. *Neuron*. 2012; 76(6): 1225–1237
- [3] Monsell S, Mizon GA: Can the task-cuing paradigm measure an endogenous task-set reconfiguration process? *Journal of Experimental Psychology: Human perception and performance*. 2006; 32.3(493).
- [4] Schneider DW, Logan GD: Task-switching performance with 1: 1 and 2: 1 cue–task mappings: Not so different after all. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2011; 37(2):405.
- [5] Periañez JA, Barceló F: Updating sensory versus task representations during task-switching: Insights from cognitive brain potentials in humans. *Neuropsychologia*. 2009; 47(4): 1160–1172.
- [6] Karayanidis F, Whitson LR, Heathcote A, Michie PT: Variability in proactive and reactive cognitive control processes across the adult lifespan. *Frontiers in psychology*. 2011; 2:318.
- [7] Allport DA, Wylie GR: Task-switching, stimulus-response bindings, and negative priming. In S. Monsell and J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII*. Cambridge, MA: MIT Press. 2000. 35-70

Competing Interests: We have no competing interests to disclose

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research