

The optimisation of functional brain network dynamics when learning to perform complex tasks

Parkin, B.L, Daws, R.E., Scott, G., Soreq, E., Leech, R., Hellyer, P.J. and Hampshire, A.

NOTICE: this is the authors' version of a work that was accepted for publication in Neuropsychologia. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Neuropsychologia, volume 216, 9 September 2025, 109188.

The final definitive version in Neuropsychologia is available online at:

<https://doi.org/10.1016/j.neuropsychologia.2025.109188>

© 2025. This manuscript version is made available under the CC-BY-NC-ND 4.0 license

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

The WestminsterResearch online digital archive at the University of Westminster aims to make the research output of the University available to a wider audience. Copyright and Moral Rights remain with the authors and/or copyright owners.

Title: The optimisation of functional brain network dynamics when learning to perform complex tasks

Short running Title: Network dynamics of complex task learning

Authors:

B, L Parkin^{*1}, Richard E. Daws^{2,3*}, Gregory Scott³, Eyal Soreq³, Robert Leech², Peter J. Hellyer² & Adam Hampshire^{2,3}.

*Joint first authors

Affiliations:

¹Department of Psychology, School of Social Sciences, University of Westminster, London, UK

²Department of Neuroimaging, Institute of Psychiatry, Psychology and Neurosciences, King's College London, London, UK.

³Department of Brain Sciences, Imperial College London, UK.

Corresponding authors:

Name: Adam Hampshire. Email: a.hampshire@imperial.ac.uk

Name: Beth Parkin. Email: b.parkin@westminster.ac.uk

Acknowledgements:

We would like to thank the participants who took part in this study

Abstract

We present two novel self-ordered switching (SOS) fMRI paradigms designed to investigate how brain networks facilitate the establishment of structured human behaviour during the learning of complex tasks with multiple goals. In study 1, SOS was performed with minimal pretraining and detailed feedback to capture the learning process, while in study 2 substantial pretraining and minimal feedback were used as a control where the potential for ongoing optimisation of behaviour is reduced. Study 1 revealed changes in the learning process characterised by a decrease in task-switching decision frequency, resulting in superior task performance for individuals who minimised switch frequency and ordered their behaviour in simple structured routines. Additionally, with practice, multiple-demand cortex activation became less responsive, and the default mode network became more responsive when performing discrimination trials. Strikingly, the opposite pattern was observed for SOS events, with multiple-demand cortex activation becoming more responsive and default mode network activation becoming less responsive with practice. These neural changes correlated with the degree of structure of behavioural routines. The neural signatures of learning were less evident in study 2, where the task was practiced prior to entering the scanner. Our studies demonstrate that the default mode network and multiple demand cortex complement each other when people learn to perform complex tasks by becoming differentially fine-tuned to discrimination task trials vs. executive-switching demands.

Key words: cognitive control, learning, behavioural structure, task-switching, fMRI, multiple-demand network, default mode network.

Highlights

We examine how brain-network resources are optimised during a novel Self-Ordered Switching (SOS) paradigm.

Reaction times decreased as SOS was practised, with overall task performance correlating with low algorithmic complexity routines.

As the task was learnt, activation in multiple-demand regions decreased for discrimination trials but increased for SOS events, whereas the default mode network showed the opposite relationship.

Introduction

The ability to rapidly learn complex tasks is a definitive feature of human cognition. Studies of cognitive control suggest that the human brain achieves this by quickly adapting to encode behavioural routines (Finc et al., 2020; Khambhati et al., 2018), also known as a ‘task set’, comprised of the combinations of rules, strategies and stimulus-response mappings that enable optimal performance (Sakai, 2008). This encoding is typically associated with distributed activation of the multiple-demand (MD) system (Duncan, 2013).

During instruction-based learning (IBL), generating and establishing simple routine behaviour imposes an initial cognitive demand that rapidly decreases with practise. This learning, quantified by a reduction in reaction times, is accompanied by a decay in MD activation, especially in the anterior prefrontal cortex (aPFC) (Hampshire et al., 2016, 2019). This occurs alongside changes in other brain networks, particularly the default mode network (DMN) which is initially deactivated in the early stages of learning and returns to baseline as behavioural routines become established (Hampshire, 2019).

Importantly, in everyday life an optimal behavioural routine is rarely evident when first performing novel tasks. This can be due to a lack of clear instruction, the inherent complexity of the task, or shifting demands. Consequently, it is necessary to monitor outcomes and progressively update routines to optimise behaviour (Sharp et al., 2010; Swick & Turken, 2002; Ionescu, 2012). This updating of routines is costly and is associated with increases in reaction time and heightened MD activation (Hampshire & Owen, 2006).

In addition, it is often necessary to juggle multiple concurrent goals, or to manage sub-routines when performing complex tasks, and in the process, to monitor not only one, but multiple outcomes. Increased task complexity poses a challenge for flexible cognitive systems due to their intrinsic limitations in capacity (Duncan, 2013). Under such conditions, humans tend to impose hierarchical structure to their behavioural routines (Fallon et al., 2013).

This tendency towards self-ordered structure is underpinned by decisions regarding when and in what order to switch attention between competing tasks. Consequently, task switching paradigms provide a natural choice for studying the cognitive control processes that enable flexible human behaviour. However, these paradigms typically dictate when switches occur (Kiesel et al., 2010) and do not require participants to simultaneously manage sub-goals within each task; thus, they do not capture the self-ordered nature of real-world behaviour (Hampshire & Owen, 2006).

Despite these limitations, there is evidence that cognitive demand is minimised with learning, including by structuring behaviour where possible. For example, task-

switching imposes a considerable cognitive demand but switching costs shown by reaction times decrease with practise (Yeung & Monsell, 2003), although a residual cost persists even after tens of thousands of trials (Stoet & Snyder, 2007). Furthermore, switching costs decrease with preparation time and when switching events are predictable (Rogers & Monsell, 1995).

During voluntary task-switching (VTS), participants are less likely to switch under high cognitive load (Demanet et al., 2010) and can exhibit a repetition bias despite being instructed to switch between tasks evenly and randomly (Arrington & Logan, 2004, 2005). These findings indicate that a key characteristic of how humans learn complex tasks is the application of structured behavioural routines that minimise the cognitive load associated with switching.

Here, we use two novel event-related fMRI Self Ordered Switching (SOS) paradigms to investigate the brain mechanisms by which people learn complex multi-component tasks, with a particular focus on examining voluntary switching. Specifically, the SOS paradigms require participants to learn through a process of feedback-driven trial and error how to manage their time between two tasks, each of which has two rules, where to score maximum points, the participant must elect to spend time on all four tasks. One paradigm requires participants to establish an optimal routine based on ongoing feedback (study 1). The other paradigm, performed by a different set of participants after receiving extensive pre-scanner training, has minimal in-scanner feedback (study 2), thereby acting as a control where the opportunity for ongoing optimisation of behaviour is reduced.

First, we examine how the behaviour differs between ongoing reinforced learning and pre-learned conditions. Then we examine how the response of task active and default mode brain networks respond to discrimination trials and switch decisions, and how these responses under those same conditions change with practice. Finally, we test whether behavioural structure correlates with changes in functional network dynamics across time and across participants, and undertake exploratory analysis to determine the degree to which those changes predict individual differences in overall task performance scores.

1. Materials and Methods

2.1 Participants

31 healthy right-handed participants were recruited in total: **Study 1** - n=16, mean age 24.75 years, range 19-40, 10 female; **Study 2** - n=15, mean age 26.47 years, range 19-42, 11 female. Approval for this study was received by the Cambridge university research ethics committee and participants gave written informed consent. Participants were reimbursed £20 for their time.

2.2 Self-ordered switching (SOS) paradigm

The SOS paradigm required participants to manage their time, between two concurrent tasks each with two subordinate rules, with the goal to gain maximum points (Figure 1).

There was a matching task (MT) and an odd-one-out (OOO) task, both involving simple visual discriminations of colour-shape objects (Figure 1a). The subordinate rules of each task required participants to apply the task rules to either the dimensions of colour or shape.

When performing the matching task, a target object was presented on screen above two probe objects. The participant indicated, via a left or right button press, which of the two probe objects matched the target according to the currently selected shape or colour rule (Figure 2c).

Similarly, during the trials of the odd-one-out task, visual arrays of objects arranged in a 2x2 grid were presented. The participant indicated which side of the display contained an object that differed from the others according to the currently selected colour or shape rule. Again, a left or right button press was used to indicate their responses. In each case, trials stayed on screen until a response was made, and all responses were made with a button box positioned under the right hand (Figure 2f).

Crucially, to keep gaining points participants had to organise their time across all four sub-tasks (more details below). This meant that after each trial, participants were required to choose whether to repeat the same condition (repeat- discrimination) or to switch either between tasks or rules (task-switch discrimination, rule-switch discrimination) (Figure 2d-e). There were no changes between trials unless the participant opted to make a task-switch decision or rule-switch decision by pressing the up or down button respectively. Participants could not make a task and a rule switch simultaneously.

The paradigm was designed so that the task and rule discriminations were simple enough to be performed with high accuracy and minimal effort; therefore, individuals could focus on the executive requirement of learning to divide their time across the four tasks to maximise gaining points.

Two studies were performed with different variants of the SOS paradigm to capture different aspects of the learning process (learning the task- in study 1, versus executing the task – in study 2). Therefore, the studies differed with respect to the amount of pre-scanner training, task presentation and the feedback. In both cases the paradigm was continuously performed inside the scanner in a single 20-minute acquisition block.

In study 1, participants performed only a short practice prior to scanning so that learning process related activity could be observed in the scanner (Figure 1b). In addition, feedback was presented during task performance via a comprehensive set of indicators so that participants could use these to optimise their behaviour.

The initial demo outside the scanner presented on-screen instructions to orient the participant towards the feedback elements of the task and button presses required. Participants then undertook each combination of the four task-rule discrimination trials and each of the switch decisions once (rule or task switch).

Inside the scanner, both tasks were displayed on the screen with the matching task at the top and the odd-one-out task at the bottom (Figure 1c). The currently active task was highlighted with a red border and the inactive task had a grey border - these flipped when the task-switch button was pressed. The active rule was displayed with a text cue in the centre of the screen ("COLOUR" or "SHAPE", in yellow) and flipped when the rule-switch button was pressed.

On each trial, new stimuli were displayed and participants were able to repeat the condition or could opt to task-switch or rule-switch. Stimuli stayed on screen until a button press was made, following which a feedback array was displayed for 1.6 seconds, immediately after this the next trial was presented (Figure 1c).

Participants learnt to optimally sequence their responses based on a set of performance meters (Figure 2a-b). Presented on screen continuously was an overall score located at the centre of the screen alongside two meters indicating the number of points currently accumulated for each task (OOO task & MT, Figure 2b). In addition, after each trial a feedback screen was presented which comprised of text indicating whether the participants response was correct or not ("CORRECT" in green or "INCORRECT" in red), and four meters showing points currently accumulated on each sub-task (i.e., MT-colour rule; MT-shape rule; OOO task -colour rule; OOO task - shape rule) (Figure 2a).

Importantly, the scoring of the sub-tasks interacted hierarchically. Specifically, within each task the colour or shape meters would increase after a correct response to the corresponding condition and would decrease if an incorrect response was made. All meters ranged between 1-10 and were initialised with 5 points at the start of the experiment.

To gain maximum points, within each task the aim was to respond to, and switch between, the colour and shape rules in a sequence that would build up both meters to 10 points (Figure 2c-e). To encourage switching, if either the colour or shape meter got further than 2 points ahead, then the other meter decreased by a point (Figure 2d). When both the colour and shape meters reached the maximum level of 10, they would reset to 5, and the corresponding task and overall score incremented by 1.

Comparably, switching between tasks was encouraged in a similar way in that the inactive task meter would decrease if the active task meter was 2 or more points ahead. Successfully dividing time between each task would increase both task meters, thereby increasing the overall score. When both task meters reached the maximum level of 10, they would reset and a +10 point bonus was awarded to the overall score (Figure 2f).

Given that switching would be cognitively costly in terms of both effort and time, to gain maximum points optimal behaviour involved a routine that used the minimal number of switches to ensure sufficient dividing of time between each task and between colour shape sub-rules. Those who learnt the most efficient way of doing this, via trial and error, would be expected to score the most points during the experiment. An example optimal sequence is shown in Figure 3e.

In study 2, participants performed the SOS paradigm with the same stimuli, rules and tasks; however, there were modifications in the presentation of the tasks, extensive pre-training and no feedback during task performance, to ensure that learning primarily occurred prior to entering the scanner.

Most importantly, prior to the scanning session participants received the same initial demo as in study 1 and then performed a 20-minute training session. This additional training session was designed to enable participants to arrive at a structured routine for performing the task prior to the scanner session as it involved performing the task described in study 1 with feedback scales, so that participants could learn to optimise their behaviour to gain maximum points (Figure 1d).

Importantly, the task and colour and rule sub-task meters were only present during the training. Whilst in the scanner, the only feedback available was the overall score (Figure 1e). There were also no +10 bonus points awarded for increasing both task meters to their limit. New trials were presented after a delay with no stimulus being shown, this lasted for the same fixed duration as the feedback from study 1 (1.6 seconds). These modifications were designed to minimise the information that the participant could use to further optimise their behavioural routine inside of the scanner whilst ensuring that they had sufficient information to remain engaged.

Finally, as opposed to presenting both tasks on screen simultaneously (as in study 1), only the currently selected task was presented on the screen (Figure 1e). This modification was intended to further control for activity related to switching between spatial locations as opposed to tasks.

1.3 Behavioural Analysis

2.31 Response times and accuracy

The response times (RTs) of three trial types were analysed: repeat, rule-switch, and task-switch discriminations. Response times for a **repeat-discrimination trial** was the time between the stimuli presentation and the left or right response being made. **Rule-switch** or **task-switch discrimination** response times were measured as the time between the switch button being pressed and the subsequent response to the new trial. For each participant, response times for each trial type were summarised using the median from correct trials. The number of trials completed would naturally vary across participants.

Response accuracy was also calculated for each trial type as the percentage of correct responses. Separate one-way repeated-measures analysis of variance (RM-ANOVA) were conducted for response times and accuracy.

2.32 Algorithmic complexity

Trial conditions could be repeated or switched between as each participant saw fit. To quantify the extent to which participants structured their behaviour, we estimated the algorithmic complexity (Zenil et al., 2018) of each behavioural timecourse.

Algorithmic complexity estimates the size of the smallest algorithm that could generate a given sequence and tends to be smaller for sequences with consistent structure vs. random sequences. Here, we use the framework, outlined by Zenil et al. (2018), which combines a block decomposition method (BDM) and the coding theorem method (CTM) to provide an estimate of Kolmogorov-Chaitin complexity (K) (Chaitin, 1966; Kolmogorov, 1968) for large strings in a computationally tractable way. The advantage of BDM is that it provides an estimate that not only considers statistical regularities, (e.g., as is captured by Shannon entropy), but it is also sensitive to segments of an algorithmic nature. Timecourse length naturally varied across individuals, and to account for this, we normalised an individual's algorithmic complexity by the mean algorithmic complexity calculated from 100 randomly shuffled versions of the individual's data, which are expected to produce high algorithmic complexity due to their randomness.

A more focused behavioural analysis was conducted to examine how response times and algorithmic complexity varied over the course of the 20-minute experiment. Here, each participant's behavioural timecourse were split into three uniform non-overlapping time-windows, or 'learning stages' of 400 seconds each. This provided early, mid and late timeseries of sufficient duration to estimate behavioural complexity. Specifically, for each learning stage the median response times for each trial type

(correct trials) as well as the normalised algorithmic complexity (using the same procedure outlined above) were extracted.

Repetition bias was also calculated as a simpler measure of behavioural structure by taking the proportion of consecutive trials from the same condition relative to the total number of trials completed.

2.4 fMRI analysis

2.41 Acquisition and preprocessing

Whole-brain blood oxygenation level dependent (BOLD) fMRI recordings were acquired using the same Siemens 3T Trim Trio scanner and following parameters in both studies: TR=2s, TE=30ms, FA=78°, 3mm³ voxels, 3.75mm slice-gap, 32 axial slices. 630 volumes were acquired during the 20-minute SOS task run. A T1-weighted magnetization prepared rapid acquisition gradient echo (MPRAGE) sequence was also acquired using the following parameters: TR=2250ms, TE=2.99ms, IT=900ms, FA=9°, 1mm³ voxels, 256 axial slices.

fMRI pre-processing was conducted using standard parameters with the Statistical Parametric Mapping 12 (SPM) toolbox in MATLAB 2018b. The first 10 volumes were discarded to account for T1-equilibrium effects. The remaining images were slice-time and motion-corrected, and the mean echo planar image (EPI) co-registered to the tissue segmented T1 image, normalised to 2mm³ MNI space and spatially smoothed (full half width maximum = 8 mm kernel).

2.42 Activation modelling

General Linear Models (GLM) were constructed from predictor functions and a constant term at the individual subject level. These consisted of the onset and duration of task events convolved with the canonical haemodynamic response function. Four predictor functions related to the responses of the different discrimination trial types (MT-colour rule, MT shape-rule, OOO task colour rule, OOO task shape rule). The onset for these trials were modelled as the time at which participants responded to the trial with a button press, with no duration being defined. Six predictor functions related to the switching decisions, two of which captured the decision to switch task (MT to OOO and OOO to MT), and the remaining four captured the decisions to switch rule (from MT colour to MT shape, from MT shape to colour, from OOO colour to shape, from OOO shape to colour). The onset for these trials were modelled as the time at which the participant made a button press indicating a switch, with no duration being defined. Errors were infrequent but were captured in an additional predictor. Models in study 1 included an additional three predictor functions quantifying feedback events,

two of these were when the rule or task performance meters changed and the other captured the moment when bonus points were awarded.

Head motion and nuisance variables were modelled using the rigid-body realignment parameters, the mean timecourses from the white matter and cerebrospinal fluid tissue segmentations, along with their 1st order temporal derivatives. This resulted in 16 nuisance variables. Additional spike regressors were included where framewise displacement (Power et al., 2012) exceeded the voxel acquisition size (3mm).

The 1st-level beta maps were estimated for each psychological predictor relative to the model intercept (as there were no resting periods). Modelling against this “implicit baseline” results in significant activation differences being relative to the average activation across time.

An additional set of 1st-level GLM models were created to examine how activation varied over time as behaviour became more practiced. Specifically, the predictors relating to the different trial types were broken down into three consecutive learning stages, i.e., aligned to those defined in the behavioural analysis. All discrimination task trials were collapsed into a single predictor, and switching decision events were collapsed into a single task-switch decision and a single rule-switch decision predictor. The same nuisance predictors were used in these models.

The group-level modelling of activation associated with switching decisions involved a voxelwise one-way RM-ANOVA of the beta maps representing the two task-switch and four rule-switch decisions. To examine how switching activation varied over time, switch decision>discrimination trial contrasts were generated from each learning stage and modelled with a group-level two-way RM-ANOVA with learning stage (3 levels) as the within-subject factor and study (2 levels) as a covariate. All group-level activation maps were initially thresholded at an uncorrected $p < 0.01$, followed by a $p < 0.05$ false discovery rate (FDR) cluster correction.

Our imaging analysis therefore focuses on activation changes relating to making a decision to switch tasks or rules, rather than the cost-related activity of the switch when responding to a discrimination trial which can not be delineated in this study.

3.Results

3.1 Behavioural analysis

SOS costs to reaction time decrease with practise

As intended, trials were performed with near-ceiling accuracy (**Study 1** - mean=94.87%, SD=4.18; **Study 2** - mean=94.69%, SD=4.04) and accuracy did not

vary across the repeat discriminations, rule-switch or task-switch discriminations (one-way rm-anova: **Study 1** – $F_{(2,30)}=1.218$, $p=0.310$, $\eta^2=0.08$; **Study 2** – $F_{(2,28)}=1.867$, $p=0.173$, $\eta^2=0.12$).

Response times varied substantially between the repeat- discriminations, rule-switch and task-switch discriminations in both studies (one-way rm-anova: **Study 1** - $F_{(2,30)}=48.588$, $p<0.001$, $\eta^2=0.76$; **Study 2** – $F_{(2,28)}=52.886$, $p<0.001$, $\eta^2=0.79$).

Specifically, in study 1 performing discriminations following task-switching slowed response times on average by 883ms relative to repeat-discriminations (SD=538) (paired t-test: $t_{(15)}=6.573$, $p<0.001$, confidence interval 95% (CI)=0.60 to 1.17, Cohen's $d=1.64$), while discriminations following rule-switches were 179ms (SD=273) faster relative to repeat-discriminations (paired t-test: $t_{15}=-2.630$, $p=0.019$, CI=0.03 to 0.33, $d=0.66$) (FDR corrected) (Figure 3a). In study 2 discriminations following rule switching slowed response times by 267ms (SD=143) relative to repeat discriminations (paired t-test: $t_{(14)}=7.214$, $p<0.001$, CI=0.19 to 0.35, $d=1.86$), and discriminations following task switching slowed response times by a further 362ms (SD=268) ($t_{14}=5.227$, $p<0.001$, CI=0.21 to 0.51, $d=1.35$) (FDR corrected) (Figure 3b).

A two-way repeated-measures analysis of variance (RM-ANOVA) was used to determine whether the number repeat and switch discrimination trials differed across studies. Trial type was the within-subject factor (repeat-discrimination trial, rule-switch discrimination, task-switch discrimination) and study was the between-subject factor (2 levels). There was a significant interaction of number of trials with study ($F_{(2,58)}=62.128$, $p<0.001$; $\eta^2=0.68$). Participants in study 1 performed fewer trials overall (two-sample t-test: $t_{(29)}=-5.262$, $p<0.001$, CI=-81.87 to -36.04, $d=1.38$) and performed more discriminations following rule-switches ($t_{(29)}=8.125$, $p<0.001$, CI=49.85 to 83.39, $d=1.64$) than participants in study 2. The number of discriminations following task-switches did not differ between studies ($t_{(29)}=0.024$, $p=0.981$, CI=-11.426 to 11.69, $d=0.01$).

To examine how switching-costs changed with learning, the trials were segmented into the three consecutive time-windows, or 'learning stages'. Switching-costs for each learning stage were calculated by subtracting the mean repeat-trial response times from the task-switch or rule-switch discrimination response times. Task-switching costs varied over learning stages in both studies (one-way RM-ANOVA: **Study 1** – $F_{(2,30)}=3.579$, $p=0.040$, $\eta^2=0.19$; **Study 2** – $F_{(2,28)}=5.413$, $p=0.010$, $\eta^2=0.28$), reducing between the first and last stage by 369ms (SD=607) in study 1 (paired t-test: $t_{(15)}=-2.433$, $p=0.028$, CI=0.05 to 0.69, $d=0.61$), and by 278ms (SD=418) in study 2 ($t_{(14)}=-2.570$, $p=0.022$, CI=-0.51 to -0.05, $d=0.66$). Rule-switching costs did not vary significantly across learning stages in study 1 ($F_{(2,30)}=0.128$, $p=0.881$, $\eta^2=0.01$), but did in study 2 ($F_{(2,28)}=5.543$, $p=0.009$, $\eta^2=0.28$) where there was a 217ms reduction (SD=286)

between the first and last learning stage ($t_{(14)}=-2.945$, $p=0.011$, $CI=-0.38$ to -0.06 , $d=0.76$).

In an exploratory analysis, individual differences in switching costs were examined to determine the relationship between response times and the number of switches made. The mean switch cost to response times (calculated by averaging across task and rule switch discrimination trials minus repeat discriminations) showed a significant negative correlation with the total number of switches performed by participants (Pearson: **Study 1** – $r_{(15)}=-0.557$, $CI=-0.83$ to -0.09 , $p=0.025$; **Study 2** – $r_{(14)}=-0.658$, $CI=-0.88$ to -0.22 , $p=0.008$) (Figure 4a-b). Therefore, there were substantial switching response time costs, these switching costs were optimised over time with learning, and individuals who made more switches had lower switch costs.

Response times to repeat discriminations varied between studies and with practise in a manner that indicated learning effects. A two-way RM-ANOVA of repeat discrimination RTs showed main effects of study ($F_{(1,29)}=16.866$, $p<0.001$, $\eta^2=0.37$), learning stage ($F_{(1,29)}=309.770$, $p<0.001$, $\eta^2=0.91$) and a significant study * learning stage interaction ($F_{(1,29)}=14.516$, $p<0.001$, $\eta^2=0.33$). In comparison to study 2, response time to repeat discrimination were significantly slower in study 1 (two-sample t-test: $t_{(29)}=4.119$, $CI = 0.29$ to 0.85 , $p<0.001$, $d=1.20$) and decreased at a greater rate between the first and last learning stages (two-sample t-test: $t_{(29)}=-2.878$, $CI = -0.37$ to -0.06 , $p=0.007$, $d=0.93$).

Task performance relates to structured behavioural strategies

Exemplar behavioural time courses from each study showing differing degrees of algorithmic complexity are presented in Figure 3e-f. Low Algorithmic complexity is characterised by ‘chunks’ of repeat discrimination trials that were consistently punctuated by switches between rules and tasks.

As expected, when normalising algorithmic complexity estimates, the shuffled data exhibited significantly greater complexity (paired t-test: Study 1 – $t_{(15)}=-9.775$, $p<0.001$, $d=3.54$; Study 2 – $t_{(14)}=-35.329$, $p<0.001$, $d=14.29$). Thus indicating that participants were structuring their behaviour.

The normalised algorithmic complexity of individuals’ behavioural routines showed a significant negative correlation with the points scored (partial Pearson: **Study 1** – $r_{(15)}=-0.646$, $CI=-0.87$ to -0.22 , $p=0.017$; **Study 2** – $r_{(14)}=-0.805$, $CI=-0.93$ to -0.50 , $p=0.002$) (Figure 3c-d) and the number of errors made (partial Pearson: **Study 1** – $r_{(15)}=-0.695$, $CI=-0.89$ to -0.31 , $p=0.004$; **Study 2** – $r_{(14)}=-0.785$, $CI=-0.93$ to -0.46 , $p<0.001$). That is, people whose behavioural routines were more structured, achieved higher points and made fewer errors. Here, partial correlations were used to control for the number of events completed as this naturally varied across participants.

Furthermore, the overall number of switches performed did not significantly correlate with the overall points scored (Pearson: Study 1 - rule-switch $r_{15}=0.473$, CI=-0.03 to 0.79, $p=0.065$; **Study 1** - task-switch $r_{15}=0.213$, CI=-0.32 to 0.64, $p=0.428$; **Study 2** - rule-switch $r_{14}=0.160$, CI=-0.38 to 0.62, $p=0.569$; Study 2 - task-switch $r_{14}=0.034$, CI=-0.49 to 0.54, $p=0.905$) (uncorrected). This was also the case when controlling for the number of trials completed (partial Pearson: **Study 1** - $r_{15}=0.407$, CI=-0.11 to 0.75, $p=0.132$; **Study 2** - $r_{14}=0.151$, CI=-0.39 to 0.62, $p=0.605$).

Running a two-way RM-ANOVA on normalised algorithmic complexity scores where learning stage was the within-subject factor (first, middle & last) and study was the between-subject factor (2 levels) showed no significant interaction of study and learning stage ($F_{(1.538, 58)}=1.235$ $p=2.92$). There was a significant main effect of study ($F_{(1,29)}=23.766$ $p<0.001$ $=0.450$), with algorithmic complexity scores significantly higher in Study 1 (mean=0.577) compared to Study 2 (mean=0.300), and a significant main effect of learning stage ($F_{(1.538, 58)}=5.558$ $p=0.012$ $=0.161$). Post hoc tests showed the relationship between learning and algorithmic complexity was non-linear, with a significant reduction between the first (mean=0.472) and second learning stage (mean=0.409) ($p=0.015$ Bonferroni correction), but no significant differences between the first and second relative to the third stage. In accordance with the main effect of task for algorithmic complexity, responses in study 2 exhibited a significantly greater repetition bias than study 1 (two-sample t-test: $t_{(29)}=6.643$, CI=0.19 to 0.36, $p<0.001$, $d=1.53$). Therefore, in study 2 participants were more likely to repeat the conditions than switch between rules or tasks.

3.2 fMRI Analysis

SOS activates multiple-demand cortex and reduces default mode network activation

We sought to identify brain regions where activation increased when individuals opted to switch between tasks or rules (vs discrimination-trials). Collapsing across switch types (rule and task switches), the cluster-corrected group-level t-contrast associated self-ordered switching vs response to discrimination trials with activation increases in the dorsolateral frontoparietal cortices, cerebellum, insula, putamen, and thalamus (Figure 5a). This pattern was consistent across studies (dice similarity = 0.58). The inverse t-contrast revealed greater relative activation in regions associated with DMN during discrimination trials compared to switch decisions (Figure 5b), and this pattern was consistent across studies (dice similarity = 0.56).

Multiple-demand and default mode network activation changes with learning

Previous studies have examined activation changes when simple tasks are established through practice (Hampshire et al., 2016, 2019; Petersen et al., 1998;

Ruge et al., 2019). To expand on this, voxelwise activation was compared across time during SOS. For each of the three learning stages, a 1st-level contrast was estimated for switching decisions minus discrimination trial activation. These contrasts were examined using group level random effects analyses with a 3 * 2 RM-ANOVA design where the learning stage (first, middle & last) was the within-subject factor and study (1 & 2) was the between-subject factor.

The main effect of the learning stage contrast rendered six bilateral regions where the difference between switching decisions and discrimination trial activation significantly differed across the learning stages. For interpretation a post hoc paired t-test was conducted for each cluster, to determine the direction of the activation change using the mean bilateral activation from the first and last learning stages. The regions that showed increased activation were the anterior cingulate ($t_{(30)}=5.267$, $p<0.001$, $CI=1.37$ to 0.61 , $d=1.59$), dorsolateral parietal cortex ($t_{(30)}=4.657$, $p<0.001$, $CI=1.65$ to 0.64 , $d=1.56$) and clusters within the occipital cortices ($t_{(30)}=4.707$, $p<0.001$, $CI=1.42$ to 0.56 , $d=1.57$), we combined these regions into a composite MD region of interest (ROI) for display purposes (Figure 5c).

Conversely, the post hoc t-tests showed that those regions where activation decreased included the posterior cingulate ($t_{(30)}=4.097$, $CI=0.31$ to 0.94 , $p<0.001$, $d=0.74$), putamen ($t_{(30)}=3.412$, $CI=0.23$ to 0.91 , $p=0.0019$, $d=0.62$) and temporal lobe ($t_{(30)}=5.401$, $CI=-1.12$ to 0.51 , $p<0.001$, $d=0.97$). These accord with the DMN and were combined into an ROI for display purposes (Figure 5d).

Notably, MD activation increasing and DMN activation decreasing with learning is the exact opposite direction to what has been observed during simple stimulus-response learning tasks (Hampshire et al., 2019). To understand why this learning effect appeared to be inverted, we examined discrimination trial activation and switching decision activation in the MD ROI across the learning stages in Figure 5e. Here we observe that the discrimination response events are more comparable to simple learning curve paradigms and show increased activity in the first learning stage compared to the last stage ($CI=-0.99$ to -0.38). Conversely for the switching decisions activation shows an increase in the last learning stage, compared to the first stage (Task-switch decision $CI=0.03$ to 0.67 ; Rule-switch decision $CI=0.09$ to 0.66). Interestingly, the opposite pattern was evident for the DMN ROI activation which we present in Figure 5f. Specifically, activation for performance of discrimination trials was increased for the last learning stage compared to the first ($CI=0.11$ to 0.40), while for switching decisions activation was increased in the first learning stage compared to the last (Task-switch decision $CI=-0.73$ to -0.27 ; Rule-switch decision $CI=-0.46$ to -0.20).

Additional exploratory analysis showed that the sum of differences in mean MD activation over learning stage negatively correlated with individuals overall normalised algorithmic complexity scores (Pearson, $r_{(15)}=-0.499$, $p=0.049$, $CI=-0.798$ to -0.004) in study 1 (Figure 5g) but not in study 2 ($r_{(14)}=-0.240$, $p=0.932$, $CI=-0.53$ to 0.49).

Therefore, participants with more structured behavioural routines showed larger changes in MD activation. The equivalent analysis of the DMN ROI activation (Figure 5h) did not correlate with algorithmic complexity (Pearson, **Study 1** - $r_{(15)}=0.295$, $p=0.268$, $CI=-0.24$ to 0.69 ; **Study 2** - $r_{(14)}=0.290$, $p=0.294$, $CI=-0.26$ to 0.70). However there was no significant correlation when examining the change in algorithmic complexity scores.

SOS differences between studies

The main effect of study contrast rendered regions where the difference between switching decisions and discrimination trial activation varied for study 1 and 2 (figure 6g). Study 1 showed greater activation in regions associated with the default mode network. Study 2 showed greater activation in regions associated with the multiple-demand (MD) network.

Activation modelling of the tasks and colour rule subtasks

Analyses of brain activation related to performance of the different tasks and colour rule sub-tasks were derived using the 1st level trial response contrasts maps which were modelled at the group-level with a two-way repeated-measures (RM-ANOVA). The within-subject factors were task (MT, OOO) and rule (colour, shape). In both studies, the average effect of performance (collapsed across the different tasks and rules) contrast rendered activation in the occipital cortex, the multiple-demand (MD) cortex including the dorsolateral frontoparietal cortex, the supplementary motor area (SMA) anterior cingulate cortex (ACC) and anterior insula (figure 6a/d). The main effect of task contrast showed an activation difference in the occipital cortex with greater activation during odd one out trials (figure 6b/e). The main effect of rule contrast showed an activation difference in occipital cortex with greater activation during shape trials (figure 6c/f).

4. Discussion

Our novel self-ordered switching paradigm enabled us to derive new insights into the processes by which people learn complex multi-component tasks. In this executively challenging context, the learning process was characterised by ongoing optimisation of behavioural and concurrent redeployment of functional-brain network activity.

More specifically, higher scores were achieved by participants who showed a greater degree of hierarchically structured behaviour and fewer behaviourally costly switches. Concurrently, in line with what has been observed for the learning of simple tasks we found that MD network activity decreased and DMN network increased as the participant learnt to perform the discrimination trials relative to switch decisions.

The most striking finding of this study was that this characteristic pattern of changes in brain network activation was not observed for decisions to switch; instead, the reverse pattern was evident, with MD regions increasing and DMN regions decreasing in activity with practice. Therefore, functional network resources supporting executive demand were redeployed toward the executive switching points around which structured behavioural routines were organised.

In everyday life we often are confronted with situations where multiple concurrent tasks must be performed. Furthermore, real-world tasks are often intrinsically hierarchical in their demands, requiring sub-goals to be completed to achieve overarching goals (Poljac et al., 2018). Most experimental paradigms that have been used to study cognitive control processes enforce careful control of the flow of events, for example, specifying when the participant must switch between sub-tasks (Kiesel et al., 2010; Koch et al., 2018). These types of design facilitate the careful balancing of experimental conditions, but they lack ecological validity insofar as they do not provide the participant with the opportunity to apply their executive abilities to organise their own behaviour (Goldberg & Podell, 2000). This is important, because structured routines are a defining characteristic of human executive behaviour, with disorganisation being symptomatic across prominent neurological and psychiatric populations (Averbeck et al., 2011; Collins & Koechlin, 2012; Fallon et al., 2013).

Moreover, the majority of cognitive control studies do not take into account the shifting role that brain systems play as a task is learnt. We argue that this is a critical oversight because the purpose of cognitive control systems is not just to support complex behaviour, but more definitively, is to enable the learning of such behaviours in order that they can be performed with minimal effort.

We and others have previously reported that the changes in brain activity associated with learning during commonly applied cognitive control paradigms can dwarf the control processes that those paradigms were designed to examine (Badre et al., 2010; Erika-Florence et al., 2014). More specifically, in a previous sequence of articles, we and others reported that during the learning of simple stimulus-response mappings from either instruction or feedback, as learning progressed, there was a concurrent reduction in event-related activity within frontoparietal networks associated with effortful cognition and an increase in default mode network activity (Hampshire et al., 2016, 2019; Ruge et al., 2018, 2019; Ruge & Wolfensteller, 2010). This likely reflected a steady transition from more deliberate to more automated modes of behaviour. This is also in line with previous work that has shown that more automatic retrieval of strongly encoded memory traces are associated with greater DMN activity in both semantic and episodic memory retrieval contexts (Vatansever et al. 2021).

The results of the current study extended previous work by examining the shifting role of brain networks during the learning of tasks that are ecologically valid insofar as they require multiple sub-tasks to be efficiently self-organised. When this learning process

was supported by detailed ongoing feedback in study 1, the behavioural routines that participants tended to develop had notable characteristics, these being the minimisation of number of switches performed, the hierarchical sequencing of task and rule switches measured by the tendency towards algorithmically simpler sequences. Interestingly, these characteristics correlated with overall task performance, both in terms of the number of points awarded across learning stages and individual differences in behavioural routines. From a simple computational perspective this relationship would not necessarily have been the case.

For example, task switches could have been performed at the same frequency, or subordinate to rule switches and would still have produced the same number of points per trial completed. Nonetheless, participants in study 1 tended to organise the rule switches subordinate to the task switches, a structure that was implicitly encouraged by the structure of the feedback elements (+10 point bonus for when both task meters reached their maximum). Notably, task switches were more effortful, as evidenced by greater response time costs; therefore, the bias towards more rule than task switches, and fewer switches overall, was adaptive insofar as it reduced overall executive costs. Relatedly, the application of algorithmically simpler routines was demonstrably less effortful, being accompanied by smaller costs and the achievement of higher overall scores. From this perspective, the application of hierarchically structured behaviour can be seen to reflect a tendency to seek modes of behaviour that require minimal switching effort to produce maximum reward.

Indeed, evidence of a tendency to self-order behaviour in a structured manner even when such structure is not a task requirement has been seen previously in some previous studies. A prominent example is voluntary task-switching (VTS) paradigms, which instruct participants to randomly choose between two tasks on each trial. Interestingly, the original VTS reports imply that participants did not fully adhere to this instruction and exhibited a significant repetition bias (Arrington & Logan, 2004, 2005). This is evidence of the behavioural self-ordering that reduced the frequency of task-switching that may have resulted in an advantageous reduction in cognitive demand. Similarly, here, participants who were better practised at SOS (study 2) exhibited a significantly greater trial repetition bias, and also behaved with a more structured routine. Relatedly, we have also observed that when identifying targets during a novel self-ordered intradimensional-extradimensional switching paradigm, participants tend to hierarchically order their search behaviour between stimuli of different dimensions. Importantly, in this context where there is no intrinsic computational advantage to sequencing the search by stimulus dimension but switching between dimensions nonetheless incurs a greater response time costs (Hampshire & Owen, 2006).

Perhaps the most striking result of the current study pertains to how the underlying brain network resources associated with performance of discrimination trials and switch decisions changed as the task was learnt. More specifically, the application of self-ordered switching was concomitant with a functional activation that was more

complex in its dynamics than that which has previously been observed during simple stimulus- response learning (Hampshire et al., 2016, 2019; Ruge et al., 2019; Ruge & Wolfensteller, 2010). In accordance with those past findings, frontoparietal activation in response to discrimination trials decreased with practise whilst the degree of default mode network deactivation was also seen to reduce. Unexpectedly though, we also observed the opposite pattern of change for switching decisions, with frontoparietal activation increasing and DMN activation decreasing as a function of learning to self-order behaviour.

The ACC and DLPFC regions that were mapped based on switching decision activity here correspond closely to parts of the MD system, which generally is more active when performing tasks that require more cognitive effort (Crittenden & Duncan, 2014). Therefore, this pattern of results likely reflects a redeployment of flexible and limited capacity network resources towards the executive switching events that enable structured behaviour as the simple stimulus-response mappings underpinning performance of trials become automatic through practice.

A plausible possibility is that frontoparietal systems operate at the leading edge of the automatization process, enabling progressively more complex but efficient hierarchical behaviours to be applied. A testable prediction of this is that the switching activation would also decrease with sufficient practise. This would make an interesting prediction that could be tested across multiple SOS scanning sessions, perhaps with paradigms that have scope for deeper hierarchies that limited capacity resources can progressively develop.

The key difference between study 1 and study 2 was that the latter had more substantial training prior to entering the scanner and only minimal feedback within the scanner. These modifications were intended to provide a control condition where the opportunity for further optimisation of the task performance was reduced, with participants in study 2 having progressed further along the learning curve and exploiting a more established routine. The behavioural results supported this view, as in study 2, SOS was performed faster and with routines that had lower algorithmic complexity compared to study 1. However, repeat discrimination trials and switching costs still showed significant reduction with practice across study 2, therefore some learning due to practice was ongoing.

More broadly, we believe that paradigms such as the SOS that enable participants the freedom to self-organise their behaviour have untapped potential for understanding normal and pathological cognitive control functions. The natural variation in the behavioural efficiency of healthy adults suggests that there is a potential to examine self-ordered behaviours in patient populations that suffer from cognitive impairments. It could be predicted that some patient groups may take longer to generate a stable behavioural routine and that this routine would be more complex, inefficient and refined at a slower rate, compared to healthy controls. These features may dissociate

across patient populations for example, Parkinson's patients show difficulty in updating strategies during set-shifting and can show a reduced tendency towards structured dimensional switching routines (Fallon et al., 2013), whereas patients with schizophrenia may prematurely implement a strategy before sufficiently exploring the options (Averbeck et al., 2011). Future work should also focus on the learning of structured routines over longer time frames and within paradigms that can allow deeper hierarchical self-organisation of behaviour and should examine how the responsiveness of functional networks evolves as those hierarchical behaviours are established.

Limitations

Response times to switch and repeat discrimination trials may not be directly comparable due to our task design. To perform a discrimination trial, a participant will process the visual features, make the discrimination decision and perform a motor response. In repeat discriminations, this process starts when the stimulus first appears on the screen, which is the point that RT are measured from. For responses to switch discrimination trials, RT are measured from when a switch decision is made, however the stimuli appeared on screen prior to this and so the processing of the stimuli could be underway. It is important to note that all event related activation estimates in this study are from a rapid event related design without rest blocks; therefore, as is common for rapid event related tasks, they relate to event related changes in activation relative to the implicit baseline of ongoing activity at those times. Further we may be underpowered for the exploratory individual differences imaging analysis, a future study with a larger cohort could help examine these questions with more power.

In summary, our results show that when learning to perform complex multi component tasks participants have a tendency to hierarchically structure their behavioural routines to minimise task switching costs. As participants learnt to perform simple discrimination task demands, changes in brain dynamics were characterised by reductions in MD network activity and increases in DMN activity. Importantly this was not the case for deciding to switch, where the opposite pattern, of increased MD network activation and reduced DMN occurs with practise. These changes to brain network dynamics may reflect the redeployment of flexible and limited capacity resources towards decisions to switch between multiple competing tasks and away from the routine execution of behaviour that has been learnt.

Figures

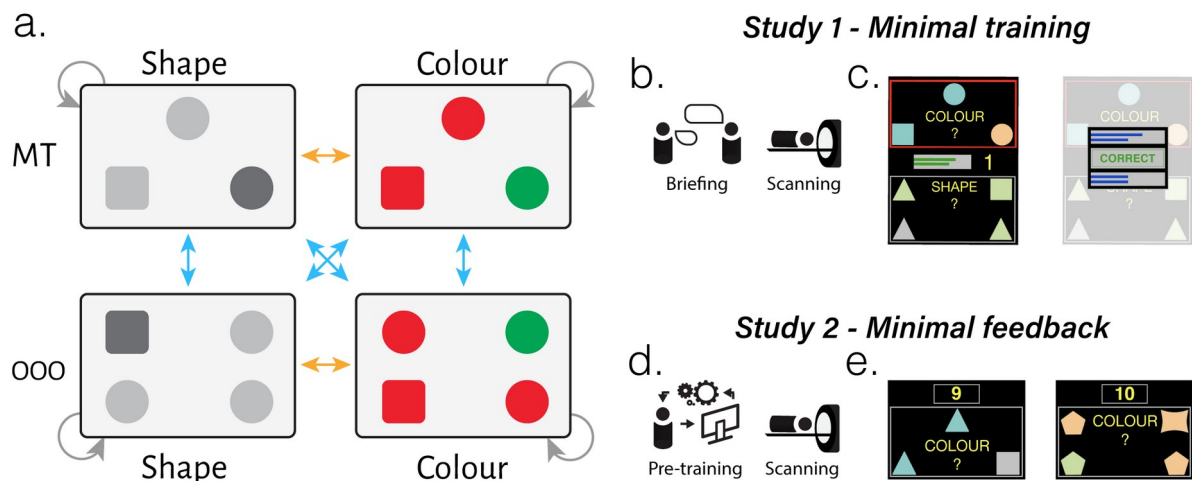


Figure 1. Self-ordered switching (SOS) experiments. a) Schematic of the two tasks (MT - matching task, OOO - odd one out) and the two rules (colour, shape) of the SOS paradigm. Grey arrows represent repeat-trials where the same condition was repeated. Task-switching and rule-switching decisions are represented by the blue and orange arrows, respectively. b) In study 1, SOS was performed after minimal training with a short demo of button presses. c) Inside the scanner both tasks were simultaneously displayed on the screen and hierarchical response feedback was presented with progress meters relating to each task and each of the task's colour and shape subordinate rules. d) In study 2, the task was performed after extensive training with full hierarchical feedback and then e) performed in the scanner with minimal feedback involving the overall score only.

Figure 2: SOS hierarchical response feedback and trial types in study 1.

- a) **The feedback panel appearing at the end of each trial** displayed response feedback (CORRECT in green ink or INCORRECT in red ink) along with progress meters of the number of points scored for each subtask (blue bars). A correct response increased the bar, and an incorrect response decreased it, by one point. If either the colour or shape meter got two points ahead of the other, the other would decrease by a point. When both the colour and shape meter reached 10, they reset, and the corresponding task meter incremented by one.
- b) **The overall task progress meter** displaying the number of points currently accumulated per task were displayed continuously on screen, in between the two task panels. These increased when both the colour and rule subtask reached 10 points.
- c) **Example sequence for the matching task colour trial.** The left button press indicated the blue square on the left-hand side of the screen matched the colour of the probe stimuli (blue circle) at the top of the screen.
- d) **Example sequence of a rule-switch trial.** The downward button press indicates a rule-switch from the MT-colour trial to the MT-shape trial. The

subsequent right button press indicated the grey rectangle on the right-hand side of the screen matched the shape of the probe stimuli (the green rectangle) at the top of the screen.

- e) **Example sequence of task-switch trial.** The upward button press indicated a task switch from the MT to the OOO task, the matching task border turns grey, and the OOO task turns red. The subsequent left-hand button press successful indicates the only stimuli in the array of a grey colour is on the left of the screen.
- f) **Example sequence of OOO shape trial.** The right button press indicates the only rectangle in the visual array is on the right of the screen. This correct response increased the shape meter by one and now both the colour and rule meters reach 10 points (indicated in blue), this causes the OOO task meter to increase by one and now both task meters reach the maximum (indicated in green), the overall score increments by 1 and a + 10 bonus is added.

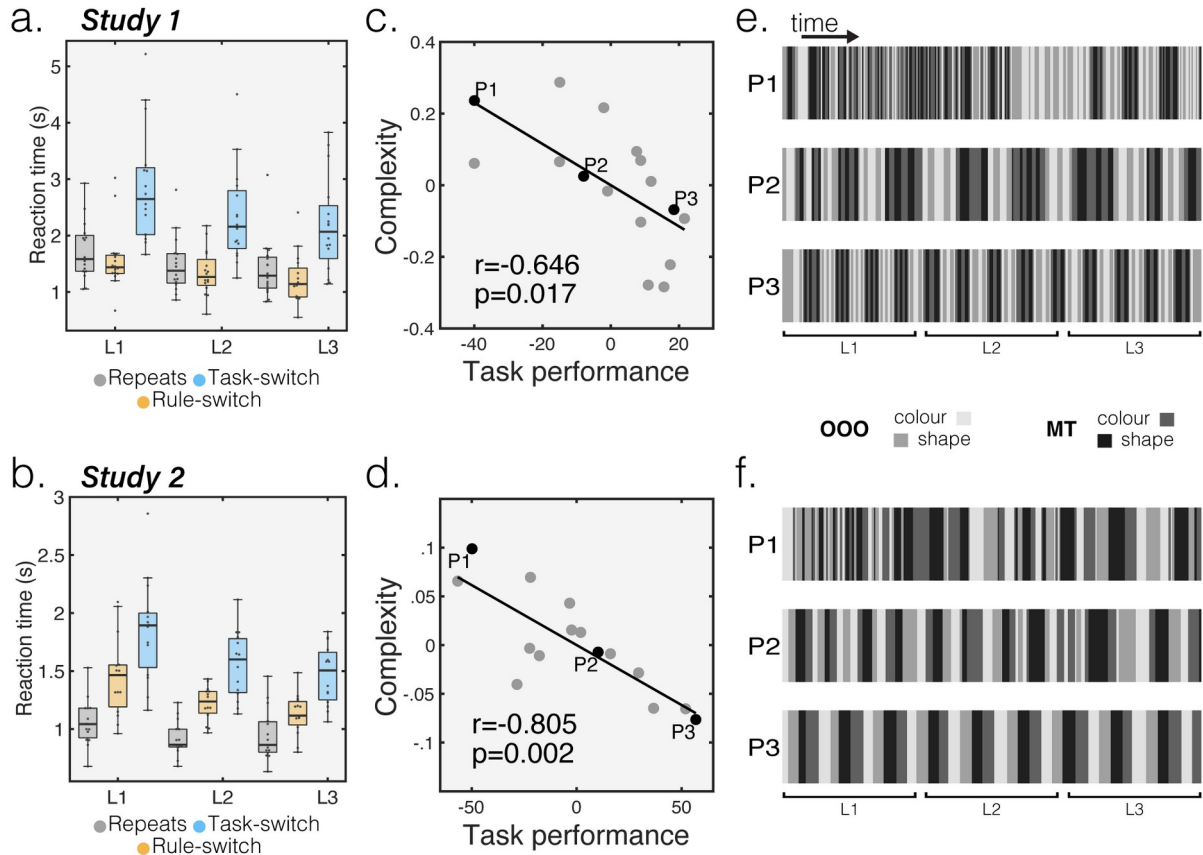


Figure 3. Behavioural results: SOS switching costs and behavioural structuring.

a/b) Response times (RT) for each discrimination condition were summarised within each ~7-minute learning stage (L). Each point represents a participant's median RT for a condition for study 1 (a) and study 2 (b). Boxplots represent the group median (central line), and the 25-75th percentiles (boxes). RT Switching costs significantly decreased with practise in study 1 and 2 (a/b).

c/d) Normalised algorithmic complexity (y-axis) of participants' SOS behaviour plotted against overall task performance, in terms of points scored (x- axis) for study 1 (c) and study 2 (d). The residuals are plotted from a partial correlation controlling for the number of trials completed.

e/f) For each study, three participants are highlighted and their behavioural timecourses are displayed. Of those, participant 1 (P1) generated the most complex SOS behaviour and P3 the least. Reductions in complexity was characterised by chunks of repeat-trials that were consistently punctuated by switches.

Figure 4. Mean task-switch and rule-switch cost (y-axis, seconds) correlates with the total number of switches made. This was the case for both study 1 (a) and study 2 (b).

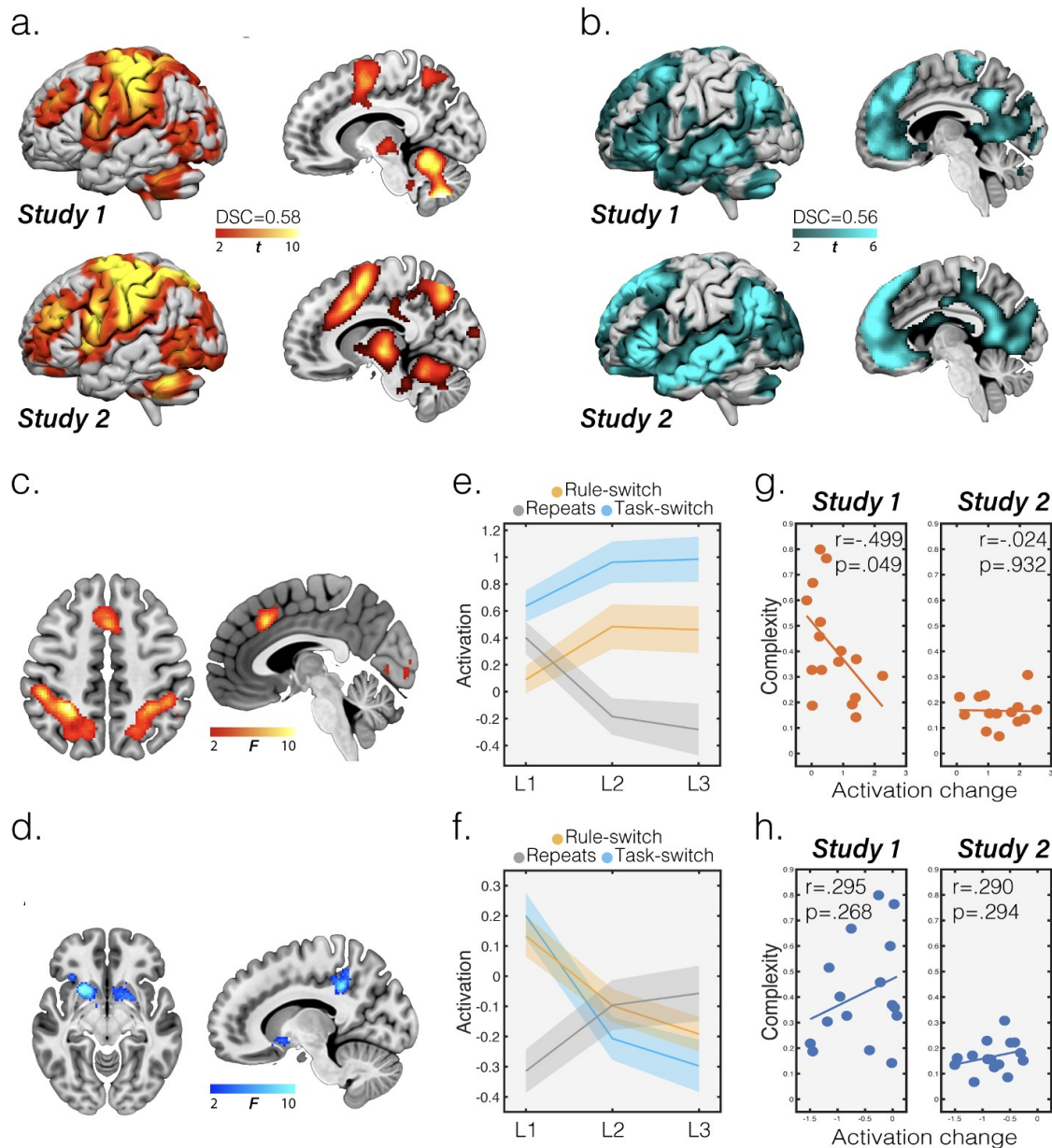


Figure 5. Multiple-demand and default mode activation varies with SOS learning.

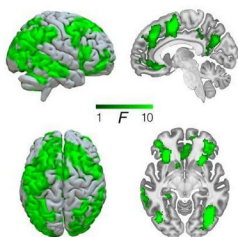
Group-level activation increases (a) during switching decision > discrimination trial contrast and (b) during discrimination trial > switching decisions contrast.

Dice similarity coefficients (DSC) represent the spatial similarity of significant voxels between the studies. Clusters of activation that change over learning stages for (c) switching decisions > discrimination trial contrast and (d) discrimination trial > switching decisions contrast. These clusters formed the composite MD ROI (c) and DMN ROI (d). **e/f** Mean activation for each learning stage (L) and each trial type for the MD (e) and DMN (f) composite ROIs (shaded area represents the standard error) in study 1. The discrimination trials > switching decisions showed decreased activation across the learning stages in the MD ROI, but increased activation across the learning stages for the DMN ROI. The opposite pattern was found for the switch decisions > discrimination trials, with increased activation across the learning stages in the MD ROI, and decreased activation for the DMN ROI. **g/h** Correlations between the sum of activation change over learning stage and normalised algorithmic complexity of SOS behaviour

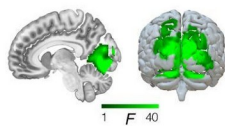
for the MD (g) and DMN (h) composite ROIs. All activation maps are thresholded at an uncorrected $p < 0.01$, followed by a $p < 0.05$ FDR cluster correction

Study 1

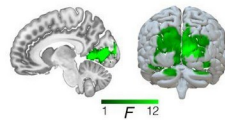
a - Average effect of trials



b - Main effect of task

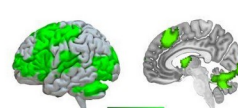


c - Main effect of rule

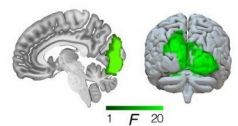


Study 2

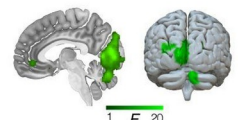
d - Average effect of trials



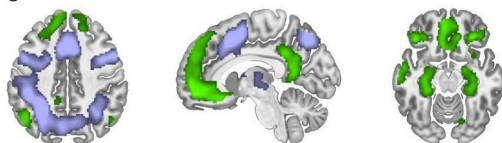
e - Main effect of task



f - Main effect of rule



g.



Study 1 > Study 2
t: 1 5



Study 1 < Study 2
t: 1 5

Figure 6. Activation modelling of tasks and colour rule subtasks:

A two-way repeated-measures voxelwise analysis of variance (RM-ANOVA) modelled activation change with task (matching task, odd one out) and rule (colour, shape) as the within-subject factors. This model was used to calculate the F contrast for **(a)** activation during all trials, **(b)** the main effects of task, and **(c)** the main effect of rule. **(d-f)** The RM_ANOVA analysis was repeated for study 2. **(g)** The main effect of study 1 vs study 2 for the contrast of the switch > trial. To determine the direction of the main effect, two t-contrasts were generated to display study 1 > 2 (green) and study 2 > 1 (lilac) differences. All maps were thresholded at $p < 0.01$ followed by a false discovery rate (FDR) cluster-level correction for multiple comparisons.

Data Sharing and Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of interest

The authors have no conflict of interest to declare.

Funding

This work was funded by an ITMAT grant from the Imperial NIHR Biomedical Research Centre. Adam Hampshire's team is supported by the Dementia Research Institute. Richard Daws was supported by a Marie Curie CIG awarded to Adam Hampshire.

This work was supported by funding to RED from the Engineering and Physical Sciences Research Council (EPSRC) under supervision of AH via the centre for doctoral training (CDT) Neurotechnology programme at Imperial College London. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

References

Antzoulatos, E. G., & Miller, E. K. (2014). Increases in Functional Connectivity between Prefrontal Cortex and Striatum during Category Learning. *Neuron*, 83(1), 216–225. <https://doi.org/10.1016/j.neuron.2014.05.005>

Arrington, C. M., & Logan, G. D. (2004). The Cost of a Voluntary Task Switch. *Psychological Science*, 15(9), 610–615. <https://doi.org/10.1111/j.0956-7976.2004.00728.x>

Arrington, C. M., & Logan, G. D. (2005). Voluntary Task Switching: Chasing the Elusive Homunculus. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(4), 683–702. <https://doi.org/10.1037/0278-7393.31.4.683>

Averbeck, B. B., Evans, S., Chouhan, V., Bristow, E., & Shergill, S. S. (2011). Probabilistic learning and inference in schizophrenia. *Schizophrenia Research*, 127(1–3), 115–122. <https://doi.org/10.1016/j.schres.2010.08.009>

Badre, D., Kayser, A. S., & D'Esposito, M. (2010). Frontal Cortex and the Discovery of Abstract Action Rules. *Neuron*, 66(2), 315–326. <https://doi.org/10.1016/j.neuron.2010.03.025>

Bassett, D. S., Wymbs, N. F., Porter, M. A., Mucha, P. J., Carlson, J. M., & Grafton, S. T. (2011). Dynamic reconfiguration of human brain networks during learning. *Proceedings of the National Academy of Sciences*, 108(18), 7641–7646. <https://doi.org/10.1073/pnas.1018985108>

Buckner, R. L., Krienen, F. M., Castellanos, A., Diaz, J. C., & Yeo, B. T. T. (2011). The organization of the human cerebellum estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106(5), 2322–2345. <https://doi.org/10.1152/jn.00339.2011>

Chaitin, G. J. (1966). On the Length of Programs for Computing Finite Binary Sequences. *Journal of the ACM*, 13(4), 547–569. <https://doi.org/10.1145/321356.321363>

Collins, A., & Koechlin, E. (2012). Reasoning, Learning, and Creativity: Frontal Lobe Function and Human Decision-Making. *PLoS Biology*, 10(3), e1001293. <https://doi.org/10.1371/journal.pbio.1001293>

Crittenden, B. M., & Duncan, J. (2014). Task Difficulty Manipulation Reveals Multiple Demand Activity but no Frontal Lobe Hierarchy. *Cerebral Cortex*, 24(2), 532–540. <https://doi.org/10.1093/cercor/bhs333>

Demanet, J., Verbruggen, F., Liefoghe, B., & Vandierendonck, A. (2010). Voluntary task switching under load: Contribution of top-down and bottom-up factors in goal-directed behavior. *Psychonomic Bulletin & Review*, 17(3), 387–393. <https://doi.org/10.3758/PBR.17.3.387>

- Duncan, J. (2013). The Structure of Cognition: Attentional Episodes in Mind and Brain. *Neuron*, 80(1), 35–50. <https://doi.org/10.1016/j.neuron.2013.09.015>
- Erika-Florence, M., Leech, R., & Hampshire, A. (2014). A functional network perspective on response inhibition and attentional control. *Nature Communications*, 5(1), 4073. <https://doi.org/10.1038/ncomms5073>
- Fallon, S. J., Williams-Gray, C. H., Barker, R. A., Owen, A. M., & Hampshire, A. (2013). Prefrontal Dopamine Levels Determine the Balance between Cognitive Stability and Flexibility. *Cerebral Cortex*, 23(2), 361–369. <https://doi.org/10.1093/cercor/bhs025>
- Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A. R., Fox, P. T., Eickhoff, S. B., Yu, C., & Jiang, T. (2016). The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture. *Cerebral Cortex*, 26(8), 3508–3526. <https://doi.org/10.1093/cercor/bhw157>
- Finc, K., Bonna, K., He, X., Lydon-Staley, D. M., Kühn, S., Duch, W., & Bassett, D. S. (2020). Dynamic reconfiguration of functional brain networks during working memory training. *Nature Communications*, 11(1), 2435. <https://doi.org/10.1038/s41467-020-15631-z>
- Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and Modulatory Interactions in Neuroimaging. *NeuroImage*, 6(3), 218–229. <https://doi.org/10.1006/nimg.1997.0291>
- Goldberg, E., & Podell, K. (2000). Adaptive Decision Making, Ecological Validity, and the Frontal Lobes. *Journal of Clinical and Experimental Neuropsychology*, 22(1), 56–68. [https://doi.org/10.1076/1380-3395\(200002\)22:1;1-8;FT056](https://doi.org/10.1076/1380-3395(200002)22:1;1-8;FT056)
- Hampshire, A., Daws, R. E., Neves, I. D., Soreq, E., Sandrone, S., & Violante, I. R. (2019). Probing cortical and sub-cortical contributions to instruction-based learning: Regional specialisation and global network dynamics. *NeuroImage*, 192, 88–100. <https://doi.org/10.1016/j.neuroimage.2019.03.002>
- Hampshire, A., Hellyer, P. J., Parkin, B., Hiebert, N., MacDonald, P., Owen, A. M., Leech, R., & Rowe, J. (2016). Network mechanisms of intentional learning. *NeuroImage*, 127, 123–134. <https://doi.org/10.1016/j.neuroimage.2015.11.060>
- Hampshire, A., MacDonald, A., & Owen, A. M. (2013). Hypoconnectivity and Hyperfrontality in Retired American Football Players. *Scientific Reports*, 3(1), 2972. <https://doi.org/10.1038/srep02972>
- Hampshire, A., & Owen, A. M. (2006). Fractionating Attentional Control Using Event-Related fMRI. *Cerebral Cortex*, 16(12), 1679–1689. <https://doi.org/10.1093/cercor/bhj116>
- Ionescu, T. (2012). Exploring the nature of cognitive flexibility. *New Ideas in Psychology*, 30(2), 190–200. <https://doi.org/10.1016/j.newideapsych.2011.11.001>

Khambhati, A. N., Medaglia, J. D., Karuza, E. A., Thompson-Schill, S. L., & Bassett, D. S. Kiesel, (2018). Subgraphs of functional brain networks identify dynamical constraints of cognitive control. *PLOS Computational Biology*, 14(7), e1006234. <https://doi.org/10.1371/journal.pcbi.1006234>

A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching—A review. *Psychological Bulletin*, 136(5), 849–874. <https://doi.org/10.1037/a0019842>

Koch, I., Poljac, E., Müller, H., & Kiesel, A. (2018). Cognitive structure, flexibility, and plasticity in human multitasking—An integrative review of dual-task and task-switching research. *Psychological Bulletin*, 144(6), 557–583. <https://doi.org/10.1037/bul0000144>

Kolmogorov, A. N. (1968). Three approaches to the quantitative definition of information^{*}. *International Journal of Computer Mathematics*, 2(1–4), 157–168. <https://doi.org/10.1080/00207166808803030>

Lempel, A., & Ziv, J. (1976). On the Complexity of Finite Sequences. *IEEE Transactions on Information Theory*, 22(1), 75–81. <https://doi.org/10.1109/TIT.1976.1055501>

McLaren, D. G., Ries, M. L., Xu, G., & Johnson, S. C. (2012). A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. *NeuroImage*, 61(4), 1277–1286. <https://doi.org/10.1016/j.neuroimage.2012.03.068>

Petersen, S. E., van Mier, H., Fiez, J. A., & Raichle, M. E. (1998). The effects of practice on the functional anatomy of task performance. *Proceedings of the National Academy of Sciences*, 95(3), 853–860. <https://doi.org/10.1073/pnas.95.3.853>

Poljac, E., Kiesel, A., Koch, I., & Müller, H. (2018). New perspectives on human multitasking. *Psychological Research*, 82(1), 1–3. <https://doi.org/10.1007/s00426-018-0970-2>

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, 59(3), 2142–2154. <https://doi.org/10.1016/j.neuroimage.2011.10.018>

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124(2), 207–231. <https://doi.org/10.1037/0096-3445.124.2.207>

Rosenberg, M. D., Finn, E. S., Scheinost, D., Papademetris, X., Shen, X., Constable, R. T., & Chun, M. M. (2016). A neuromarker of sustained attention from whole-brain

functional connectivity. *Nature Neuroscience*, 19(1), 165–171. <https://doi.org/10.1038/nn.4179>

Ruge, H., Karcz, T., Mark, T., Martin, V., Zwosta, K., & Wolfensteller, U. (2018). On the efficiency of instruction-based rule encoding. *Acta Psychologica*, 184, 4–19. <https://doi.org/10.1016/j.actpsy.2017.04.005>

Ruge, H., Schäfer, T. A., Zwosta, K., Mohr, H., & Wolfensteller, U. (2019). Neural representation of newly instructed rule identities during early implementation trials. *eLife*, 8, e48293. <https://doi.org/10.7554/eLife.48293>

Ruge, H., & Wolfensteller, U. (2010). Rapid Formation of Pragmatic Rule Representations in the Human Brain during Instruction-Based Learning. *Cerebral Cortex*, 20(7), 1656–1667. <https://doi.org/10.1093/cercor/bhp228>

Sakai, K. (2008). Task Set and Prefrontal Cortex. *Annual Review of Neuroscience*, 31(1), 219–245. <https://doi.org/10.1146/annurev.neuro.31.060407.125642>

Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X.-N., Holmes, A. J., Eickhoff, S. B., & Yeo, B. T. T. (2018). Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cerebral Cortex*, 28(9), 3095–3114. <https://doi.org/10.1093/cercor/bhx179>.

D. J., Bonnelle, V., De Boissezon, X., Beckmann, C. F., James, S. G., Patel, M. C., & Mehta, M. A. (2010). Distinct frontal systems for response inhibition, attentional capture, and error processing. *Proceedings of the National Academy of Sciences*, 107(13), 6106–6111. <https://doi.org/10.1073/pnas.1000175107>

Sheffield, J. M., Mohr, H., Ruge, H., & Barch, D. M. (2021). Disrupted Salience and Cingulo- Opercular Network Connectivity During Impaired Rapid Instructed Task Learning in Schizophrenia. *Clinical Psychological Science*, 9(2), 210–221. <https://doi.org/10.1177/2167702620959341>

E., Leech, R., & Hampshire, A. (2019). Dynamic network coding of working-memory domains and working-memory processes. *Nature Communications*, 10(1), 936. <https://doi.org/10.1038/s41467-019-08840-8>

Vatansever, D., Smallwood, J., & Jefferies, E. (2021). Varying demands for cognitive control reveals shared neural processes supporting semantic and episodic memory retrieval. *Nature communications*, 12(1), 2134.

E., Violante, I. R., Daws, R. E., & Hampshire, A. (2021). Neuroimaging evidence for a network sampling theory of individual differences in human intelligence test performance. *Nature Communications*, 12(1), 2072. <https://doi.org/10.1038/s41467-021-22199-9>

Stoet, G., & Snyder, L. H. (2007). Extensive practice does not eliminate human switch costs. *Cognitive, Affective, & Behavioral Neuroscience*, 7(3), 192–197. <https://doi.org/10.3758/CABN.7.3.192>

Stokes, M. G. (2015). 'Activity-silent' working memory in prefrontal cortex: A dynamic coding framework. *Trends in Cognitive Sciences*, 19(7), 394–405. <https://doi.org/10.1016/j.tics.2015.05.004>

Swick, D., & Turken, A. U. (2002). Dissociation between conflict detection and error monitoring in the human anterior cingulate cortex. *Proceedings of the National Academy of Sciences*, 99(25), 16354–16359. <https://doi.org/10.1073/pnas.252521499>

Yeung, N., & Monsell, S. (2003). The effects of recent practice on task switching. *Journal of Experimental Psychology: Human Perception and Performance*, 29(5), 919–936. <https://doi.org/10.1037/0096-1523.29.5.919>

Zenil, H., Hernández-Orozco, S., Kiani, N., Soler-Toscano, F., Rueda-Toicen, A., & Tegnér, J. (2018). A Decomposition Method for Global Evaluation of Shannon Entropy and Local Estimations of Algorithmic Complexity. *Entropy*, 20(8), 605. <https://doi.org/10.3390/e20080605>

Supplementary Materials

Supplementary MRI analysis

When comparing task vs rule switch decisions increased activation was observed in the dorsal visual streams in study 1 (supplementary figure 1a). The same contrast in study 2 showed focal increases in the dorsal motor cortex (supplementary figure 1b), modest spatial similarity between study 1 and 2 was observed (DSC=0.210). The inverse contrast testing for increased activation for rule vs task switch decisions showed no differences that survived correction in either study.

Supplementary Figure 1 : T-contrast maps testing for increased activation during task-switch decisions compared to rule-switch decisions in study 1 (supplementary figure 1a) and study 2 (supplementary figure 2b) (the inverse contrast showed no significant differences in either study).

