Reward boosts neural coding of task rules to optimise cognitive flexibility

Abbreviated Title: Reward prospect and cognitive flexibility

Hall-McMaster, S., Muhle-Karbe, P.S, Myers, N.E.*, & Stokes, M.G.*

Department of Experimental Psychology, University of Oxford,

New Radcliffe House, OX2 6AE

Corresponding author: sam.hall-mcmaster@psy.ox.ac.uk

Number of pages: 27

Number of figures: 8

Conflict of interest: The authors declare no competing financial interests.

*Authors made equal contribution to this work

Abstract

Cognitive flexibility is critical for intelligent behaviour. However, its execution is effortful and often suboptimal. Recent work indicates that flexible behaviour can be improved by the prospect of reward, suggesting that rewards optimise flexible control processes. Here we investigated how different reward prospects influence neural encoding of task rule information to optimise cognitive flexibility. We applied representational similarity analysis (RSA) to human electroencephalograms, recorded while participants performed a rule-guided decision-making task. During the task, the prospect of reward varied from trial to trial. Participants made faster, more accurate judgements on high reward trials. Critically, high reward boosted neural coding of the active task rule and the extent of this increase was associated with improvements in task performance. These results suggest that reward motivation can improve cognitive performance by strengthening neural coding of task rule information, improving cognitive flexibility during complex behaviour.

Significance Statement

The importance of motivation is evident in the ubiquity with which reward prospect guides adaptive behaviour and the striking number of neurological conditions where motivation is impaired. In this study, we investigated how dynamic changes in motivation, as manipulated through reward, shape neural coding for task rules during a flexible decision-making task. The results of this work suggest that motivation to obtain reward modulates encoding of task rules needed for flexible behaviour. The extent to which reward increased task rule coding also tracked improvements in behavioural performance under high reward conditions. These findings help inform how motivation shapes neural processing in the healthy human brain.

Introduction

Flexible cognitive control is critical to human intelligence. When vying to win a card game, we can use arbitrary rules to play the best hand. When navigating a new city, we can apply navigation rules to sensory input from the world around us to arrive at the next tourist attraction. Cognitive flexibility is suboptimal compared to behaviours that require less flexible processing (Kleinsorge & Rinkenauer, 2012; Shen & Chun, 2011) but can be improved by motivational factors, such as the prospect of reward for fast and accurate performance (Braem & Egner, 2018). When the prospect of reward is high, performance improves on flexible rule-based tasks (Etzel et al., 2016; Kleinsorge & Rinkenauer, 2012), suggesting that reward might optimise cognitive flexibility by tuning rule-based neural coding patterns.

The existing neuroimaging literature provides support for this perspective, indicating that reward prospect leads to stronger recruitment of frontoparietal brain regions implicated in cognitive control (Parro, Dixon, & Christoff, 2018). When reward cues are presented at the beginning of each trial, activity in these regions is typically enhanced prior to the onset of a target stimulus, suggesting that proactive control mechanisms contribute to reward-induced performance benefits (Engelmann et al., 2009; Krebs et al., 2011; Padmala & Pessoa, 2011). Critically, a recent study combining fMRI with pattern classification methods found that frontoparietal activity prior to target onset encoded abstract task rules with greater fidelity on reward trials than on no-reward trials (Etzel et al., 2016). The extent to which reward prospect enhanced the decodability of task rules also mediated behavioural benefits, suggesting that performance improvements may result from reward-driven tuning in cognitive control processes that prioritise task-relevant processing.

This proposal has considerable theoretical appeal. Increased prioritisation of task-relevant information would provide a general explanation for the beneficial effects of reward on flexible cognitive control that is grounded in theories of prefrontal cortex function (Duncan, 2001; Miller & Cohen, 2001) and extends across a wide array of cognitive demands. Nonetheless, constraints on the temporal resolution of fMRI make it challenging to isolate task rule coding from subsequent perceptual processing because the slow hemodynamic response can make it difficult to pinpoint effects in time and distinguish sustained anticipatory activity from transient stimulus-evoked responses. To address this, high temporal resolution methods such as magnetoencephalography and electroencephalography are needed because their ability to distinguish rapid stimulus-evoked dynamics makes them ideal for isolating the effects of reward on task rule coding from subsequent neural coding patterns.

The present study aimed to characterise how trial-by-trial changes in reward motivation influence neural coding for task rule information in the service of flexible behaviour. To do so, we

developed a behavioural task that manipulated reward motivation dynamically by varying the magnitude of reward for fast and accurate performance. We then applied representational similarity analysis (RSA; Kriegeskorte, Mur, & Bandettini 2008) to human electroencephalograms to examine changes in neural coding as a function of reward prospect. This high temporal resolution, multivariate approach allowed us to track both task rule and sensory variables while separating task preparation from target-evoked neural processing in the lead up to adaptive behavioural responses.

To summarise the main results, we found that high reward prospect produced significant performance improvements in accuracy and reaction time (RT). Consistent with the view that reward prospect increases proactive cognitive control, we found a significant increase in neural coding for task rules under high reward conditions. High reward prospect also modulated downstream sensorimotor coding, significantly increasing neural coding for task-relevant perceptual features of target stimuli and accurate motor responses. However, the only variables that tracked improvements in performance were neural coding of reward prospect and the increase in task coding on high reward trials.

Method

Participants

We set a target sample size of 30 participants. During recruitment three participants were excluded, one due to a corrupt EEG recording and two due to excessive artefacts that led to the rejection of more than 120/650 trials. We therefore collected three more participants to reach the 30 participant target. The final sample were between 18 and 35 years of age (mean age = 23, 11 female), with normal or corrected-to-normal vision, who reported no history of neurological or psychiatric illness. Participants received £8 per hour or course credit for taking part and could earn up to £10 extra for their performance. This study was approved by the Central University Research Ethics Committee at the University of Oxford and all participants signed informed consent before taking part.

Materials

Stimuli were presented on a 22-inch screen with a spatial resolution of 1280 x 1024 and refresh rate of 60Hz. Stimulus presentation was controlled using Psychophysics Toolbox-3 (Kleiner, Brainard & Pelli, 2007) in MATLAB (MathWorks, version R2015b). Reward cues and feedback shown during the task were presented in size 30 Arial font. Task cues and target stimuli had approximate visual angles of 2.52° (100x100 pixels) and 1.26° (50x50 pixels) respectively, with visual angles calculated based on an approximate viewing distance of 60cm. F and J keys on a standard QWERTY keyboard were used to record left and right hand responses. EEG data were recorded with 61 Ag/AgCl sintered electrodes (EasyCap, Herrsching, Germany), a NeuroScan SynAmps RT amplifier, and Curry 7 acquisition software (Compumedics NeuroScan, Charlotte, NC). EEG data were pre-processed in EEGLAB (Delorme & Makeig, 2004, version 14.1.1b), behavioural analyses were performed using JASP (https://jasp-stats.org, version 0.8.1.3) and EEG analyses were performed in MATLAB (MathWorks, version R2015b) using the EEGlab and Fieldtrip toolboxes as well as custom scripts.

Code Accessibility

Task and analysis code will be made available prior to publication.

Experimental Design and Statistical Analysis

In this task (Figure 1), participants' overarching goal was to gain as many points as possible. To do so, participants categorized bi-dimensional target stimuli based on their colour (yellow vs blue) or shape (square vs circle). On each trial, only one feature dimension of the target (colour or shape) was relevant to gaining points, while the other feature served as an irrelevant distractor. The relevant feature dimension was signalled through a visual task cue prior to target onset. In addition to a single relevant feature dimension, each trial offered a high or low reward magnitude for making

a correct response. This was signalled to participants at the beginning of each trial by a single pound sign (low reward: 5-10 points) or three pound signs (high reward: 50-100 points).

The experimental sequence consisted of a reward cue, task cue, target, feedback screen and an inter-trial interval (ITI). The reward cue (£ or £££) was first presented for 800ms, followed by a 400ms delay. The task cue (one of four possible abstract shapes) was then presented for 200ms. The mapping of cues to tasks was counterbalanced between participants. Task cue offset was followed by a 400ms delay. The target (a yellow square, blue square, yellow circle or blue circle) was then presented and remained on screen until a response was given or for a maximum of 1400ms. If the active task rule was colour, the correct response mapping was 'f' or 'j' for yellow or blue targets respectively. If the active rule was shape, the correct response mapping was 'f' or 'j' for square or circle targets respectively. The response phase was followed by feedback lasting 200ms. An incorrect response or omission resulted in feedback showing '+0' points. A correct response resulted in feedback showing '+X', where X was a value within the high or low reward point ranges, the precise value of which was determined by RT. More specifically, RT criteria for different points were initialised so that responses faster than 400ms, 600ms, 800ms, 1000ms, 1200ms and 1400ms earned 100/10, 90/9, 80/8, 70/7, 60/6 and 50/5 points on high and low reward trials respectively. For correct trials, the current trial RT was added to an array for its reward condition. When each array contained more than six values, individualised points criteria were calculated for that condition and were calculated again every time a new value entered the array. The individualised points criteria followed criteria outlined by Shen & Chun (2011), in which the most (to least) points are rewarded for correct responses faster than 95%, 80%, %65, %50 and 35% of the median RT. The trial concluded with a randomly selected ITI duration, drawn from a uniform distribution with values of 1000, 1100, 1200, 1300 or 1400ms. Participants were trained to reach a criterion of 70% accuracy before completing 10 experimental blocks of 65 trials. Excluding the first trial in each block, equal numbers of reward cues, task cues, stimuli and ITI durations were presented within each block. Presentation was pseudo-randomised to ensure trials were balanced based on task, target-congruency and task sequence for each reward condition. Target congruency refers to whether task-relevant and irrelevant target features are mapped to the same (congruent) or different (incongruent) response hands. Task sequence refers to whether the task rule was different from the previous trial (switch trial) or the same as the previous trial (repeat trial).

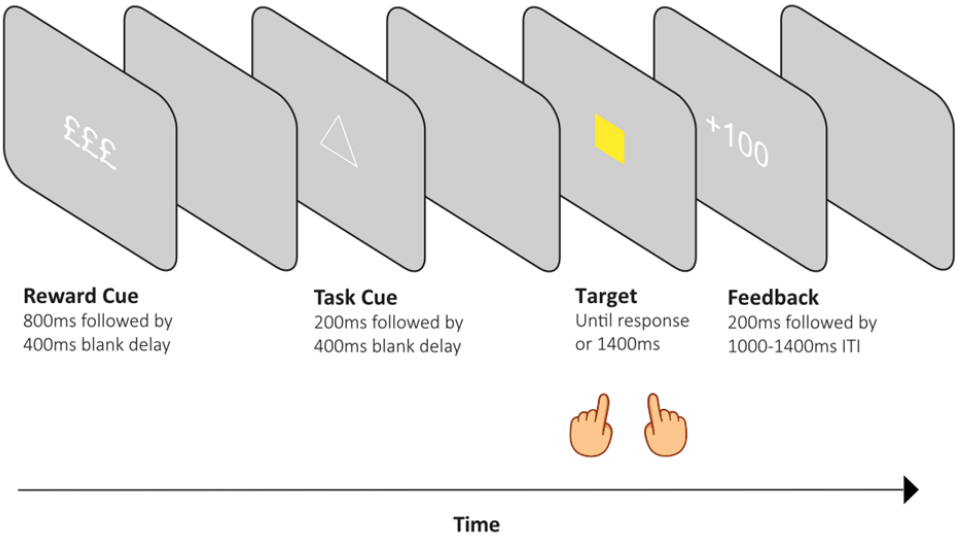

Figure 1. Task design. On each trial, a high or low reward cue was presented followed by a blank delay. A task rule cue was then presented, indicating whether participants should respond to the upcoming target based its colour or shape. Following a second blank delay, a bi-dimensional target (a coloured shape) appeared until a response was given or for a maximum duration of 1400ms. This was followed by feedback (based on accuracy and reaction time) and a variable inter-trial-interval.

Behavioural measures. Our main dependent measures were reaction time (RT) and proportion correct (accuracy). We calculated median RTs for the respective design cells to account for the skewness of RT distributions (see Ratcliff, 1993). For all RT analyses, we only included trials in which the current and previous trials were correct to mitigate potential effects of post-error slowing, in which participants tend respond more slowly after making an error (Dutilh et al., 2012; Notebaert et al., 2009). For accuracy analyses, we included all trials in which a response was made within the 1.4s response window. Behavioural data were analysed using 2x2 repeated measures ANOVAs with factors of reward and task sequence. The task sequence factor consisted of switch trials, where the task rule was different from the previous trial, and repeat trials, where the task rule was the same as the previous trial. In addition, we used a paired t-test for normally distributed data and a Wilcoxon signed-rank test for data showing significant deviation from normality, to compare performance on task-switch trials and task-repeat-cue-switch trials. This allowed us to verify that behavioural effects were driven by task set shifts and not simple changes in visual cues (Logan & Schneider, 2006).

EEG pre-processing. EEG data were down-sampled from 1000 to 250Hz and filtered using 40Hz low-pass and 0.01Hz high-pass filters. For each participant, channels with excessive noise were identified by visual inspection and replaced via interpolation, using a weighted average of the

surrounding electrodes. Data were then re-referenced by subtracting the mean activation across all electrodes from each individual electrode at each time point. Data were divided into epochs from -1 to +5 seconds from the onset of the reward cues. Epochs containing artefacts (such as muscle activity) were rejected based on visual inspection. In the final stage of pre-processing, data were subjected to an Independent Component Analysis (ICA). Structured noise components, such as eye blinks, were removed, resulting in the data set used for subsequent analyses. Prior to each analysis, data were z-scored over the trial dimension and baseline-corrected using a time window of 200ms to 50ms prior to the trial event of interest (e.g. cue or target presentation).

EEG analyses. We used RSA (Kriegeskorte et al., 2008) to investigate how reward prospect influenced neural coding for different kinds of task-relevant information. The logic of this approach was to characterise neural coding patterns elicited by different trial conditions and test whether reward prospect led to more distinct task representations for the two abstract tasks being performed (colour vs shape judgements; Etzel et al., 2016; Westbrook & Braver, 2016) or higher neural coding for task-relevant compared to irrelevant target features (Pessoa, 2017). Incorrect and omission trials were excluded from EEG analyses.

Neural coding across the trial. For each participant, trials were divided into conditions based on reward condition (low, high), as well as task-relevant and irrelevant target features (e.g. yellow[rel]-square[irrel]). Dividing the trials this way also implicitly divided trials by task. If the task-relevant target feature was yellow, for example, then the task must be a colour judgement. We then averaged trials in each condition to get an array with channels x time points x conditions. Mahalanobis distances were calculated between all scalp topographies for the sixteen conditions, separately at each time point (using the covariance across all trials at that time point). This procedure yielded a 16 x 16 representational dissimilarity matrix (RDM) of multivariate condition distances for each time point and participant. We then constructed a set of five 16 x 16 model RDMs to capture neural coding patterns related to different task variables. These variables included reward coding, task coding, task-relevant feature coding, task-irrelevant feature coding, and motor coding. The logic of all models was to place zeroes in cells of a 16x16 matrix where conditions matched on the variable of interest and ones in all remaining cells. For example, the task coding model was a 16x16 matrix containing zeros in cells where two conditions involved the same task (e.g. both colour judgements) and ones in cells where two conditions involved different tasks (e.g. shape vs colour judgements). The task-relevant feature model was a 16x16 matrix containing zeros in cells where two conditions had the same task-relevant target feature (e.g. both yellow on colour trials) and ones in cells where two conditions had different task-relevant features (e.g. yellow vs blue or yellow vs square). Data RDMs (not z-scored and averaged for illustration) and model RDMs for all task

variables are presented in Figure 3. Data and model RDMs were z-scored and entered into a multiple regression analysis that was conducted at each time point (4ms apart after down-sampling). This included a constant to model the intercept of the linear regression equation. This regression procedure was performed three times, once with a baseline window prior to reward cue onset, once prior to task cue onset and once prior to the target onset.

Neural coding as a function of reward prospect. We repeated the analyses above separately for high and low reward trials. For each participant, trials were divided based on the task-relevant and irrelevant features of the target. Conditions were then averaged over the trial dimension. Mahalanobis distances were calculated between all scalp topographies for the eight conditions, separately at each time point, generating a set of 8x8 RDMs. For these analyses, model RDMs were generated to reflect condition differences based on task coding, the task-relevant feature of the target, the task-irrelevant feature of the target and the motor response. These models followed the same logic as those described in the previous section: zeroes were placed in cells of an 8x8 matrix where two conditions did not differ in the variable of interest (e.g. the task-relevant target feature) and ones were placed in cells where conditions differed in the variable of interest. Model data RDMs were z-scored and entered into a single regression at each time point (including a constant). The regression procedure was performed twice, once using only low reward trials and once using only high reward trials. For this set of analyses, it was particularly important to match the number of high and low reward trials, so that regression coefficients were not higher for high reward conditions because more data was being included in the regression. To address this, these analyses subsampled trials to match the number of high and low reward trials. 100 iterations were performed per participant and final regression coefficients resulted from averaging the estimated regression coefficients over iterations. To test whether the observed results were driven by correlations between coding models, we ran a control analysis in which we first regressed task coding, task-irrelevant feature coding and motor coding against participants' neural dissimilarity matrices. We then regressed the task-relevant feature coding model against the residual variance, which had not been accounted for by the other models. Finally, we examined reward effects on switch and repeat trials. This was motivated by previous work showing reward prospect benefits switching performance (Shen & Chun, 2011) and can exert a greater performance benefit on switch trials (Kleinsorge & Rinkenauer, 2012), where there is highest interference between task sets and thus where neural coding of the active task set could be most critical in determining behavioural performance. Our own pilot work also showed a significant reduction in RT switch costs on high reward trials. These analyses followed the analysis procedure outlined above but only used switch or repeat trials, instead of all trials.

Relationships between neural coding of task, sensory and motor information. To test whether reward-induced sensorimotor modulations arose from upstream changes in task coding, we performed a series of correlations between average task and sensorimotor coding regression coefficients. To do so, we selected time windows of interest based on results from the previous analysis section. For task coding, we averaged regression coefficients for each participant from 1400-1800ms. We did not include 1200-1400ms in this average due to a stimulus-evoked signal artefact shown in Figure 4B. For feature coding and motor coding, we averaged regression coefficients from 2000-2400ms, where we observed the peak difference between reward conditions (Figures 5A & 6A). For motor coding locked to the response, we averaged regression coefficients -200-0ms from the response, within which we found the peak difference in motor coding between reward conditions. As initial tests, we correlated average task coding with average feature coding and average motor coding, independent of reward. Significant correlations from this first step were followed up by correlating the mean difference in task coding (high-low reward) with the mean difference in the relevant sensorimotor variable (high-low reward). Alpha values for these analyses were corrected for multiple comparisons using the Bonferroni correction.

***Neural coding and cognitive performance*.** To systematically test relationships between neural coding and cognitive performance, we took participants' regression coefficients for different task variables at each time point along the trial and correlated this value with their difference in accuracy (high-low reward) and difference in RT (low-high reward). Except for reward coding (Figure 3C), the neural data for these correlations were differences in regression coefficients between reward conditions. This included the difference in task coding (Figure 4B), relevant feature coding (Figure 5A), the interaction between reward and relevant feature prioritisation (Figure 5C), as well as motor coding locked to the reward cue (Figure 6A) and the response (Figure 6B). Correlations involving reward-differences in neural regression coefficients used the same test windows as their corresponding figure, from which differences were computed. All analyses applied non-parametric Spearman correlations to reduce the influence of outliers in behavioural difference measures, which were more than three scaled median absolute deviations away from the median (RT difference distribution: 2 outliers, lower threshold=-42ms, upper threshold=68ms, median=13ms; accuracy difference distribution: no outliers, lower threshold=-0.06, upper threshold=0.10, median=0.02).

Statistical testing for neural analyses. Data were smoothed with a 12ms Gaussian kernel immediately prior to non-parametric cluster-based permutation testing, which was used to correct for multiple comparisons (Maris and Oostenveld, 2007; Spaak, Watanabe, Funahashi, & Stokes, 2017). 10,000 permutations were performed to generate null distributions for each analysis.

Results

High reward decreased reaction time and improved accuracy

To assess the impact of reward prospect on behavioural performance, we performed a 2x2 repeated measures ANOVA on RT, with factors of reward (low x high) and task sequence (switch x repeat). This revealed a significant main effect of reward, $F(1,29) = 18.676$, $p < 0.001$, $\eta 2p = 0.392$, which reflected lower RTs on high reward trials (M = 415ms, SD = 55.45) compared with low reward trials (M = 438ms, SD = 62.28). There was also a significant effect of task sequence, $F(1,29) = 35.546$, $p < 0.001$, $\eta 2p = 0.551$, driven by lower RTs on repeat trials (M = 412ms, SD = 57.64) compared with switch trials (M = 440ms, SD = 59.36). While switch costs were numerically reduced on high compared with low reward trials, the interaction between reward and task sequence did not reach significance, $F(1,29) = 2.662$, $p = 0.114$, $\eta 2p = 0.084$. A control analysis comparing switch trials with repeat trials involving different task cues confirmed the main effect of task sequence was driven by genuine changes in task-set and not simple changes in visual cues (Shapiro-Wilk test for normality: W=0.822, p=0.003; Wilcoxon T = 438.0, p<0.001, matched rank biserial correlation = 0.884).

The equivalent 2x2 ANOVA for proportion correct showed the same pattern of results. In particular, the analysis showed a significant main effect of reward, $F(1,29) = 14.268$, $p < 0.001$, $\eta 2p = 0.330$, driven by greater accuracy on high reward trials (M = 0.87, SD = 0.06) compared with low reward trials (M = 0.85, SD = 0.07). It also showed a significant main effect of task sequence, $F(1,29) = 51.894$, $p < 0.001$, $\eta 2p = 0.642$, reflecting greater accuracy for repeat (M = 0.89, SD = 0.06) compared with switch trials (M = 0.84, SD = 0.07). We did not detect a significant interaction between reward and task sequence, $F(1,29) = 0.317$, $p = 0.578$, $\eta 2p = 0.011$. Like RT, a control analysis comparing switch trials with repeat trials involving different task cues confirmed the effect of task sequence on accuracy was driven by genuine changes in task-set (Shapiro-Wilk test for normality: W=0.977, p=0.744; t(29)=-6.110, p<0.001, d=-1.116). To summarise, our behavioural analyses show that high reward prospect improved both RT and accuracy performance.

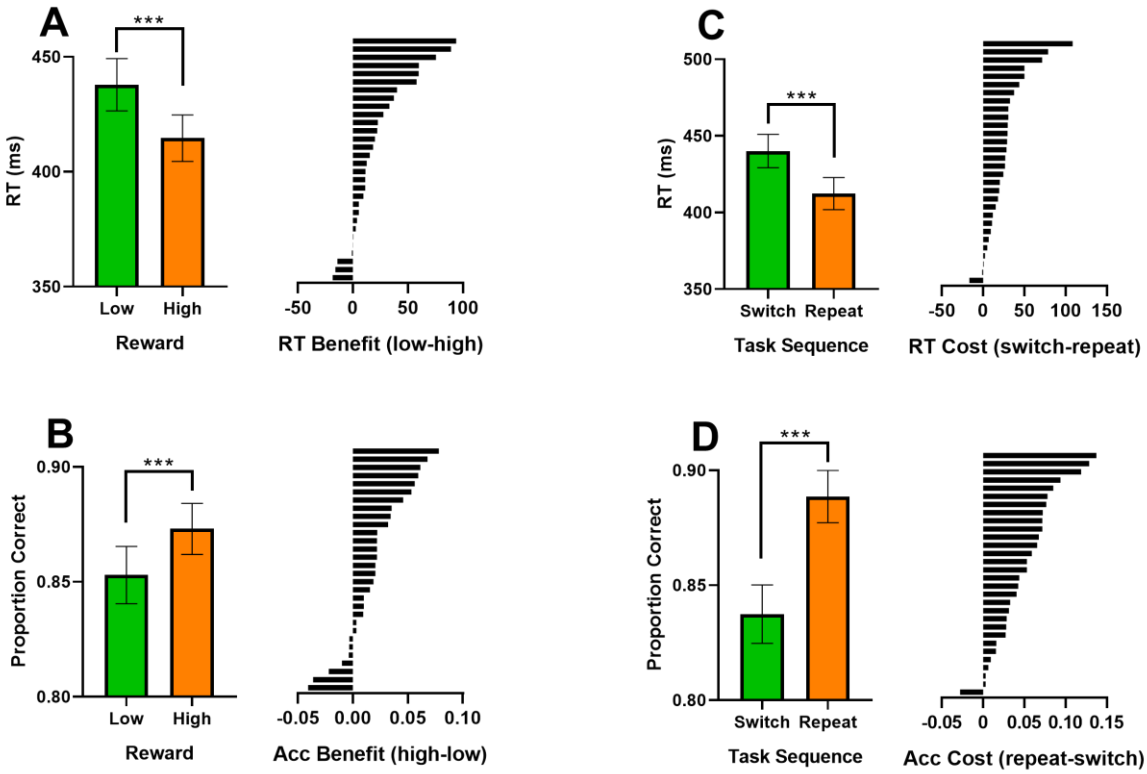

Figure 2. Behavioural performance as a function of reward and task sequence. A: Reaction time for high and low reward trials (left) and the difference between reward conditions for each participant (right). B: Proportion correct (accuracy) for high and low reward trials (left) and the difference between reward conditions per participant (right). C: Reaction time as a function of task sequence (left) and the difference between switch and repeat trials for each participant (right). D: Proportion correct (accuracy) as a function of task sequence (left) and the difference between switch and repeat trials (right) per participant. Error bars show standard error of mean.

Neural coding across the trial

Having established the beneficial impact of reward prospect on cognitive performance, we tested neural coding of task variables across the trial. Reward coding emerged shortly after reward cue onset and was sustained throughout the trial (Figure 3; window tested = 0-3500ms from reward cue onset, cluster = 68-2964ms, p = 0.0002). Task coding peaked shortly after the task rule cue was presented and continued into the response phase (window tested = 1200-3500ms from reward cue onset, first cluster = 1260-2340ms, first cluster p = 0.0002, second cluster = 2412-2944ms, second cluster p = 0.0122). Finally, coding of task-relevant and irrelevant target features, as well as motor response coding rose shortly after target presentation (relevant feature coding: window tested = 1800-4500ms, cluster window = 1900-3020ms, p = 0.0002; irrelevant feature coding: window tested = 1800-4500ms from reward cue onset, cluster window = 1892-2284ms, cluster p = 0.0032; motor

coding: window tested = 1800-4500ms, cluster window = 1816-4076ms, cluster p = 0.0002). In summary, we verified our multivariate analysis approach was sensitive to dynamic temporal changes in neural coding of task variables at sensible stages within the trial.

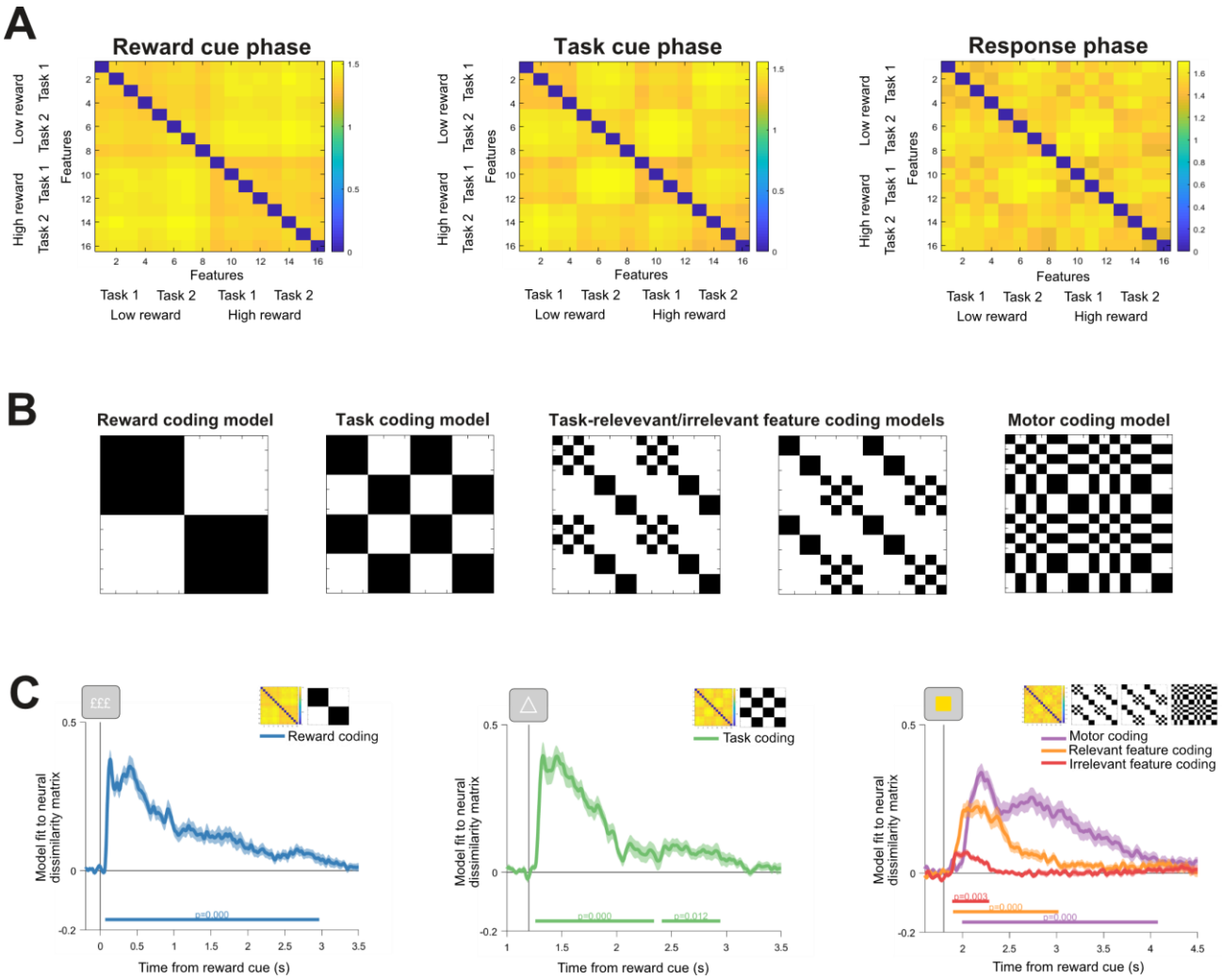

Figure 3. A: Representational dissimilarity matrices averaged across participants and trial periods following: reward cue onset (0-800ms), task cue onset (1200-1800ms) and target onset (1800-2600ms). B: Model representational dissimilarity matrices for: reward coding, task coding, coding of the task-relevant target feature, task-irrelevant target feature and motor coding. C: General linear model regression coefficients from regressing model and neural dissimilarity matrices across time. Shading around principle lines indicates standard error of mean. Cluster-corrected p-values are shown below each time course.

Neural coding as a function of reward prospect

Reward prospect increased proactive task coding.

Having verified core task variables were encoded in the EEG signal at the plausible stages within the trial, we tested which variables were influenced by the reward manipulation. We found that average task coding was significantly greater on high reward trials prior to target onset (Figure 4A: window averaged=1400-1800ms; Shapiro-Wilk test for normality: W=0.958, *p*=0.278; one-tailed t(29)=-2.447, *p*=0.010, *d*=-0.447). Time-resolved permutation analyses confirmed robust encoding of task rules prior to the target under both reward conditions (Figure 4B: window tested=1200-2500ms, low reward cluster=1280-2016ms, low reward *p*=0.0002; high reward cluster=1268-2032ms, high reward *p*=0.0002; biggest difference cluster=1884-1936ms, *p*=0.1566).

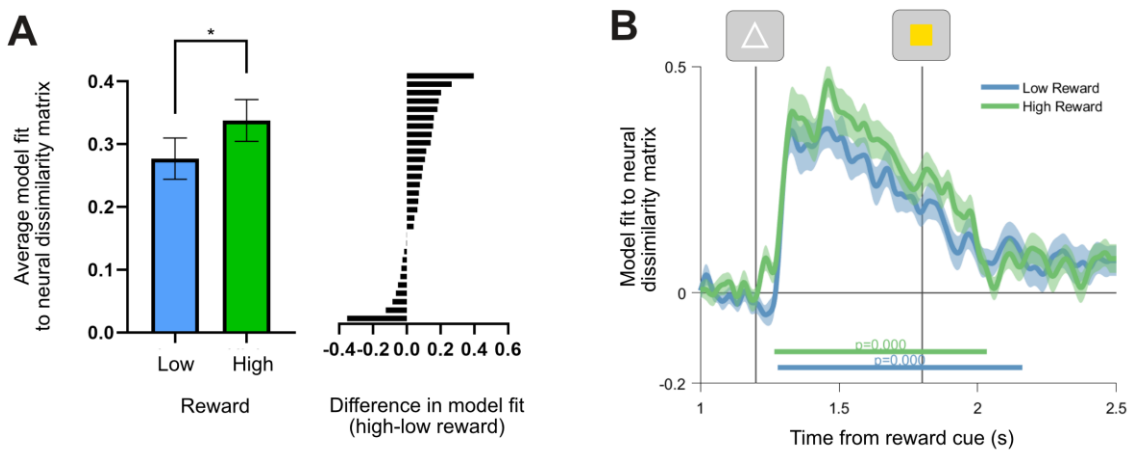

Figure 4. Generalised linear model regression coefficients for task coding as a function of reward. A: coefficients averaged over time points in the pre-target interval (1400-1800ms). The left sub-panel shows group means for average regression betas as a function of reward. The right sub-panel shows the difference between high and low reward regression coefficients for each participant. B: Time resolved regression coefficients for task coding as a function of reward. Vertical lines show onset of the task cue and target respectively. Shading around principle lines indicates standard error of mean. Cluster-corrected p-values are shown below time courses.

Reward prospect selectively increased coding of task-relevant target features.

After evaluating the impact of reward prospect on task coding, we examined the effect of reward prospect on neural coding of task-relevant and irrelevant target features. Relevant feature coding was observed shortly after target presentation on both high and low reward trials (Figure 5A: window tested=1800-3000ms, low reward cluster=1916-2808ms, p=0.0002; first high reward cluster=1912-2652ms, p=0.0002; second high reward cluster=2660-2896, p=0.0192. In addition,

neural coding for task-relevant features was significantly higher under high reward conditions (window tested=1800-3000ms, first difference cluster=2040-2168, p=0.0180; second difference cluster=2252-2408ms, difference p=0.0070). A control analysis which regressed the task-relevant feature model against the residual variance, not explained by any of the other coding models, confirmed this result was not driven by correlated regressors (window tested=1800-3000; low reward cluster=1920-2812ms, p =0.0002; high reward cluster=1912-2896ms, p =0.0002; difference cluster=2036-2168ms, p=0.0178; second difference cluster=2248-2408, p=0.0066). Task-irrelevant information was also represented following target onset for both reward levels (Figure 5B: window tested=1800-3000ms, low reward cluster=2008-2284ms, p=0.0002; second low reward cluster=2464-2568ms, p = 0.0420; high reward cluster=1916-2216ms, p=0.0028). However, the strength of these coding patterns did not differ as a function of reward (window tested=1800-3000ms, no candidate clusters). As a consequence of these target-related effects, reward prospect showed a significant interaction with the difference in task-relevant and irrelevant feature coding. This reflected a greater difference between task-relevant and irrelevant coding under high reward conditions (Figure 5C: window tested=1800-3000ms, first cluster=2032-2176ms, p=0.0220, second cluster=2240-2368ms, p=0.0232). In summary, we found evidence that high reward prospect increased the difference in neural coding for task-relevant and irrelevant target information. This difference was due to a selective increase in the neural coding of task-relevant feature information in high reward contexts.

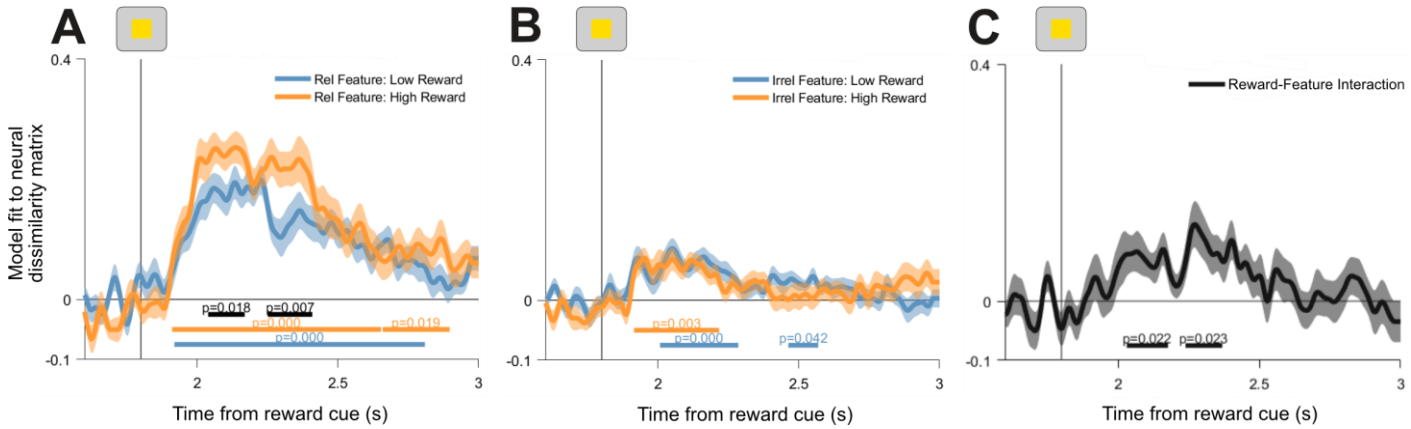

Figure 5. General linear model regression coefficients for coding of target features as a function of reward. A: Task-relevant target feature regression coefficients. B: Task-irrelevant feature regression coefficients. C: Interaction between reward and feature regression coefficients. This time course is computed by taking the difference between task-relevant and task-irrelevant regression coefficients within each reward condition. Resulting time courses are subtracted to show the extent to which high reward increases the difference between task-relevant and irrelevant target features. Vertical lines show target onset. Shading around principle lines indicates standard error of mean. Cluster-corrected p-values are shown below time courses.

Reward prospect increased neural encoding of task-relevant motor output.

After demonstrating reward prospect modulated neural activity coding for task-relevant target features, we examined the effect of reward prospect on activity patterns related to the upcoming motor response. Motor coding appeared after target onset during both high and low reward conditions (Figure 6A: window tested=1800-4500ms, low reward cluster=1992-4008ms, low reward *p*=0.0002; high reward cluster=2016-3940ms, high reward p = 0.0002). When the analysis was locked to the onset of the reward cue, we observed a trend towards higher motor coding under high reward conditions (Figure 6A: window tested=1800-4500ms, difference cluster=2040-2152ms, difference *p* =0.0916). When the analysis was locked to the onset of the motor response itself, we found significantly higher motor coding (figure 6B: window tested =-500-2000ms from response, low reward cluster =-180-1352ms, low reward *p*=0.0002; high reward cluster =-180-1824ms, high reward *p*=0.0002; difference cluster = -88-132ms, difference *p*=0.0296). To summarise, our analyses demonstrate that high reward prospect increased neural coding for task-relevant motor output.

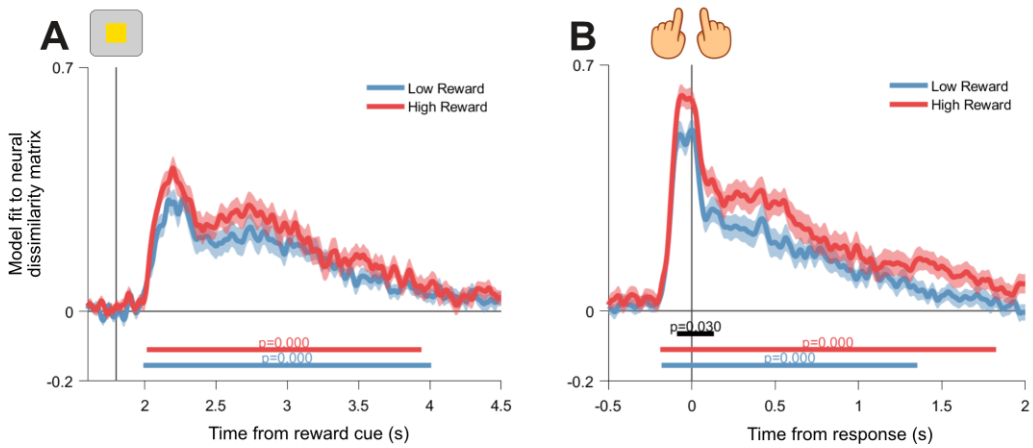

Figure 6. General linear model regression coefficients for coding of the upcoming motor response as a function of reward. A: Motor model regression coefficients for data locked to reward cue onset. The vertical line indicates target onset. B: Motor model regression coefficients for data locked to the response. The vertical line indicates the point responses were made. Shading around principle lines indicates standard error of mean. Cluster-corrected p-values are shown below time courses.

Reward-induced increases in task coding were higher on switch trials.

Having shown reward prospect modulated task, sensory and motor coding, we tested whether the effect of reward on task coding differed on switch trials because previous work has indicated a stronger influence of reward on performance when task sets must be updated (Kleinsorge & Rinkenauer, 2012). This revealed a significant increase in task coding on high reward switch trials compared with low reward switch trials (Figure 7A: window tested=1200-2500, first difference cluster=1452-1584ms, *p*=0.0172; second difference cluster=1704-1916ms, *p*=0.0040). By contrast, we did not detect a difference in task coding as a function of reward on repeat trials (Figure 7B: window tested=1200-2500ms, longest candidate cluster=1224-1244ms, *p*=0.6041). When performing exploratory analyses to examine subsequent reward effects, we did not detect a significant difference in task-relevant feature coding between reward conditions on switch trials (window tested=1800-3000ms, longest candidate cluster=2064-2100ms, *p*=0.3002) or repeat trials (window tested=1800-3000ms, longest candidate cluster=2100-2128ms, *p*=0.4124). For motor coding locked to the response, we found a significant difference in coding between reward conditions on switch trials (window tested=-500-2000ms from response, difference cluster=-76-96ms, *p*=0.0412) and repeat trials (window tested=-500-2000ms from response, first difference cluster=-104-92ms, *p*=0.0332; second difference cluster=344-844ms, *p*=0.0046; third difference cluster=856-1060ms, *p*=0.0354; fourth difference cluster=1304-2000, *p*=0.0026).

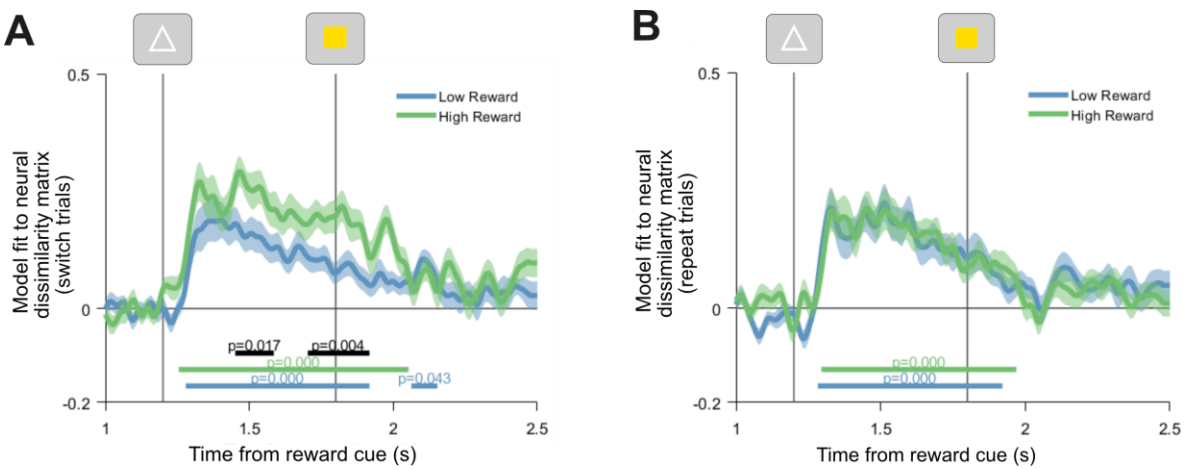

Figure 7. Time resolved regression coefficients for task coding as a function of reward on switch trials (A) and repeat trials (B). Vertical lines show onset of the task cue and target respectively. Shading around principle lines indicates standard error of mean. Cluster-corrected p-values are shown below time courses.

Relationships between neural coding of task, sensory and motor information

Reward-induced changes in task coding were not associated with sensorimotor effects.

Having established that reward prospect modulated both proactive coding of task information and reactive coding of target features and accurate responses, we investigated how proactive and reactive coding were related to one another. To do so, we first tested whether average task coding during the pre-target interval (1400-1800ms) was correlated with average feature coding (2000-2400ms) and average motor coding locked to the reward cue (2000-2400ms) and the response (-200-0ms). Task coding was significantly correlated with coding of the relevant target feature (r=0.5196, p=0.0033) and the difference between task-relevant and irrelevant features (r=0.4815, p=0.0071), indicating that participants with greater task coding also tended to exhibit greater prioritisation of task-relevant target features. We did not detect a significant correlation between task coding and motor coding locked to the reward cue (r=-0.2231, p=0.2360) or the response (r=0.3014, p=0.1055). Having established these basic relationships, we tested whether the difference in average task coding between reward conditions was correlated with the average difference in feature coding between reward conditions. We did not detect a significant correlation between the average difference in task coding (high-low reward) and coding of the relevant target feature (high-low reward) (r=0.2490, p=0.1845). The same pattern of results was found for the average difference in task coding (high-low reward) and the extent to which high reward increased the representation of the relevant target feature over the irrelevant feature (r=0.0949, p=0.6179).

Neural coding and cognitive performance

Neural encoding of reward and task set were associated with performance improvements.

To understand the strong performance benefits observed in behaviour, we first tested whether reward encoding was associated with performance improvements. We found that the magnitude of reward coding was significantly correlated with participants' difference in RT on high compared with low reward trials (Figure 8A: window tested=0-5000ms, first cluster window=0344-1984ms, mean rho=0.4808, p=0.0005; second cluster window=2000-3468ms, mean rho=0.6023, p=0.0003; third cluster window=3492-3896ms, mean rho=0.5059, p=0.0144; fourth cluster window=4020-4532ms, mean rho=0.5066, p=0.0092; fifth cluster=4544-4764ms, mean rho=0.5182, p=0.0352). Control analyses correlating reward coding with RT separately on high or low reward trials confirmed this reward coding association was specific to the difference in performance (window tested=0-5000ms, high reward RT only: longest cluster=1048-1136, mean rho=-0.4308, p=0.2362; low reward RT only: longest cluster=4484-4520ms, mean rho=0.3992, p=0.5454). We did

not detect a significant correlation between reward coding and the difference in accuracy between reward conditions (window tested=0-5000ms, longest candidate cluster=0-20ms, mean rho=-0.3892, *p*=0.7621).

The relationship between task coding and cognitive performance recapitulated the same pattern of results. The difference in task coding between reward conditions was significantly correlated with reward-induced changes in RT (Figure 8B: window tested=1200-2500ms, first cluster window=1856-2000ms, mean rho=0.4994, *p*=0.0043; second cluster window=2060-2100ms, mean rho=0.4708, *p*=0.0375). Control analyses showed this result reflected the difference in RT (window tested=1200-1800ms, high reward RT only: longest candidate cluster=1652-1656ms, mean rho=0.4024, *p*=0.8138; low reward RT only: longest candidate cluster=1628-1676ms, mean rho=0.4922, *p*=0.1466). We did not detect a significant correlation between the difference in task coding and the difference in accuracy between reward conditions (window tested=1200-2500ms, longest candidate cluster=2268-2340ms, mean rho=-0.4643, *p*=0.0677).

In contrast to encoding of reward and task set, we did not detect any significant correlations between reward-induced changes in sensorimotor coding and reward-induced changes in cognitive performance. This was the case for relevant feature coding (window tested=1800ms-3000ms, correlation with RT: longest cluster=1872-1884ms, mean rho=-0.4197, *p*=0.6551; correlation with accuracy: longest cluster=2668-2668ms, *p*=0.8615), the interaction between reward and relevant feature prioritisation (Figure 5C) (window tested=1800-3000ms, correlation with RT: longest cluster=1864-1880ms, mean rho=-0.4167, p=0.6924; correlation with accuracy: no candidate clusters), motor coding locked to the reward cue (window tested=1800-4500ms, correlation with RT: longest candidate cluster=2112-2160ms, mean rho=0.4686, *p*=0.2629; correlation with accuracy: longest candidate cluster=2860-2928ms, mean rho=-0.4132, *p*=0.1904), and motor coding locked to the response (window tested=-500-2000ms from response, correlation with RT: longest cluster=52-156ms, mean rho=0.4143, *p*=0.0978; correlation with accuracy: longest candidate cluster=1672-1772ms, *p*=0.1016).

In summary, we found that neural encoding of reward prospect and the difference in task coding between reward conditions correlated significantly with the benefits of reward prospect on cognitive performance. While reward prospect also modulated encoding of sensorimotor information, we did not find evidence that these changes were associated with reward-induced performance improvements.

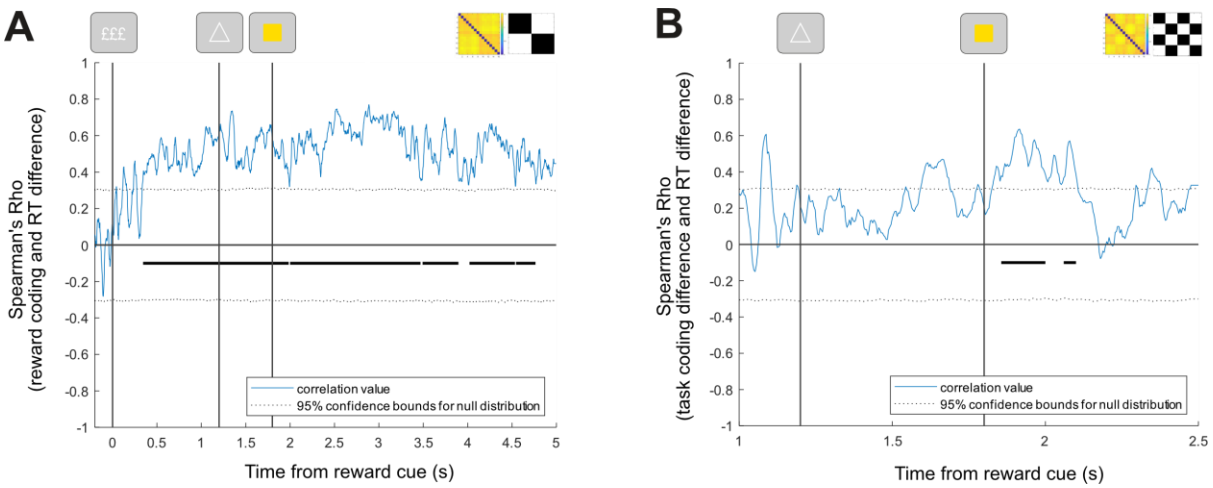

Figure 8. A: Spearman Rho values for correlation between participant reward coding regression coefficients and their difference in reaction time between reward conditions. B: Spearman Rho values for correlation between the difference in task coding regression coefficients (high-low reward) and their difference in reaction time between reward conditions. In both panels, black lines indicate significant correlation clusters, corrected for multiple comparisons using cluster-based permutation testing. Grey dotted lines indicate 95% confidence intervals for the null distribution.

Discussion

The present study aimed to identify stages of neural processing that mediate the performance benefits of reward motivation. We employed RSA to examine changes in neural geometry as the prospect of reward changed dynamically, from trial to trial. Using this method, we were able to track neural representations for reward prospect, abstract task rules, task-relevant and irrelevant perceptual features of target stimuli and neural representations related to accurate motor output. We found that high reward prospect boosted the encoding of multiple task variables. Critically, reward increased encoding of the active task rule in preparation for the target. After the target, we observed increased encoding of task-relevant perceptual information and more distinct neural representations for different behavioural responses. Neural encoding of reward prospect and the reward-related boost in task rule coding correlated with reward-based reductions in RT.

Consistent with the results of previous fMRI decoding studies (Etzel et al., 2016; Qiao, Zhang, Chen, & Egner, 2017; Woolgar et al., 2011; Waskom et al., 2014; Wisniewski et al., 2015), RSA successfully tracked neural representations for task rule information. In using this approach, our results replicate the fMRI findings of Etzel et al. (2016) in human electroencephalographic data, showing that high reward prospect increased average task coding prior to target onset, and that the difference in task coding between reward conditions was associated with improvements in cognitive performance. The effect of reward on task coding was most pronounced on switch trials, where abstract task rules needed to be updated relative to the previous trial. This could suggest that dynamic increases in reward prospect primarily promote flexible rule updating, as opposed to maintenance of existing task rule representations in prefrontal regions. Such updating might be mediated by phasic dopamine release in the striatum, which is thought to be important for driving flexible and targeted updating of contextual information in PFC (Westbrook & Braver, 2016; Yee & Braver, 2018). Overall, these results provide evidence that high reward prospect can improve cognitive performance - at least in part - by increasing neural representations for task rule information.

Consistent with the results from previous studies focusing on neural responses to perceptual targets (Padmala & Pessoa, 2011; Serences, 2008; Serences & Saproo, 2010; Hickey & Peelen, 2015), we found that reward enhanced the representation of task-relevant perceptual features. The present results complement this literature in two ways. First, many previous studies have focused on perceptual stimuli associated with reward over many trials (Serences, 2008; Serences & Saproo, 2010; Hickey & Peelen, 2015). Here we show that transient changes in prospective reward can also modulate task-relevant perceptual feature representations. Second, research using prospective reward cues has led to the proposal that reward motivation might benefit attentional filtering, either

by enhancing task-relevant perceptual representations or supressing task-irrelevant representations (Pessoa, 2017). Hickey & Peelen (2015) found evidence for both of these mechanisms, showing that reward could either enhance or suppress perceptual representations as a function of task-relevance. In the present study, we report a more selective effect on perceptual representations, wherein high reward prospect increased the neural representation of task-relevant information without impacting the representation of irrelevant information. This might suggest that transient reward coupling with sensory target features has a different impact on perceptual encoding than reward-associations established over many trials.

While task coding and the prioritisation of task-relevant target features were strongly correlated, we did not find evidence that reward-driven modulations in these variables were associated. We are cautious not to over-interpret these null-effects. These results do not rule out the possibility that reward-modulated task coding affects downstream perceptual representations. However, they do raise the possibility of an alternative mechanism, wherein reward prospect acts independently on several neural variables. Among these variables, our results suggest that neural encoding of reward prospect and task rule information are important factors associated with dynamic shifts in improving performance. The present results to do not permit conclusive interpretations about the functional role of reward-driven perceptual and motor changes; although we did not detect significant associations between perceptual or motor representations and behavioural measures, this does not imply that these variables were functionally irrelevant to task performance.

How might reward motivation translate into performance improvements more broadly? Previous studies have pointed to the idea that reward motivation might upregulate attention (Etzel et al., 2016; Padmala & Pessoa, 2011; Pessoa & Engelmann, 2010). For instance, reward and attention have been shown to recruit overlapping frontoparietal control regions (Pessoa & Engelmann, 2010) and have analogous effects on electrophysiological signatures of task preparation (van den Berg, Krebs, Lorist, & Woldorff, 2014). One possibility in our study is that reward prospect had additional effects that were not captured by the conditions of our task. For instance, high reward prospect could have increased alertness and temporal attention to information proximal to reward cue presentation. This may explain why associations between behavioural measures and downstream reward effects, such as target feature and motor representations, could then be noisier and less reliable than the strong correlation between behaviour and reward itself.

To conclude, previous work has shown that reward motivation may improve cognitive performance by boosting abstract task codes (Etzel et al., 2016) or biasing perceptual representations to prioritise sensory information relevant to current task goals (Hickey & Peelen,

2015; Pessoa, 2017; Serences, 2008; Serences & Saproo, 2010). Here we show that high reward prospect can enhance neural coding of task information as well as task-relevant sensory information, without altering representations for task-irrelevant sensory information. While reward modulated neural coding for multiple task variables, reward-driven increases in task rule coding were strongly associated with performance improvements. These results suggest that reward motivation can improve cognitive performance by sharpening neural representations for task rule information.

Acknowledgments

This research was funded by a Biotechnology and Biological Sciences Research Council (BB/M010732/1) and James S. McDonnell Foundation Scholar Award (220020405) to Mark G. Stokes, and by the NIHR Oxford Health Biomedical Research Centre. It was also made possible through support from several additional funding groups. Sam Hall-McMaster is funded by the Rutherford Foundation and William Georgetti Trusts, through the New Zealand Government. Paul S. Muhle-Karbe is funded by the Research Foundation Flanders (grant 12R8817N), the Wellcome Trust (grant 210849/Z/18/Z) and Linacre College Oxford (EPA Cephalosphorin Junior Research Fellowship). Nicholas E. Myers is funded by the Wellcome Trust (201409Z/16/Z) and University College Oxford. The Wellcome Centre for Integrative Neuroimaging is supported by core funding from the Wellcome Trust (203139/Z/16/Z).

References

Braem, S., & Egner, T. (2018). Getting a grip on cognitive flexibility. *Current Directions in Psychological Science*, 27(6), 470-476.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9-21.

Dutilh, G., Vandekerckhove, J., Forstmann, B. U., Keuleers, E., Brysbaert, M., & Wagenmakers, E. J. (2012). Testing theories of post-error slowing. *Attention, Perception, & Psychophysics*, *74*(2), 454-465.

Duncan, J. (2001). An adaptive coding model of neural function in prefrontal cortex. *Nature Reviews Neuroscience*, 2(11), 820.

Engelmann, J. B., Damaraju, E., Padmala, S., & Pessoa, L. (2009). Combined effects of attention and motivation on visual task performance: transient and sustained motivational effects. *Frontiers in Human Neuroscience*, 3, 4.

Etzel, J. A., Cole, M. W., Zacks, J. M., Kay, K. N., & Braver, T. S. (2015). Reward motivation enhances task coding in frontoparietal cortex. *Cerebral cortex*, *26*(4), 1647-1659.

Hickey, C., & Peelen, M. V. (2015). Neural mechanisms of incentive salience in naturalistic human vision. *Neuron*, *85*(3), 512-518.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, *36*(14), 1.

Kleinsorge, T., & Rinkenauer, G. (2012). Effects of monetary incentives on task switching. *Experimental Psychology*. 59, 216–226.

Krebs, R. M., Boehler, C. N., Roberts, K. C., Song, A. W., & Woldorff, M. G. (2011). The involvement of the dopaminergic midbrain and cortico-striatal-thalamic circuits in the integration of reward prospect and attentional task demands. *Cerebral cortex*, 22(3), 607-615.

Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*, 4.

Logan, G. D., & Schneider, D. W. (2006). Interpreting instructional cues in task switching procedures: The role of mediator retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(2), 347.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177-190.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24(1), 167-202.

Notebaert, W., Houtman, F., Van Opstal, F., Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: an orienting account. *Cognition*, *111*(2), 275-279.

Padmala, S., & Pessoa, L. (2011). Reward reduces conflict by enhancing attentional control and biasing visual cortical processing. *Journal of Cognitive Neuroscience*, 23(11), 3419-3432.

Parro, C., Dixon, M. L., & Christoff, K. (2018). The neural basis of motivational influences on cognitive control. *Human Brain Mapping*, 39(12), 5097-5111.

Pessoa, L. (2017). Cognitive-motivational interactions: Beyond boxes-and-arrows models of the mind-brain. *Motivation Science*, *3*(3), 287.

Pessoa, L., & Engelmann, J. B. (2010). Embedding reward signals into perception and cognition. *Frontiers in Neuroscience*, *4*, 17.

Qiao, L., Zhang, L., Chen, A., & Egner, T. (2017). Dynamic Trial-by-Trial Re-Coding of Task-Set Representations in Frontoparietal Cortex Mediates Behavioral Flexibility. *Journal of* Neuroscience, 0935-17.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, 86(3), 446.

Shen, Y. J., & Chun, M. M. (2011). Increases in rewards promote flexible behavior. *Attention, Perception, & Psychophysics*, *73*(3), 938-952.

Serences, J. T. (2008). Value-based modulations in human visual cortex. *Neuron*, *60*(6), 1169-1181.

Serences, J. T., & Saproo, S. (2010). Population response profiles in early visual cortex are biased in favor of more valuable stimuli. *Journal of neurophysiology*, *104*(1), 76-87.

Spaak, E., Watanabe, K., Funahashi, S., & Stokes, M. G. (2017). Stable and dynamic coding for working memory in primate prefrontal cortex. *Journal of Neuroscience*, 3364-16.

van den Berg, B., Krebs, R. M., Lorist, M. M., & Woldorff, M. G. (2014). Utilization of reward-prospect enhances preparatory attention and reduces stimulus conflict. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(2), 561-577.

Waskom, M. L., Kumaran, D., Gordon, A. M., Rissman, J., & Wagner, A. D. (2014). Frontoparietal representations of task context support the flexible control of goal-directed cognition. *Journal of Neuroscience*, *34*(32), 10743-10755.

Wisniewski, D., Reverberi, C., Momennejad, I., Kahnt, T., & Haynes, J. D. (2015). The role of the parietal cortex in the representation of task–reward associations. *Journal of Neuroscience*, *35*(36), 12355-12365.

Woolgar, A., Hampshire, A., Thompson, R., & Duncan, J. (2011). Adaptive coding of task-relevant information in human frontoparietal cortex. *Journal of Neuroscience*, *31*(41), 14592-14599.