Functional MRI signatures of Pavlovian and instrumental valuation systems during a modified orthogonalized go/no-go task.

Abbreviated title:

fMRI signatures of Pavlovian and instrumental value

Authors:

Filippo Queirazza[1,4,*], J. Douglas Steele[2], Rajeev Krishnadas[4], Jonathan Cavanagh[3] and Marios G. Philiastides[4,*]

Affiliations:

[1]School of Health and Wellbeing, University of Glasgow, Glasgow, G12 0XH, UK

[2]Division of Imaging Science and Technology, University of Dundee, Dundee, DD1 9SY, UK

[3]School of Infection, Immunity and Inflammation, University of Glasgow, Glasgow, G12 8TA, UK

[4]School of Psychology and Neuroscience, University of Glasgow, Glasgow, G12 8AD, UK

***Correspondence:**

Filippo.Queirazza@glasgow.ac.uk; Marios.Philiastides@glasgow.ac.uk

Number of pages: 56

Number of figures: 7

Number of words (Abstract): 250

The authors declare no competing financial interests.

Abstract

Motivational (i.e. Pavlovian) values interfere with instrumental responding and can lead to suboptimal decision-making. In humans, task-based neuroimaging studies have only recently started illuminating the functional neuroanatomy of Pavlovian biasing of instrumental control. To provide a mechanistic understanding of the neural dynamics underlying the Pavlovian and instrumental valuation systems, analysis of neuroimaging data has been informed by computational modelling of conditioned behaviour. Nonetheless, due to collinearities in Pavlovian and instrumental predictions, previous research failed to tease out haemodynamic activity that is parametrically and dynamically modulated by coexistent Pavlovian and instrumental value expectations. Moreover, neural correlates of Pavlovian to instrumental transfer effects have so far only been identified in extinction (i.e. in the absence of learning). In this study we devised a modified version of the orthogonalized go/no-go paradigm which introduced Pavlovian only catch trials to better disambiguate trial-by-trial Pavlovian and instrumental predictions in both sexes. We found that haemodynamic activity in the ventromedial prefrontal cortex covaried uniquely with the model-derived Pavlovian value expectations. Notably, modulation of neural activity encoding for instrumental predictions in the supplementary motor cortex was linked to successful action selection in conflict conditions. Furthermore, haemodynamic activity in regions pertaining to the limbic system and medial prefrontal cortex was correlated with synergistic Pavlovian and instrumental predictions and improved conditioned behaviour during congruent trials. Altogether, our results provide new insights into the functional neuroanatomy of decision-making and corroborate the validity of our variant of the orthogonalized go/no-go task as a behavioural assay of the Pavlovian and instrumental valuation systems.

Introduction

According to the two-process theory of associative learning the interaction between the Pavlovian and instrumental valuation systems underpins adaptive behaviour (Mackintosh, 1983; Rescorla & Solomon, 1967). Correspondingly, the Pavlovian corrupting influence on optimal instrumental control is thought to underlie well-known behavioural anomalies observed in animals (i.e. autoshaping (Brown & Jenkins, 1968), negative automaintenance (Williams & Williams, 1969)) and humans (i.e. framing (De Martino, Kumaran, Seymour, & Dolan, 2006) and endowment (Kahneman, Knetsch, & Thaler, 1990) effect). Furthermore, there is growing empirical evidence that Pavlovian learning biases may account for maladaptive behaviours associated with depression (Dayan & Huys, 2008; Huys et al., 2016), addiction (Robinson & Berridge, 2003), trauma (Ousdal et al., 2018) and other common mental disorders (Garbusow et al., 2022).

While the Pavlovian system supports the acquisition of associations between temporally contiguous stimuli and outcomes, the instrumental system enables adaptive learning of stimulus-response-outcome pairings. Crucially, since hard-wired, Pavlovian-mediated preparatory responses (i.e. approach versus avoidance) are tightly tied to outcome valence (i.e. reward versus punishment), the Pavlovian system invariably prescribes approach to reward-predicting stimuli and avoidance of punishment-predicting stimuli. Conversely, the instrumental system flexibly selects actions based on their contingent outcomes and independently of outcome valence. Therefore, when the required action is not congruent with the anticipated outcome valence (i.e. approach punishment-predicting stimuli / avoid reward-predicting stimuli) Pavlovian and instrumental predictions diverge and compete for behavioural control. Alternatively, when the required action and anticipated outcome valence are congruent

(i.e. approach reward-predicting stimuli / avoid punishment-predicting stimuli), Pavlovian and instrumental predictions converge and exert synergistic effects on behaviour.

The two most popular experimental paradigms that leverage the asymmetries of Pavlovian and instrumental predictions as a function of the two axes of behavioural control (i.e. valence and action) are the Pavlovian to instrumental transfer (PIT) (Cartoni, Puglisi-Allegra, & Baldassarre, 2013) and orthogonalized go/no-go task (Crockett, Clark, & Robbins, 2009). In the PIT task Pavlovian and instrumental predictions coupled with different cues are acquired in separate experimental stages and transfer effects are subsequently tested in extinction (i.e. in the absence of new learning). In the orthogonalized go/no-go task Pavlovian and instrumental predictions coupled with the same cues are continuously updated via probabilistic feedback.

To illuminate the functional neuroanatomy of the Pavlovian-instrumental dichotomy human studies have exploited haemodynamic responses to Pavlovian and instrumental predictions during acquisition of functional magnetic resonance imaging (fMRI) data (Guitart-Masip, Huys, et al., 2012). Moreover, analysis of fMRI recordings has capitalised on the mechanistic insights afforded by the use of reinforcement learning (RL) theory to account for experimentally elicited motor responses (Guitart-Masip, Huys, et al., 2012). In the most successful implementations of this modelling work the linear superposition of Pavlovian and instrumental values into a common decision variable guides action selection (Guitart-Masip, Huys, et al., 2012; Huys et al., 2011). Computational models of the PIT effect are usually built on the offline linear combination of static (learnt) instrumental and (fitted) Pavlovian value expectations and thus parametric modulation of haemodynamic responses as a function of PIT effects is based on behavioural measures such as the strength of instrumental

104 responding (Geurts, Huys, den Ouden, & Cools, 2013). Although RL models of the
105 orthogonalized go/no-go task incorporate online updating of Pavlovian and
106 instrumental values by means of separate prediction errors, the time courses tracking
107 their temporal evolution are highly collinear, thus precluding fMRI modelling of the
108 neural correlates of Pavlovian learning biases (Guitart-Masip, Huys, et al., 2012).
109 In this study we have addressed this shortcoming and developed a modified version of
110 the orthogonalized go/no-go task to effectively tease apart Pavlovian and instrumental
111 value representations and their interaction in the brain. We included catch trials
112 whereby only Pavlovian expectations were being updated and no instrumental
113 responding was required. We demonstrated parametric encoding of instrumental and
114 Pavlovian predictions in the supplementary motor (SMC) and ventromedial prefrontal
115 cortex (vmPFC) respectively and of synergistic Pavlovian by instrumental interaction
116 (PII) effects in regions pertaining to the limbic system. Crucially, we revealed brain-
117 behaviour correlations further validating the role of these regions in modulating
118 learning and choice behaviour.
119

Materials and Methods

Participants. We recruited forty-five participants for our study. Five participants were excluded due to significant incidental clinical brain imaging finding (n=1) or due to poor task performance (i.e. choice accuracy < 55%) (n=3) or due to technical problems with data acquisition (n=1). The remaining forty subjects (32 females) were included in the analyses presented in this paper. Participants were aged between 18 and 60 years (mean= 22.07, sd ± 2.32). All participants provided written, informed consent. Based on a previously documented large behavioural effect size (Cohen's d = 1.4) (Guitart-Masip, Huys, et al., 2012), we estimated a sample size of 40 subjects would be sufficient to achieve 80% power at an alpha level of 0.05. The study protocol was approved by the University of Glasgow College of Science and Engineering Ethics Committee (300160098).

Experimental procedures. To better tease apart the dynamic (that is, trial-by-trial) neural representations of instrumental and Pavlovian systems in the presence of learning we modified the popular orthogonalized go/no-go task (Guitart-Masip, Huys, et al., 2012). Indeed, the classic version of this task does not allow a straightforward dissociation of Pavlovian and instrumental predictions and previous work highlighted this shortcoming (Guitart-Masip, Huys, et al., 2012). Likewise, classic PIT paradigms are designed to elicit Pavlovian effects on instrumental responding in the absence of new learning and thus are not suited to investigate the 'dynamic' effects of the Pavlovian and instrumental systems on decision making (Huys et al., 2011).

Our task consisted of four blocks. Each block had 40 mixed (i.e. instrumental and Pavlovian) and 20 Pavlovian only trials (60 trials per block and 240 trials in total) (Figure 1A-B). Within each block mixed trials were randomly interspersed with

Pavlovian only trials. In the mixed trials there were three events: presentation of a fractal cue, target detection and probabilistic outcome. There were four fractal cues and each denoted a specific combination between action requirement (i.e., go versus no-go) and outcome valence (i.e., win versus lose). The association between the four fractals and the resulting combinations of action and valence (i.e., go to win / no-go to win / go to avoid losing / no-go to avoid losing) were randomised across participants. In the target detection phase a circle was shown on either the right- or left-hand side of the screen (target positions were counterbalanced across trials within each block). In the outcome phase, possible outcomes were an upward pointing green arrow (1 point) or a horizontal yellow bar (0 point) in the win trials and a downward pointing red arrow (-1 point) or a horizontal yellow bar (0 point) in the lose trials. Response-outcome contingencies were probabilistic as shown in Figure 1B. Correct responses were rewarded with the best possible outcome (i.e., green arrow for the win cues and yellow bar for the lose cues) 80% of the time. Reverse outcome contingencies were applied to incorrect responses. In the Pavlovian only trials presentation of a fractal cue was directly followed by a probabilistic outcome (Figure 1B). Outcomes were selected as if participants had made the correct response. We did not yoke outcomes in the Pavlovian only trials to the individual win/lose outcome rates on the mixed trials so that the updating of the Pavlovian value was not tied to instrumental performance during the Pavlovian only trials. This experimental manipulation allowed us to better decorrelate Pavlovian and instrumental predictions. Moreover, we capitalised on the trial-by-trial variability of fitted Pavlovian and instrumental predictions to enhance power to detect covarying fMRI activations. Fractal cues were counterbalanced across both mixed and Pavlovian only trials.

Participants were instructed that during mixed trials correct responses could be either 'go' or 'no-go'. For the 'go' responses they were advised to press either a right or left button depending on the side of the screen the cue was shown on. They were informed of the probabilistic nature of the task and had to learn stimulus-response-outcome contingencies by trial and error. Participants were advised they could win up to £10 based on their task performance. Moreover, they were given the opportunity to practice an example block of the task outside the scanner so that they could familiarise themselves with the speed requirements of the task. The task was programmed using Presentation® (Neurobehavioural Systems) stimulus delivery software.

Behavioural analyses. For our behavioural analyses we conducted maximal by-subject random intercept and random slopes generalised and loglinear mixed-effects models (Barr, Levy, Scheepers, & Tily, 2013) using the *lme4* package in R (http://www.r-project.org) and allowing for random correlations between independent variables. We tested the statistical significance of the fixed effects using the likelihood ratio test (Barr et al., 2013). Moreover, in the presence of an interaction between a factorial predictor and a continuous covariate, we derived estimates of the covariate slopes for each level of the factorial predictor and tested their statistical significance using the *emmeans* package in R. We dealt with non-convergence issues due to model singularity (or near-singularity) by dropping terms in the random effects structure of the model.

To test task related learning effects as a function of task block, valence and action requirement we conducted the following mixed-effects regression model:

$$logit(Accuracy) = 1 + Block + Valence * Action + (1 + Block + Valence * Action|Subject)$$

To further assess behavioural effects of motivational biases we regressed valence and accuracy on response times (RT) as per the following model:

$$log(RT) = 1 + Valence * Accuracy + (1 + Valence * Accuracy|Subject)$$

Computational modelling of behavioural data. We fitted a nested sequence of differently parameterised reinforcement learning models. On each trial t action weight W represented the expected value assigned to go and no-go responses for a given cue i and determined propensity p for action $a \in \{go, no-go\}$ according to the following decision function:

$$p(a_{go}^t|\, cue_i^t) = \sigma\big((W_{go}^t|cue_i^t) - (W_{no-go}^t|cue_i^t)\big)(1 - noise) + \frac{noise}{2}$$

where σ() is standard sigmoid function and noise is a free parameter that can vary between 0 and 1. The expected value Q for instrumental actions was updated according to a Rescorla-Wagner (RW) learning rule parameterised as follows:

$$Q_i^t(a^t|\, cue_i^t) = Q_i^{t-1}(a^{t-1}|\, cue_i^{t-1}) + \alpha\,(\rho r^t - Q_i^{t-1}(a^{t-1}|\, cue_i^{t-1}))$$

where $r^t \in \{1, 0, -1\}$ represents the outcome, α is the learning rate and ρ is an outcome sensitivity parameter. Whilst in the base model RW (learning rate + noise) we did not include a sensitivity parameter, in other model parameterisations we included either a single sensitivity parameter ρ for both reward and punishment or two distinct sensitivity parameters allowing for differential scaling of reward (ρ_rew) and

punishment (ρ_pun) outcomes. To account for the observed tendency to favour go over no-go responses, especially in the early trials of the task, we incorporated a time-invariant, fixed, go bias parameter b into action weight W as follows:

$$W_i^t(a^t|\, cue_i^t) \begin{cases} Q_i^t(a^t|\, cue_i^t) + b \;\; if \;\; a^t = go \\ Q_i^t(a^t|\, cue_i^t) \qquad\qquad else \end{cases}$$

We also allowed for the initial instrumental value Q^0 of all cues to be a free parameter; otherwise we set it to 0. The initial Pavlovian value V^0 of all cues was set to 0. To estimate the biasing effect of the Pavlovian system over instrumental learning we integrated Pavlovian expected value V in the update equation of action weight W as follows:

$$W_i^t(a^t|\, cue_i^t) \begin{cases} Q_i^t(a^t|\, cue_i^t) + b + \pi V_i^t(cue_i^t) \;\; if \;\; a = go \\ Q_i^t(a^t|\, cue_i^t) \qquad\qquad\qquad\qquad else \end{cases}$$

where π is a free parameter indexing the magnitude of Pavlovian bias (the greater π, the greater the influence of the Pavlovian system on instrumental responding). The Pavlovian expected value V was updated according to the following equation:

$$V_i^t(cue_i^t) = V_i^{t-1}(cue_i^{t-1}) + \alpha \left(\rho r^t - V_i^{t-1}(cue_i^{t-1})\right).$$

Given that the sign of the Pavlovian expected value V depended on the valence of each cue (that is, positive for win cues and negative for lose cues), V enhanced instrumental responding by increasing the value of the go action during the win trials. Conversely,

V disrupted instrumental responding by decreasing the value of the go action during the lose trials. Moreover, in a further variant of the model we employed two different Pavlovian biases π (approach and avoidance) to estimate the differential effect of Pavlovian invigoration (π_app) and suppression (π_avd) on go responses. Finally, we tested the additional hypotheses that i) the Pavlovian biasing of the go action was restricted to the win trials and ii) a static (as opposed to a dynamically learnt) Pavlovian value (ψ), which can only be inferred upon the first non-neutral outcome, contributed to the action weights (Swart et al., 2018). To preserve the parameters' natural bounds, log (ρ, π, ψ) and logit (noise, α) transforms of the parameters were implemented. We set the initial value of the free parameters' $(\mathrm{noise}, \alpha, b, \rho, \rho_{\mathrm{rew}}, \rho_{\mathrm{pun}}, \pi, \pi_{\mathrm{app}}, \pi_{\mathrm{avd}}, Q^0, \psi)$ prior means in their native space to (0.5,0.5,0,1,1,1,1,1,1,0,1) and their prior variances to 100.

Importantly, while updating of Pavlovian value in the absence of any instrumental learning during the Pavlovian only trials enabled partial decorrelation of Pavlovian and instrumental value estimates, Pavlovian learning was still tied to instrumental performance during mixed trials.

Model fitting and validation. To compute parameter estimates we implemented a type II maximum likelihood fitting procedure as previously described in (Huys et al., 2011). We optimised the log likelihood of observed data Y by performing k iterations of an expectation-maximization routine until convergence. Briefly, at each iteration k, in the expectation step we optimised the log likelihood with respect to the distribution over the parameters θ holding prior parameters η fixed:

$$q^k(\theta) = p(\theta|Y, \eta^{k-1})$$

We used the Laplace approximation for $q^k(\theta) \sim N = (m^k, s^k)$ and for each subject i updated the mean m and variance s of the normal distribution as follows:

$$m_i^k = \underset{\theta}{argmax}\ p(Y_i, \theta_i | \eta^{k-1})$$

$$s_i^k = \left(\frac{\partial^2 p(Y_i, \theta_i | \eta^{k-1})}{\partial \theta_i^2} |_{\theta_i = m_i^k} \right)^{-1}$$

Subsequently in the maximization step we optimised the log likelihood with respect to prior parameters η holding the distribution over the parameters θ fixed:

$$\eta_m^k = \frac{1}{N} \sum_i^N m_i^k$$

$$\eta_s^k = \frac{1}{N} \sum_i^N (m_i^k)^2 + s_i^k - \eta_m^k$$

The advantage of this hierarchical approach was to prevent overfitting and avoid noisy parameters estimates since poorly constrained parameters were regularised by prior parameters.

We estimated the log-likelihood as the cross-entropy loss function:

$$\sum^N y log\hat{y} + (1 - y) log(1 - \hat{y})$$

where y and $\hat{y}$ represent observed and predicted choices respectively.

To verify the model's goodness of fit we computed the 20% bend correlation coefficient between observed and model-predicted group-level trial-wise probabilities of go action. Furthermore, we performed parameter recovery and thus tested whether parameters of the best fitting model were identifiable. We simulated new data using fitted parameters and estimated parameters again on simulated data. Subsequently, we computed the 20% bend correlation coefficient between true and recovered parameters.

Model comparison and model falsification. To select the best fitting model, we evaluated both predictive and generative performance of the candidate models (Palminteri, Wyart, & Koechlin, 2017). We initially performed model comparison by estimating the group-level BIC_{int} for N individuals as described in (Huys et al., 2011). Briefly, to estimate the log model evidence $p(Y|\eta)$ we approximated the integral over the parameters by sampling θ from the prior distribution $N(\eta_m, \eta_s)$ 1000 times:

$$\int logp(Y|\theta)logp(\theta|\eta)d\theta \sim \frac{1}{K}\sum^{K} logp(Y|\hat{\theta})$$

$$BIC_{int} = \sum^{N} logp(Y|\hat{\theta}) - \frac{1}{2}|\theta|log|Y|$$

where $\hat{\theta}$ is the sampled parameters, $|\theta|$ is the number of parameters in the model and $|Y|$ is the number of data points. The BIC_{int} represents a parsimonious estimate of a model's goodness-of-fit based on both optimised parameters and hyperparameters. Nonetheless, the BIC_{int} is a relative measure of a model's goodness-of-fit and only provides information on whether a given model outperforms competing modelling hypotheses.

To assess generative performance, we subsequently simulated behavioural data and compared observed and simulated trial- and cue-wise group-level probabilities of the go action. To generate surrogate behavioural data we played the task 100 times resampling fitted parameters without replacement. We stochastically determined action using model-derived choice propensities and randomly selected outcomes according to the ground-truth feedback schedule. We then averaged group-level trial-wise probabilities of the go action and correlated it with observed probability of the go action using 20% bend correlation test.

fMRI data acquisition. We used a 3-Tesla Siemens TIM Trio MRI scanner (Siemens, Erlangen, Germany) with a 12-channel head coil to record MRI data. Cushions were placed around the head to minimize head motion. We acquired a high-resolution T1-weighted structural image (1 mm isotropic voxels, 128 axial slices, TI=900 ms, TR=2300 ms, TE=2.96 ms, flip angle=90^{o}), a T2*-weighted echo planar imaging (EPI) functional scan (2 mm isotropic voxels, 68 axial slices, TR=2000 ms, TE=26 ms, flip angle=60^{o}) using multiband 2 acquisition, phase and magnitude fieldmaps (3.3 x 3.3 x 3 mm voxels, 46 axial slices, TR=488 ms, short TE=4.92 ms, long TE=7.38 ms) for distortion correction of the acquired EPI images (Weiskopf, Hutton, Josephs, & Deichmann, 2006). Slice orientation was tilted -30^{o} from the AC-PC plane to reduce susceptibility induced signal drop out (Weiskopf et al., 2006).

fMRI data pre-processing and analysis. MRI data were pre-processed and analysed using FSL software (Smith et al., 2004). The pre-processing pipeline involved B0 unwrapping (Jenkinson, 2003), intra-modal motion correction using MCFLIRT (Jenkinson, Bannister, Brady, & Smith, 2002), slice timing correction, spatial

smoothing with an isotropic 5 mm FWHM Gaussian kernel, high-pass temporal filtering with 110 sec. cut-off frequency and grand-mean intensity normalisation of each entire 4D dataset. EPI scans were subsequently co-registered with skull-stripped structural images using boundary-based registration (FLIRT) (Greve & Fischl, 2009; Jenkinson & Smith, 2001) and spatially normalised into MNI152 space using FNIRT non-linear registration.

fMRI data analysis. We performed whole brain statistical analyses of fMRI data using a multilevel mixed-effects approach as implemented in FLAME1+2 (FSL) (Beckmann, Jenkinson, & Smith, 2003). Regressors were convolved with a double gamma hemodynamic response function. Six additional motion parameters (three translations and three rotations) estimated during the motion correction phase were included in the design matrix as regressors of no interest. We used FSL collinearity diagnostics to ensure design matrices were well conditioned and not rank-deficient. At the first level we estimated contrasts of the parameter estimates specified in the design matrix. At the second and third level we conducted a one-sample t-test of the lower-level contrasts of the parameter estimates to account for between-blocks (second level) and between-subjects (third level) random effects. We thresholded Z statistic images using a cluster-defining threshold of $Z>3.1$ and a FWE-corrected significance threshold of $p=0.05$.

To elucidate the neural circuitry underpinning Pavlovian and instrumental valuation systems and PII effects we performed two sets of fMRI analyses: *task-informed* and *model-informed* analyses. While in the task-informed approach we built a general linear model (GLM) using only task events, in the model-informed approach we built regressors from the Pavlovian and instrumental value estimates (and their respective prediction errors) of our winning computational model fit. The goals of the model-

informed fMRI analyses were i) to validate modelling results and ii) to uncover parametric hemodynamic responses associated with instrumental and Pavlovian predictions and PII effects.

Task-informed GLM. The regressors of interest in the design matrix were four unmodulated boxcar functions, each aligned with and lasting for the duration of the cue presentation. Moreover, nuisance regressors included two boxcar regressors modelling response-time modulated go and unmodulated no-go responses in the target detection phase, one modulated boxcar regressor modelling outcome (i.e. +1 for rewards, 0 for neutral outcomes and -1 for punishments) and one unmodulated regressor modelling late response trials. We then used the parameter estimates of the cue-wise regressors to set up linear contrasts designed to capture main effects of action requirement (go vs no-go cues), inaction (no-go vs go cues), positive valence (win vs lose cues), negative valence (lose vs win cues), congruence (congruent vs incongruent cues) and incongruence (incongruent vs congruent cues).

Model-informed GLM. In this fMRI analysis we capitalised on the results of computational modelling and constructed a GLM to investigate the main effects of instrumental and Pavlovian value expectations and PII effects. We modelled the cue presentation phase of the task by building two parametric regressors encoding the instrumental value of action (i.e. Q_{go}-$Q_{no\text{-}go}$) associated with the presented cue (instrumental regressor) and its motivational value (Pavlovian regressor). It is important to note that the linear (positive) contrast of the instrumental regressor is action specific as it captures differential haemodynamic responses to the go action value compared to the no-go action value (and vice versa with the negative contrast). We modelled PII effects by means of a third regressor representing the element-wise product of the magnitude of the instrumental and Pavlovian regressors. The resulting signed

389 interaction regressor was positive in the case of congruent (i.e. same-sign) instrumental
390 and Pavlovian regressors and negative in the case of incongruent (i.e. different-sign)
391 instrumental and Pavlovian regressors. We conducted a positive and negative linear
392 contrast of the interaction regressor to uncover synergistic (positive contrast) versus
393 antagonistic (negative contrast) PII neural effects over and above Pavlovian and
394 instrumental main effects. It is important to note that our interaction regressor conflated
395 cooperation (and competition) interaction effects across appetitive and aversive
396 domains. Notably, we orthogonalized the PII regressor with respect to the instrumental
397 and Pavlovian regressors. In the outcome phase we built two parametric regressors
398 encoding instrumental and Pavlovian prediction errors. As in the task-informed GLM
399 we also included nuisance regressors accounting for visual stimulation in the decision
400 and outcome phase, go and no-go responses in the target decision phase and late
401 response trials. Moreover, we performed a supplementary analysis where we included
402 a nuisance modulated regressor (i.e. +1 for go and -1 for no-go responses) modelling
403 motor response in the cue presentation phase. The purpose of this analysis was to
404 control for any confounding effects of preparatory motor activity pertaining to the
405 encoding of the instrumental value of action. While in all GLMs we modelled the
406 outcome phase, the design of the task was optimised to uncover haemodynamic
407 responses to cue presentation. We therefore focused our analysis on the presentation
408 phase and only discuss findings pertaining to it.
409 We also conducted region of interest (ROI) analyses based on previous findings in the
410 literature implicating the striatum and inferior frontal gyrus (IFG) in instrumental and
411 Pavlovian learning (Guitart-Masip, Chowdhury, et al., 2012; Guitart-Masip et al., 2011;
412 Guitart-Masip, Huys, et al., 2012). Using FLS's Harvard-Oxford Cortical and
413 Subcortical Structural Atlases, we created anatomical masks of the striatum and IFG,

thresholded at 50%, and used them to mask second-level contrasts images of interest, which we then entered into a third level one-sample t-test where we performed cluster-level inference as in the whole-brain analyses.

Time course analysis of fMRI data. We conducted follow-up analyses and for each subject we extracted BOLD signal time courses from the FWE-corrected significant clusters identified by the model-informed GLM and therefore encoding instrumental and Pavlovian expected value and PII effects. The aim of these follow-up analyses was threefold: i) to visualise cluster-wise valence/action/congruence effects, ii) to correlate cluster-wise mean activity with subject-wise behavioural performance and iii) to predict trial-by-trial individual behaviour (i.e. choice accuracy and reaction times). For these analyses, we reverse normalised masked clusters of interest from standard into functional space to retrieve cluster- and trial-wise BOLD activity from subject-specific pre-processed functional scans. We then estimated cluster-wise BOLD percentage signal change traces locked to the onset of decision phase for all events of interest as follows (Philiastides, Biele, & Heekeren, 2010):

$$BOLD\ \%\ signal\ change_j^t = \left(\frac{BOLD_j^t - BOLD\ baseline_j}{\overline{BOLD}}\right)$$

where *j* and *t* index trial and time point respectively, BOLD baseline is defined as the average BOLD signal over the 4 seconds preceding the event of interest and $\overline{\mathrm{BOLD}}$ is the mean BOLD signal across all time points. To ascertain brain-behaviour correlations we performed 20% bend correlation between the subject-wise, cue-locked mean BOLD signal change averaged over a time window of interest and individual mean behavioural performance. To avoid erroneous inferences (also known as the interaction fallacy)

(Nieuwenhuis, Forstmann, & Wagenmakers, 2011) we compared non-overlapping dependent correlation coefficients using percentile bootstrap (Wilcox, 2016). To predict trial-by-trial behaviour we conducted single trial regression of cluster-wise peak BOLD activity against choice accuracy and reaction time using the following mixed effects models and dropping terms from the random effects in case of model singularity or near-singularity:

$$logit(Accuracy) = 1 + BOLD * cue + (1 + cue|Subject)$$

$$log(RT) = 1 + BOLD + (1 + BOLD|Subject).$$

Results

Behavioural asymmetries as a function of conflicting Pavlovian and instrumental effects.

The well-documented disrupting interference of motivational biases with instrumental responding was borne out by behavioural evidence that on average task performance was greater for congruent (go to win: 96% ± 5%; no-go to avoid losing: 79% ± 14%) than incongruent (go to avoid losing: 77% ± 17%; no-go to win: 61% ± 31%) cues (Figure 1C). Correspondingly, there was a statistically significant effect of congruence (i.e. action by valence interaction) ($\chi^2(1) = 29.2$, $p < .001$) on choice accuracy. Moreover, we found evidence for both valence-dependent and action-dependent behavioural biases as confirmed by significant valence-adjusted main effect of action ($\chi^2(1) = 14.4$, $p < .001$) and action-adjusted main effect of valence ($\chi^2(1) = 12$, $p < .001$) on choice accuracy. In sum, participants performed better in response to: i) go compared to no-go cues, ii) win compared to lose cues and iii) congruent compared to incongruent cues.

Importantly, there was evidence that on average participants successfully learned cue-response contingencies. Indeed, mean performance improved over the course of the task as shown in Figure 2A and as confirmed by the finding of a significant effect of task block on choice accuracy ($\chi^2(1) = 26.6$, $p < .001$).

Cue valence influenced speed of reaction times (Figure 2B). Participants responded faster following win (mean = .436 sec., sem = .08) compared to lose cues (mean = .448 sec., sem = .08). While the valence-adjusted main effect of accuracy was statistically significant ($\chi^2(1) = 9.35$, $p = .002$), the accuracy-adjusted main effect of valence just fell short of statistical significance ($\chi^2(1) = 3.06$, $p = .057$) and the valence by accuracy interaction term was not significant ($\chi^2(1) = 2.78$, $p = .095$). For this analysis we also

included late trials as we reasoned late responses would reflect valence behavioural effect.

Overall, behavioural results revealed the presence of motivational and action-dependent learning biases in accordance with the two-process theory of associative learning which postulates the coexistence of Pavlovian and instrumental systems during learning.

Computational modelling corroborates Pavlovian biasing of instrumental learning.

We found that while the predictive performance of the "dynamically learnt" Pavlovian value model (BIC_i = 5167) was worse than that of the "fixed" Pavlovian value model (BIC_i = 5160) (Figure 1D), its generative performance was better (Figure 1F), especially regarding the no-go to win condition. To objectively evaluate generative performance, we estimated the mean squared error (MSE) between simulated and observed mean choice behaviour for both models. We found that the "dynamically learnt" Pavlovian value model (MSE = 1.18) better reproduced observed behavioural effects than the "fixed" Pavlovian value model (MSE = 1.54) (Figure 1F). To thus arbitrate between these two models, we reasoned that generative performance should be given more weight since it represents an "absolute" rather than "relative" model comparison criterion and the ability to reproduce behavioural effects of interest is a critical aspect of model validation (Palminteri et al., 2017).

Crucially, fitted ($r_{bend}(158) = .98$, $p < .001$) and simulated data ($r_{bend}(158) = .97$, $p < .001$) of the "dynamically learnt" Pavlovian value model provided a good fit to observed choice behaviour (Figure 1E-F). Finally, we were able to successfully recover fitted parameters using our hierarchical type II maximum likelihood fitting routine

(noise: $r_{bend}(38) = .36$, $p = .02$; α: $r_{bend}(38) = .80$, $p < .001$; b: $r_{bend}(38) = .71$, $p < .001$; ρ_{rew}: $r_{bend}(38) = .45$, $p = .003$; ρ_{pun}: $r_{bend}(38) = .60$, $p < .001$; π: $r_{bend}(38) = .72$, $p < .001$; Q^0: $r_{bend}(38) = .86$, $p < .001$).

Our modelling results replicated evidence from previous modelling work denoting Pavlovian biasing of actions (Guitart-Masip, Huys, et al., 2012; Mkrtchian, Aylward, Dayan, Roiser, & Robinson, 2017; Ousdal et al., 2018; Swart et al., 2018; Swart et al., 2017) and are concordant with the observed asymmetries in behavioural performance. Interestingly, unlike previous work (Mkrtchian et al., 2017), albeit using a different experimental setting, we found that accounting for differential effects (i.e. approach versus avoidance) of the Pavlovian expected value on the action weight did not significantly improve model parsimony. It is possible that in previous work the threat posed by unpredictable electric shocks may have increased the variance of the differential Pavlovian effects on go responding.

Instrumental value and action are represented in the action space.

In accordance with previous work (Wunderlich, Rangel, & O'Doherty, 2009), converging results from our fMRI analyses revealed instrumental value and action to be represented in the action space. The task-informed approach revealed a main effect of action (i.e. go>no-go cues) in the bilateral SMC (peak Z score = 5.09; MNI space coordinates = 8,-4,64; $p < .05$ FWE), in the precentral gyri (right: peak Z score = 4.62; MNI space coordinates = 40,-10,56; $p < .05$ FWE; left: peak Z score = 4.67; MNI space coordinates = -24,-4,48; $p < .05$ FWE) and the right postcentral gyrus (peak Z score = 4.2; MNI space coordinates = 50,-18,46; $p < .05$ FWE) (Figure 6A). Conversely, the main effect of inaction (i.e. no-go>go cues) was significantly associated with a distributed group of clusters including the bilateral medial superior frontal

gyrus/paracingulate gyrus (peak Z score = 4.48; MNI space coordinates = -8,16,66; p < .05 FWE), frontal pole (right: peak Z score = 4.42; MNI space coordinates = 54,40,-6; p < .05 FWE; left: peak Z score = 4.46; MNI space coordinates = -24,58,24; p < .05 FWE), right inferior frontal gyrus (IFG) (peak Z score = 4.35; MNI space coordinates = 54,18,24; p < .05 FWE) and left IFG/fronto orbital cortex (peak Z score = 4.41; MNI space coordinates = -44,22,-10; p < .05 FWE) (Figure 6A). A complete list of all significant clusters is provided in Table 1.

Due to the action-specificity of the instrumental value positive contrast (i.e. Q_{go}-$Q_{no\text{-}go}$) we found evidence for action values in a network of areas that largely overlapped with the task-informed main effect of action, including the bilateral SMC (peak Z score = 7.79; MNI space coordinates = -6,2,48; p < .05 FWE), precentral gyri (right: peak Z score = 13.1; MNI space coordinates = 38,-10,58; p < .05 FWE; left : peak Z score = 10; MNI space coordinates = -38,-12,56; p < .05 FWE), left postcentral gyrus (peak Z score = 5.25; MNI space coordinates = -48,-24,50; p < .05 FWE), left lateral occipital gyrus (peak Z score = 4.64; MNI space coordinates = -12,-66,50; p < .05 FWE) and right cerebellum (peak Z score = 10.6; MNI space coordinates = 32,-50,-24; p < .05 FWE) (Figure 3A). Conversely, the instrumental value negative contrast (i.e. $Q_{no\text{-}go}$-Q_{go}) revealed a cluster in the left IFG (peak Z score = 13.1; MNI space coordinates = -56,18,16; p < .05 FWE) and left superior frontal gyrus (peak Z score = 9.15; MNI space coordinates = -6,16,66; p < .05 FWE). This finding is consistent with prior evidence that greater activity in the IFG is elicited by no-go compared to go cues (Guitart-Masip, Huys, et al., 2012) and that the IFG is part of a network operating as a “brake” on motor activity (Chambers et al., 2006) (Figure 6E).

Contrary to previous work (Algermissen, Swart, Scheeringa, Cools, & den Ouden, 2022; Guitart-Masip, Chowdhury, et al., 2012; Guitart-Masip et al., 2011; Guitart-

Masip, Huys, et al., 2012), we did not find evidence of a main effect of action requirement or instrumental value representations in the striatum, even when we constrained our analysis to an anatomical mask of the striatum, therefore making thresholding of statistical maps less stringent.

Instrumental value-related SMC activity correlates with behavioural performance in conflict conditions.

Unsurprisingly, as we retrieved fMRI activity encoding instrumental value from the SMC and plotted mean BOLD time courses as a function of action requirement and valence, we observed a clear action effect in this region (Figure 3B). We then assessed the relationship between SMC activity during mixed trials and accuracy of action selection by correlating mean BOLD signal change averaged over a time window capturing surge response with individual cue-wise performance. Interestingly, we found that upregulating SMC activity in response to incongruent aversive cues (go to avoid losing: $r_{bend}(38) = .47$, $p = .002$) but downregulating it in response to incongruent appetitive cues (no-go to win: $r_{bend}(38) = -.48$, $p = .002$) correlated with greater accuracy of instrumental responding (Figure 3D/E). Conversely, while SMC haemodynamic responses to congruent cues were consistent with action requirements (i.e. increased and diminished signal change for go to win and no-go to avoid losing cues respectively), they did not account for inter-individual differences in choice accuracy (go to win: $r_{bend}(38) = .12$, $p = .47$; no-go to avoid losing: $r_{bend}(38) = -.11$, $p = .48$) (Figure 3C/F). Notably, even after controlling for the potentially confounding effect of preparatory motor activity on the encoding of instrumental value we found BOLD activity in the SMC to significantly covary with positive instrumental value (i.e. Q_{go}-$Q_{no\text{-}go}$) (peak Z score = 7.45; MNI space coordinates = 6,-4,54; $p < .05$ FWE) (Figure

3A) and to be correlated with task performance associated with incongruent (go to avoid losing: $r_{bend}(38) = .45$, $p = .003$; no-go to win: $r_{bend}(38) = -.49$, $p = .001$) but not congruent (go to win: $r_{bend}(38) = .08$, $p = .62$; no-go to avoid losing: $r_{bend}(38) = -.12$, $p = .47$) cues. It is however still possible that BOLD percentage signal change captured motor activity since i) the SMC is recruited in the context of movement (or inhibition of action) (Nachev, Kennard, & Husain, 2008) and ii) the ISI between the cue presentation and target detection phases ranged from 250 to 2500 ms. We thus capitalised on the presence of Pavlovian only trials (see Figure 1B) in our task where there is no target detection phase and thus no overt motor actions are initiated. If SMC activity in our original analysis was driven primarily by movement or action inhibition, we would expect it to dissipate in the Pavlovian only trials. Yet, consistent with our original analysis, we still found BOLD activity in the SMC to be correlated with task performance in incongruent (go to avoid losing: $r_{bend}(38) = .34$, $p = .03$; no-go to win: $r_{bend}(38) = -.38$, $p = .017$) but not congruent (go to win: $r_{bend}(38) = .13$, $p = .43$; no-go to avoid losing: $r_{bend}(38) = -.09$, $p = .59$) trials.

We did not find any significant BOLD-behaviour correlations between fMRI activity recorded in the left IFG and task performance (go to avoid losing: $r_{bend}(38) = .03$, $p = .85$; no-go to win: $r_{bend}(38) = -.013$, $p = .93$; go to win: $r_{bend}(38) = .016$, $p = .92$; no-go to avoid losing: $r_{bend}(38) = -.2$, $p = .21$).

Altogether, these results suggested that differential recruitment of the SMC during conflict conditions facilitates instrumental responding. Correspondingly, it has been documented the SMC plays a key role in successful conflict monitoring and resolution (Aron, Behrens, Smith, Frank, & Poldrack, 2007; Nachev, Rees, Parton, Kennard, & Husain, 2005). Notably, a recent imaging study reported increased haemodynamic

activity in the SMC for incongruent compared to congruent cues (Algermissen et al., 2022).

Instrumental value-related SMC activity predicts trial-by-trial response time and choice accuracy.

Based on the observation that on average participants successfully learned cue-response contingencies, meaning that instrumental predictions effectively guided action selection, we reasoned that haemodynamic activity encoding instrumental value would be predictive of choice behaviour. To this end, we regressed cue-locked peak BOLD activity in the SMC on trial-by-trial response accuracy. As we then assessed the haemodynamic effects of each cue on instrumental responding, we found that while greater BOLD activity significantly predicted greater choice accuracy for the go cues (go to avoid losing: $\beta = 1.12$, $p < .001$; go to win: $\beta = .82$, $p < .001$), the reverse was the case for the no-go cues (no-go to avoid losing: $\beta = -.94$, $p < .001$; no-go to win: $\beta = -1.10$, $p < .001$), even after accounting for motor confounds. Moreover, greater peak BOLD activity in the SMC significantly predicted faster button presses ($\beta = -.06$, $p = .004$). Taken together, these findings consolidate the role of the SMC as a critical region to implement instrumental control of behaviour.

Pavlovian value is encoded in the ventromedial prefrontal cortex.

Surprisingly, in the task-informed fMRI analysis we found a main effect of positive valence (i.e. win > lose cues) in the left precentral gyrus (peak Z score = 4.64; MNI space coordinates = -22,-10,64; $p < .05$ FWE) (Figure 6B) and of negative valence (i.e. lose > win cues) in the left medial caudate/accumbens (peak Z score = 4.32; MNI space coordinates = -6,14,-2; $p < .05$ FWE) (Figure 6C). The latter cluster also significantly

correlated with negative Pavlovian value when we constrained our model-informed analysis to the striatum (peak Z score = 5.13; MNI space coordinates = --8,18,0; $p < .05$ FWE) (Figure 6F).

It is plausible that the relatively greater proportion of go responses in the win trials compared to the lose trials accounted for the observed motor preparation signal associated with positive valence. Furthermore, striatal activity in the negative valence contrast may represent an anticipatory signal of potential losses as previously documented in adults and adolescents (Beck et al., 2009; Bretzke et al., 2022) and is consistent with a similar finding in a recent fMRI study (Algermissen et al., 2022).

Notably, Pavlovian expected value was positively correlated with BOLD activity in the bilateral (but predominantly left) ventromedial prefrontal cortex (vmPFC) (peak Z score = 8.35; MNI space coordinates = -2,58,2; $p < .05$ FWE) (Figure 4A) and negatively correlated with the right dorsal ACC (peak Z score = 5.11; MNI space coordinates = 4,22,38; $p < .05$ FWE) (Figure 4E). Our finding replicates recent evidence that positive and negative valence effects are encoded in the vmPFC and dorsal ACC respectively (Algermissen et al., 2022).

Pavlovian value-related neural activity biases behavioural performance.

By sorting haemodynamic responses in the vmPFC and dorsal ACC as a function of action requirement and valence we uncovered a noticeable valence effect in both areas (Figure 4B/F). As we anticipated, given that the updating of Pavlovian value was tied to instrumental responding during the mixed trials, this valence effect was only marginally modulated by instrumental performance.

To ascertain whether diminished encoding of Pavlovian value in the vmPFC resulted in a lesser disrupting Pavlovian effect on instrumental responding we correlated mean

BOLD percent signal change with individual task performance as a function of congruence. There was a significant BOLD-behaviour correlation only for congruent ($r_{bend}(38) = .34$, $p = .03$) (Figure 4D) but not for incongruent ($r_{bend}(38) = .099$, $p = .54$) cues (Figure 4C). However, when we tested for a significant interaction effect, we found that the difference between correlation coefficients was non-significant (r_{cong}-$r_{incong} = .25$ [-.13 – .64], $p = .22$). Our findings corroborate previous evidence that optimal action selection is achieved by enhanced effort-based overriding of Pavlovian influence on behaviour rather than by attenuated Pavlovian behavioural biases (Cavanagh, Eisenberg, Guitart-Masip, Huys, & Frank, 2013). Notably, peak vmPFC BOLD activity did not predict trial-by-trial choice accuracy (go to avoid losing: $\beta = .07$, $p = .55$; go to win: $\beta = .4$, $p = .08$; no-go to avoid losing: $\beta = -.05$, $p = .65$; no-go to win: $\beta = .13$, $p = .3$). Furthermore, mean BOLD activity in the dorsal ACC was not significantly correlated with performance during congruent ($r_{bend}(38) = .15$, $p = .34$) and incongruent ($r_{bend}(38) = -.09$, $p = .56$) trials (Figure 4G-H).

We then assessed the effect of motivational biases on speed of trial-by-trial responses and found that greater (peak) BOLD activity in the vmPFC and dorsal ACC sped up ($\beta = -.04$, $p = .021$) and slowed down ($\beta = .095$, $p < .001$) motor responses respectively. Consistent with similar recent result (Algermissen et al., 2022), our findings further implicate neural activity in the vmPFC and dorsal ACC in exerting Pavlovian biases on behaviour.

Synergistic Pavlovian and instrumental predictions are represented in the limbic system and medial prefrontal cortex

The final step of our fMRI analysis was to identify neural structures encoding PII effects in both the appetitive and aversive domain. In the task-informed fMRI analysis

674 we uncovered a main effect of congruence (congruent > incongruent cues) in the
675 bilateral vmPFC (peak Z score = 4.84; MNI space coordinates = -6,48,-4; $p < .05$ FWE),
676 right frontal pole (peak Z score = 3.88; MNI space coordinates = 34,48,28; $p < .05$
677 FWE), right orbitofrontal cortex (peak Z score = 4.28; MNI space coordinates = 32,32,-
678 14; $p < .05$ FWE), right posterior cingulate cortex (peak Z score = 4.39; MNI space
679 coordinates = 6,-20,46; $p < .05$ FWE), left parietal opercular cortex (peak Z score =
680 4.12; MNI space coordinates = -54,-28,26; $p < .05$ FWE) and left superior parietal
681 cortex (peak Z score = 4.1; MNI space coordinates = -24,-42,56) (Figure 6D). No
682 clusters associated with a main effect of incongruence (i.e. incongruent > congruent
683 cues) survived multiple comparison correction.
684 Compared to the task-informed fMRI analysis, the model-informed approach
685 uncovered a broader network of (partially overlapping) activations in the brain (Figure
686 4A). When Pavlovian and instrumental value expectations converged, we detected
687 significant BOLD activity in regions pertaining to the limbic system such as the
688 bilateral perigenual ACC/medial PFC (peak Z score = 10.6; MNI space coordinates = -
689 -2,46,20), dorsal ACC (peak Z score = 7.68; MNI space coordinates = -2,36,18; $p < .05$
690 FWE), right orbitofrontal cortex (OFC) (peak Z score = 10.3; MNI space coordinates =
691 52,22,-10; $p < .05$ FWE) and right hippocampus/amygdala complex (peak Z score =
692 8.82; MNI space coordinates = 28,-16,-22; $p < .05$ FWE). A complete list of all
693 significant clusters is provided in Table 2. It is important to note again that activity in
694 these brain areas covaried with congruent Pavlovian and instrumental predictions in
695 both the aversive and appetitive domain.
696 Consistent with the task-informed analysis we did not find any significant clusters to
697 be negatively correlated with the model-informed interaction regressor. One possible
698 explanation is that competition interaction effects are being absorbed by instrumental

and Pavlovian main effects and the neural mechanisms of competition are being implemented in the premotor areas subserving instrumental control.

PII synergistic neural effects scale with behavioural performance in cooperation conditions.

We reasoned that BOLD activity reflecting Pavlovian and instrumental synergistic effects should scale with observed task performance during congruent trials when both Pavlovian and instrumental predictions prescribe the same motor responses.

Based on the extensive literature on the key role of the amygdala in the context of PIT (Cartoni, Balleine, & Baldassarre, 2016) we first investigated this region and found that its haemodynamic activity was significantly correlated with performance in the congruent ($r_{bend}(38) = .38$, $p = .015$) but not in the incongruent ($r_{bend}(38) = -.05$, $p = .77$) trials (Figure 4G/E). As we compared correlation coefficients, we found a significant difference (r_{cong}-r_{incong} = .41 [.01 – .77], p =.041) thus confirming a significant interaction and pointing to a specific effect of amygdalar activity on task performance in response to congruent cues. As we further dissected activity in this region by cue type, we found evidence of temporal disambiguation of the PII neural effects along the valence axis with an earlier peak for the aversive congruent cues and a later peak for the appetitive congruent cues (Figure 4C).

We also investigated BOLD responses in other brain regions pertaining to the limbic system (including the PCC, dorsal ACC and right OFC,) and medial PFC. In the PCC we observed an analogous activity pattern to the right amygdala/hippocampus complex (Figure 4C). Similarly, individual task performance scaled with mean PCC haemodynamic responses to congruent ($r_{bend}(38) = .35$, $p = .025$) but not incongruent ($r_{bend}(38) = .04$, $p = .79$) cues (Figure 4E/G) although in this case interaction was

724 statistically non-significant (r_{cong}-r_{incong} = .31 [-.08 – .67], p =.12). A less prominent
725 two-peak activity pattern was still discernible in the dorsal ACC (Figure 7A) although
726 it did not significantly correlate with behaviour (congruent cues: $r_{bend}(38)$ = .087, p =
727 .59; incongruent cues: $r_{bend}(38)$ = .11, p = .51). We found positive BOLD responses in
728 the medial PFC (Figure 7C) and in the right OFC (Figure 7E) to be valence dependent
729 as they were predominantly elicited by appetitive and aversive congruent cues
730 respectively. However, we did not detect any significant BOLD-behaviour correlations
731 in the medial PFC (congruent cues: $r_{bend}(38)$ = .23, p = .16; incongruent cues: $r_{bend}(38)$
732 = .18, p = .26) nor the right OFC (congruent cues: $r_{bend}(38)$ = .17, p = .29; incongruent
733 cues: $r_{bend}(38)$ = -.12, p = .44).
734

Discussion

Dissecting the functional neuroanatomy of associative learning can afford invaluable insights into the neural mechanisms underlying adaptive and maladaptive decision-making. Accordingly, the neural underpinnings of the two-process learning theory have been extensively researched predominantly in animals but more recently in humans. Computational methods have made a meaningful contribution to advancing our mechanistic understanding of how Pavlovian and instrumental valuation systems shape learning and oversee action control (Dorfman & Gershman, 2019; Guitart-Masip, Huys, et al., 2012). Furthermore, combining computational approaches with neuroimaging techniques has yielded a powerful analysis tool to probe the neural pathways mediating associative learning and conditioned behaviour (O'Doherty, Hampton, & Kim, 2007). Nonetheless, the experimental paradigms so far employed in this line of research such as the PIT and orthogonalized go/no-go tasks have failed to completely unlock the explanatory potential of modelling work. In this study we devised a variation of the popular orthogonalized go/no-go task that permitted updating Pavlovian value in the absence of any instrumental learning. This simple refinement of the task design enabled us to better decorrelate the temporal evolution of Pavlovian and instrumental predictions associated with a given cue and model trial-by-trial updating of Pavlovian and instrumental contingencies. Moreover, we were able to describe neural interaction effects dynamically rather than in the absence of any new learning as it is the case in the conventional PIT paradigms. Using a model-informed imaging analysis approach, we identified distinct fMRI activation clusters encoding instrumental and Pavlovian predictions of future payoff in the SMC and vmPFC respectively. Crucially, activity in the SMC could not be simply accounted for by impending motor responses and replicated a prior finding from our lab that

760 value-based decisions requiring an overt response are encoded in the SMC (Pisauro,
761 Fouragnan, Retzler, & Philiastides, 2017). In humans, a previous functional
762 neuroimaging study also found evidence for action value signals in the SMC
763 (Wunderlich et al., 2009). The action-specificity of our instrumental regressor
764 accounted for the observed activity in premotor areas, which are involved in
765 representing and planning movement. It has been previously suggested action values
766 are encoded in the action space and embedded in premotor processing of action
767 selection (Wunderlich, Rangel, & O'Doherty, 2010). Importantly, we have shown that
768 modulation of mean SMC haemodynamic responses to incongruent cues was linked to
769 successful instrumental responding. Furthermore, trial-by-trial oscillations in this
770 activity were predictive of response time and accuracy, therefore casting this region as
771 a focal neural hub of instrumental control.
772 Anatomically, the SMC is thought to be part of a frontal-subcortical network (Aron et
773 al., 2007; Nachev et al., 2008), which is involved in cognitive processes such as
774 conflict monitoring, detection and resolution (usually referred to as cognitive control)
775 (Aron et al., 2007). The SMC has a 'hyperdirect' connection to the subthalamic
776 nucleus (Aron et al., 2007; Tanji, Kurata, & Okano, 1985), which it recruits to inhibit
777 or slow down prepotent motor responses (Aron & Poldrack, 2006; Frank, Samanta,
778 Moustafa, & Sherman, 2007) via suppression of thalamocortical activity (Mink,
779 1996). In addition to withholding prepotent motor responses, another important
780 function of the SMC is altering movement plans and switching between actions or
781 rules (Crone, Wendelken, Donohue, & Bunge, 2006; Rushworth, Hadland, Paus, &
782 Sipila, 2002). The SMC thus seems ideally suited to enabling optimal action control
783 in high conflict situations. Unlike in a previous fMRI study where 'learners' showed

784 greater recruitment of bilateral IFG in response to no-go cues (Guitart-Masip, Huys,
785 et al., 2012), we did not find any evidence linking IFG to improved performance.
786 Altogether, our data seem to suggest that Pavlovian response biases are overcome
787 through increased goal-directed cognitive control rather than attenuated Pavlovian
788 value signals. We did in fact not find evidence that reduced neural encoding of
789 Pavlovian value was correlated with improved performance as a function of
790 congruence. Relatedly, previous research work employing electroencephalography
791 (EEG) demonstrated that frontal midline theta activity suppressed Pavlovian response
792 biases and improved choice accuracy (Cavanagh et al., 2013; Csifcsak, Melsaeter, &
793 Mittner, 2020; Swart et al., 2018). Theta oscillations have been linked to the detection
794 of response conflict (Cohen & Cavanagh, 2011; van Driel, Swart, Egner,
795 Ridderinkhof, & Cohen, 2015) and their source has been traced back to the SMC
796 (Cohen & Ridderinkhof, 2013). Furthermore, stronger EEG midfrontal-motor
797 connectivity has been shown to be associated with the reduction of Pavlovian
798 interference, suggesting that the SMC may modulate motor response threshold in
799 conflict conditions (Philiastides, Biele, Vavatzanidis, Kazzer, & Heekeren, 2010;
800 Swart et al., 2018).
801 The function of the vmPFC in human value-based decision making has so far been
802 characterised as heterogeneous. Indeed, the vmPFC has been extensively implicated
803 in signalling value (Bartra, McGuire, & Kable, 2013; Fouragnan, Retzler, Mullinger,
804 & Philiastides, 2015; Philiastides, Biele, & Heekeren, 2010) and, more specifically,
805 has been found to be involved in the representation of action-specific value
806 (FitzGerald, Friston, & Dolan, 2012), chosen stimulus value (independent of
807 stimulus-action pairing) (Wunderlich et al., 2010), probability of chosen action (Daw,
808 O'Doherty, Dayan, Seymour, & Dolan, 2006), the prior belief that a choice is correct

(Hampton, Bossaerts, & O'Doherty, 2006), reward expectations (Blair et al., 2006), expected reward value (H. Kim, Shimojo, & O'Doherty, 2006), chosen (rostral) and unchosen (caudal) action expected value (Morris, Dezfouli, Griffiths, & Balleine, 2014). To the best of our knowledge while there are no previous imaging studies that have documented Pavlovian value signals in the vmPFC or dorsal ACC, recent research work has reported positive and negative valence effects in these regions consistently with our findings (Algermissen et al., 2022). Also, outcome-locked vmPFC activity has been shown to scale with greater Pavlovian influence as a function of diminished environmental controllability (Gershman, Guitart-Masip, & Cavanagh, 2021). Notably, previous human imaging studies employed instrumental learning paradigms whereby Pavlovian predictions would have been updated based on instrumental performance and therefore resulting in instrumental predictions capturing neural activity linked to Pavlovian valuation. While in our task we were not able to completely decorrelate Pavlovian and instrumental predictions due to the presence of mixed trials, we still found a noticeable valence effect (only marginally modulated by instrumental performance) in the vmPFC and dorsal ACC. The critical role of these regions in exerting motivational biases on behaviour was corroborated by the findings that both vmPFC and dorsal ACC peak activity predicted trial-by-trial speeding up and slowing down of button presses. Furthermore, the observation that vmPFC activity did not correlate with performance as a function of congruence nor did it predict trial-wise choice accuracy suggested that the vmPFC is not directly involved in guiding action selection and resolving conflict.

Using classical or operant conditioning paradigms, previous human fMRI studies detected action (FitzGerald et al., 2012) and Pavlovian (Gottfried, O'Doherty, & Dolan, 2002; O'Doherty, Deichmann, Critchley, & Dolan, 2002) expected value

834 signals in the striatum. Furthermore, previous work using the conventional version of
835 the orthogonalized go/no-go task reported striatal haemodynamic activity during cue
836 presentation predominantly represented action rather than valence effects (Guitart-
837 Masip et al., 2011; Guitart-Masip, Huys, et al., 2012). It is worth noting that in
838 (Guitart-Masip et al., 2011) participants were fully aware of cue-action couplings,
839 which may have lessened motivational biases on action control. Nonetheless, the
840 modulatory effect of valence on action representations in the striatum was still
841 evident, albeit statistically non-significant (Guitart-Masip et al., 2011). In line with a
842 previous report (Algermissen et al., 2022), we found that BOLD activity in a cluster
843 in the left medial caudate was significantly correlated with negative valence (whole-
844 brain analysis) and negative Pavlovian value (ROI analysis). Remarkably, we did not
845 detect any striatal activity covarying with trial-by-trial instrumental predictions nor
846 positive Pavlovian value nor PII effects. It is possible that the presence of Pavlovian
847 only trials in our task may have enhanced valence signals and diluted action
848 representations in the striatum.
849 There exists a rich animal literature implicating the amygdala in different forms of
850 PIT including conditioned suppression (i.e. an aversive Pavlovian conditioned
851 stimulus decreases vigour of appetitive instrumental approach responding) and
852 conditioned facilitation (i.e. an aversive Pavlovian conditioned stimulus increases
853 vigour of aversive instrumental avoidance responding) (Campese, Gonzaga,
854 Moscarello, & LeDoux, 2015; Campese et al., 2017; I. T. Kim et al., 2022). Using
855 congruent appetitive PIT paradigms previous neuroimaging studies have found that
856 BOLD activity in the nucleus accumbens (Mendelsohn, Pine, & Schiller, 2014; Talmi,
857 Seymour, Dayan, & Dolan, 2008), amygdala (Mendelsohn et al., 2014; Prevost,
858 Liljeholm, Tyszka, & O'Doherty, 2012; Talmi et al., 2008) and striatum (Bray,

Rangel, Shimojo, Balleine, & O'Doherty, 2008; Prevost et al., 2012) is correlated with appetitive PIT effects. Furthermore, the amygdala and the nucleus accumbens have also been linked to aversive PIT and were activated in the context of behavioural inhibition elicited by aversive Pavlovian cues (Geurts et al., 2013). Our finding that the hippocampus/amygdala complex encodes congruent PII effects both in the aversive and appetitive domains thus reconciles previous neuroimaging reports and is supported by the observation of a significant correlation with performance during congruent trials. Intriguingly, the temporal dynamics of this effect are characterised by an earlier peak for aversive cues and a later peak for appetitive cues. Notably a similar activity pattern was also observed in the PCC. Converging empirical evidence has established that the amygdala acquires information about the overall motivational value and salience of environmental stimuli and regulates motivated behaviour (Morrison & Salzman, 2010). The amygdala does in fact encode information about both positively and negatively valenced stimuli (Ball et al., 2009), which it appraises in a context-depended fashion, subsequently giving rises to decision biases (De Martino et al., 2006).

Our task did not permit identification of biases in instrumental learning. Swart et al. reported asymmetries in instrumental learning alongside already known Pavlovian mediated response biases and showed that, compared to no-go responses, go responses were easier to learn and unlearn in the face of reward and punishment respectively (Swart et al., 2017). Moreover, our task did not account for action-specificity of Pavlovian influence on instrumental control across withdrawal and approach contexts (Huys et al., 2011). Finally, to conclusively rule out the biasing effect of motor confounds when assessing correlations between brain activity and performance, our task could be modified to ensure a comparable number of correct

(or incorrect) trials across subjects. Future imaging work leveraging the greater spatial resolution of ultra-high field fMRI may shed light on the human neural correlates of general and specific PIT effects which have been well-characterised in animals using lesion manipulations (Corbit & Balleine, 2011).

In conclusion, in this study we have devised a novel version of a popular orthogonalized go/no-go task to better disentangle Pavlovian and instrumental neural representations and have expanded existing knowledge on the functional neuroanatomy of Pavlovian and instrumental processes underlying decision-making. We have shown that while the SMC encodes instrumental value and facilitates optimal instrumental responses detecting and overriding Pavlovian conflict, the vmPFC underpins Pavlovian valuation and gives rise to motivational biasing of behaviour. Moreover, we have elucidated the role of the amygdala/hyppocampus (and PCC) in implementing PII synergistic neural effects at different timescales across the negative and positive valence domain.

Data Availability Statement

Data and code will be made available upon reasonable request.

Author Contribution

Filippo Queirazza: Conceptualization; Data curation; Funding acquisition; Formal analysis; Investigation; Methodology; Project administration; Software; Visualization; Writing—Original draft; Writing—Review & editing. Rajeev Krishnadas: Writing—Review & editing. J. Douglas Steele: Writing—Review & editing. Jonathan Cavanagh: Writing—Review & editing. Marios Philiastides: Conceptualization; Formal analysis; Methodology; Software; Visualization; Writing—Review & editing.

Acknowledgements

This work was supported by supported by an MRC- and MRF-funded clinical research training fellowship (PsySTAR) awarded to F.Q..

References

Algermissen, J., Swart, J. C., Scheeringa, R., Cools, R., & den Ouden, H. E. M. (2022). Striatal BOLD and Midfrontal Theta Power Express Motivation for Action. *Cereb Cortex, 32*(14), 2924-2942.

Aron, A. R., Behrens, T. E., Smith, S., Frank, M. J., & Poldrack, R. A. (2007). Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. *J Neurosci, 27*(14), 3743-3752.

Aron, A. R., & Poldrack, R. A. (2006). Cortical and subcortical contributions to Stop signal response inhibition: role of the subthalamic nucleus. *J Neurosci, 26*(9), 2424-2433.

Ball, T., Derix, J., Wentlandt, J., Wieckhorst, B., Speck, O., Schulze-Bonhage, A., et al. (2009). Anatomical specificity of functional amygdala imaging of responses to stimuli with positive and negative emotional valence. *J Neurosci Methods, 180*(1), 57-70.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J Mem Lang, 68*(3).

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage, 76*, 412-427.

Beck, A., Schlagenhauf, F., Wustenberg, T., Hein, J., Kienast, T., Kahnt, T., et al. (2009). Ventral striatal activation during reward anticipation correlates with impulsivity in alcoholics. *Biol Psychiatry, 66*(8), 734-742.

Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2003). General multilevel linear modeling for group analysis in FMRI. *Neuroimage, 20*(2), 1052-1063.

938 Blair, K., Marsh, A. A., Morton, J., Vythilingam, M., Jones, M., Mondillo, K., et al.
939 (2006). Choosing the lesser of two evils, the better of two goods: specifying
940 the roles of ventromedial prefrontal cortex and dorsal anterior cingulate in
941 object choice. *J Neurosci, 26*(44), 11379-11386.
942 Bray, S., Rangel, A., Shimojo, S., Balleine, B., & O'Doherty, J. P. (2008). The neural
943 mechanisms underlying the influence of pavlovian cues on human decision
944 making. *J Neurosci, 28*(22), 5861-5866.
945 Bretzke, M., Vetter, N. C., Kohls, G., Wahl, H., Roessner, V., Plichta, M. M., et al.
946 (2022). Is loss avoidance differentially rewarding in adolescents versus adults?
947 Differences in ventral striatum and anterior insula activation during the
948 anticipation of potential monetary losses. *Cogn Neurosci*, 1-14.
949 Brown, P. L., & Jenkins, H. M. (1968). Auto-shaping of the pigeon's key-peck. *J Exp*
950 *Anal Behav, 11*(1), 1-8.
951 Campese, V. D., Gonzaga, R., Moscarello, J. M., & LeDoux, J. E. (2015). Modulation
952 of instrumental responding by a conditioned threat stimulus requires lateral
953 and central amygdala. *Front Behav Neurosci, 9*, 293.
954 Campese, V. D., Soroeta, J. M., Vazey, E. M., Aston-Jones, G., LeDoux, J. E., &
955 Sears, R. M. (2017). Noradrenergic Regulation of Central Amygdala in
956 Aversive Pavlovian-to-Instrumental Transfer. *eNeuro, 4*(5).
957 Cartoni, E., Balleine, B., & Baldassarre, G. (2016). Appetitive Pavlovian-instrumental
958 Transfer: A review. *Neurosci Biobehav Rev, 71*, 829-848.
959 Cartoni, E., Puglisi-Allegra, S., & Baldassarre, G. (2013). The three principles of
960 action: a Pavlovian-instrumental transfer hypothesis. *Front Behav Neurosci, 7*,
961 153.

962 Cavanagh, J. F., Eisenberg, I., Guitart-Masip, M., Huys, Q., & Frank, M. J. (2013).
963 Frontal theta overrides pavlovian learning biases. *J Neurosci, 33*(19), 8541-
964 8548.
965 Chambers, C. D., Bellgrove, M. A., Stokes, M. G., Henderson, T. R., Garavan, H.,
966 Robertson, I. H., et al. (2006). Executive "brake failure" following
967 deactivation of human frontal lobe. *J Cogn Neurosci, 18*(3), 444-455.
968 Cohen, M. X., & Cavanagh, J. F. (2011). Single-trial regression elucidates the role of
969 prefrontal theta oscillations in response conflict. *Front Psychol, 2*, 30.
970 Cohen, M. X., & Ridderinkhof, K. R. (2013). EEG source reconstruction reveals
971 frontal-parietal dynamics of spatial conflict processing. *PLoS One, 8*(2),
972 e57293.
973 Corbit, L. H., & Balleine, B. W. (2011). The general and outcome-specific forms of
974 Pavlovian-instrumental transfer are differentially mediated by the nucleus
975 accumbens core and shell. *J Neurosci, 31*(33), 11786-11794.
976 Crockett, M. J., Clark, L., & Robbins, T. W. (2009). Reconciling the role of serotonin
977 in behavioral inhibition and aversion: acute tryptophan depletion abolishes
978 punishment-induced inhibition in humans. *J Neurosci, 29*(38), 11993-11999.
979 Crone, E. A., Wendelken, C., Donohue, S. E., & Bunge, S. A. (2006). Neural
980 evidence for dissociable components of task-switching. *Cereb Cortex, 16*(4),
981 475-486.
982 Csifcsak, G., Melsaeter, E., & Mittner, M. (2020). Intermittent Absence of Control
983 during Reinforcement Learning Interferes with Pavlovian Bias in Action
984 Selection. *J Cogn Neurosci, 32*(4), 646-663.
985 Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical
986 substrates for exploratory decisions in humans. *Nature, 441*(7095), 876-879.

Dayan, P., & Huys, Q. J. (2008). Serotonin, inhibition, and negative mood. *PLoS Comput Biol, 4*(2), e4.

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science, 313*(5787), 684-687.

Dorfman, H. M., & Gershman, S. J. (2019). Controllability governs the balance between Pavlovian and instrumental action selection. *Nat Commun, 10*(1), 5826.

FitzGerald, T. H., Friston, K. J., & Dolan, R. J. (2012). Action-specific value signals in reward-related regions of the human brain. *J Neurosci, 32*(46), 16417-16423a.

Fouragnan, E., Retzler, C., Mullinger, K., & Philiastides, M. G. (2015). Two spatiotemporally distinct value systems shape reward-based learning in the human brain. *Nat Commun, 6*, 8107.

Frank, M. J., Samanta, J., Moustafa, A. A., & Sherman, S. J. (2007). Hold your horses: impulsivity, deep brain stimulation, and medication in parkinsonism. *Science, 318*(5854), 1309-1312.

Garbusow, M., Ebrahimi, C., Riemerschmid, C., Daldrup, L., Rothkirch, M., Chen, K., et al. (2022). Pavlovian-to-Instrumental Transfer across Mental Disorders: A Review. *Neuropsychobiology*, 1-20.

Gershman, S. J., Guitart-Masip, M., & Cavanagh, J. F. (2021). Neural signatures of arbitration between Pavlovian and instrumental action selection. *PLoS Comput Biol, 17*(2), e1008553.

1010 Geurts, D. E., Huys, Q. J., den Ouden, H. E., & Cools, R. (2013). Aversive Pavlovian
1011 control of instrumental behavior in humans. *J Cogn Neurosci, 25*(9), 1428-
1012 1441.
1013 Gottfried, J. A., O'Doherty, J., & Dolan, R. J. (2002). Appetitive and aversive
1014 olfactory learning in humans studied using event-related functional magnetic
1015 resonance imaging. *J Neurosci, 22*(24), 10829-10837.
1016 Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using
1017 boundary-based registration. *Neuroimage, 48*(1), 63-72.
1018 Guitart-Masip, M., Chowdhury, R., Sharot, T., Dayan, P., Duzel, E., & Dolan, R. J.
1019 (2012). Action controls dopaminergic enhancement of reward representations.
1020 *Proc Natl Acad Sci U S A, 109*(19), 7511-7516.
1021 Guitart-Masip, M., Fuentemilla, L., Bach, D. R., Huys, Q. J., Dayan, P., Dolan, R. J.,
1022 et al. (2011). Action dominates valence in anticipatory representations in the
1023 human striatum and dopaminergic midbrain. *J Neurosci, 31*(21), 7867-7875.
1024 Guitart-Masip, M., Huys, Q. J., Fuentemilla, L., Dayan, P., Duzel, E., & Dolan, R. J.
1025 (2012). Go and no-go learning in reward and punishment: interactions
1026 between affect and effect. *Neuroimage, 62*(1), 154-166.
1027 Hampton, A. N., Bossaerts, P., & O'Doherty, J. P. (2006). The role of the
1028 ventromedial prefrontal cortex in abstract state-based inference during
1029 decision making in humans. *J Neurosci, 26*(32), 8360-8367.
1030 Huys, Q. J., Cools, R., Golzer, M., Friedel, E., Heinz, A., Dolan, R. J., et al. (2011).
1031 Disentangling the roles of approach, activation and valence in instrumental
1032 and pavlovian responding. *PLoS Comput Biol, 7*(4), e1002028.

Huys, Q. J., Golzer, M., Friedel, E., Heinz, A., Cools, R., Dayan, P., et al. (2016). The specificity of Pavlovian regulation is associated with recovery from depression. *Psychol Med, 46*(5), 1027-1035.

Jenkinson, M. (2003). Fast, automated, N-dimensional phase-unwrapping algorithm. *Magn Reson Med, 49*(1), 193-197.

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage, 17*(2), 825-841.

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Med Image Anal, 5*(2), 143-156.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1990). Experimental Tests of the Endowment Effect and the Coase Theorem. *Journal of Political Economy, 98*(6), 1325-1348.

Kim, H., Shimojo, S., & O'Doherty, J. P. (2006). Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *PLoS Biol, 4*(8), e233.

Kim, I. T., Farb, C., Hou, M., Prasad, S., Talley, E., Cook, S., et al. (2022). General and Specific Aversive Modulation of Active Avoidance Require Central Amygdala. *Front Behav Neurosci, 16*, 879168.

Mackintosh, N. J. (1983). *Conditioning and associative learning*. Oxford: Clarendon Press.

Mendelsohn, A., Pine, A., & Schiller, D. (2014). Between thoughts and actions: motivationally salient cues invigorate mental action in the human brain. *Neuron, 81*(1), 207-217.

1057 Mink, J. W. (1996). The basal ganglia: focused selection and inhibition of competing
1058 motor programs. *Prog Neurobiol, 50*(4), 381-425.
1059 Mkrtchian, A., Aylward, J., Dayan, P., Roiser, J. P., & Robinson, O. J. (2017).
1060 Modeling Avoidance in Mood and Anxiety Disorders Using Reinforcement
1061 Learning. *Biol Psychiatry, 82*(7), 532-539.
1062 Morris, R. W., Dezfouli, A., Griffiths, K. R., & Balleine, B. W. (2014). Action-value
1063 comparisons in the dorsolateral prefrontal cortex control choice between goal-
1064 directed actions. *Nat Commun, 5*, 4390.
1065 Morrison, S. E., & Salzman, C. D. (2010). Re-valuing the amygdala. *Curr Opin*
1066 *Neurobiol, 20*(2), 221-230.
1067 Nachev, P., Kennard, C., & Husain, M. (2008). Functional role of the supplementary
1068 and pre-supplementary motor areas. *Nat Rev Neurosci, 9*(11), 856-869.
1069 Nachev, P., Rees, G., Parton, A., Kennard, C., & Husain, M. (2005). Volition and
1070 conflict in human medial frontal cortex. *Curr Biol, 15*(2), 122-128.
1071 Nieuwenhuis, S., Forstmann, B. U., & Wagenmakers, E. J. (2011). Erroneous
1072 analyses of interactions in neuroscience: a problem of significance. *Nat*
1073 *Neurosci, 14*(9), 1105-1107.
1074 O'Doherty, J. P., Deichmann, R., Critchley, H. D., & Dolan, R. J. (2002). Neural
1075 responses during anticipation of a primary taste reward. *Neuron, 33*(5), 815-
1076 826.
1077 O'Doherty, J. P., Hampton, A., & Kim, H. (2007). Model-based fMRI and its
1078 application to reward learning and decision making. *Ann N Y Acad Sci, 1104*,
1079 35-53.

1080 Ousdal, O. T., Huys, Q. J., Milde, A. M., Craven, A. R., Ersland, L., Endestad, T., et
1081 al. (2018). The impact of traumatic stress on Pavlovian biases. *Psychol Med,*
1082 *48*(2), 327-336.
1083 Palminteri, S., Wyart, V., & Koechlin, E. (2017). The Importance of Falsification in
1084 Computational Cognitive Modeling. *Trends Cogn Sci, 21*(6), 425-433.
1085 Philiastides, M. G., Biele, G., & Heekeren, H. R. (2010). A mechanistic account of
1086 value computation in the human brain. *Proc Natl Acad Sci U S A, 107*(20),
1087 9430-9435.
1088 Philiastides, M. G., Biele, G., Vavatzanidis, N., Kazzer, P., & Heekeren, H. R. (2010).
1089 Temporal dynamics of prediction error processing during reward-based
1090 decision making. *Neuroimage, 53*(1), 221-232.
1091 Pisauro, M. A., Fouragnan, E., Retzler, C., & Philiastides, M. G. (2017). Neural
1092 correlates of evidence accumulation during value-based decisions revealed via
1093 simultaneous EEG-fMRI. *Nat Commun, 8*, 15808.
1094 Prevost, C., Liljeholm, M., Tyszka, J. M., & O'Doherty, J. P. (2012). Neural
1095 correlates of specific and general Pavlovian-to-Instrumental Transfer within
1096 human amygdalar subregions: a high-resolution fMRI study. *J Neurosci,*
1097 *32*(24), 8383-8390.
1098 Rescorla, R. A., & Solomon, R. L. (1967). Two-process learning theory:
1099 Relationships between Pavlovian conditioning and instrumental learning.
1100 *Psychol Rev, 74*(3), 151-182.
1101 Robinson, T. E., & Berridge, K. C. (2003). Addiction. *Annu Rev Psychol, 54*, 25-53.
1102 Rushworth, M. F., Hadland, K. A., Paus, T., & Sipila, P. K. (2002). Role of the
1103 human medial frontal cortex in task switching: a combined fMRI and TMS
1104 study. *J Neurophysiol, 87*(5), 2577-2592.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage, 23 Suppl 1*, S208-219.

Swart, J. C., Frank, M. J., Maatta, J. I., Jensen, O., Cools, R., & den Ouden, H. E. M. (2018). Frontal network dynamics reflect neurocomputational mechanisms for reducing maladaptive biases in motivated action. *PLoS Biol, 16*(10), e2005979.

Swart, J. C., Frobose, M. I., Cook, J. L., Geurts, D. E., Frank, M. J., Cools, R., et al. (2017). Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in)action. *Elife, 6*.

Talmi, D., Seymour, B., Dayan, P., & Dolan, R. J. (2008). Human pavlovian-instrumental transfer. *J Neurosci, 28*(2), 360-368.

Tanji, J., Kurata, K., & Okano, K. (1985). The effect of cooling of the supplementary motor cortex and adjacent cortical areas. *Exp Brain Res, 60*(2), 423-426.

van Driel, J., Swart, J. C., Egner, T., Ridderinkhof, K. R., & Cohen, M. X. (2015). (No) time for control: Frontal theta dynamics reveal the cost of temporally guided conflict anticipation. *Cogn Affect Behav Neurosci, 15*(4), 787-807.

Weiskopf, N., Hutton, C., Josephs, O., & Deichmann, R. (2006). Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: a whole-brain analysis at 3 T and 1.5 T. *Neuroimage, 33*(2), 493-504.

Wilcox, R. R. (2016). Comparing dependent robust correlations. *Br J Math Stat Psychol, 69*(3), 215-224.

Williams, D. R., & Williams, H. (1969). Auto-maintenance in the pigeon: sustained pecking despite contingent non-reinforcement. *J Exp Anal Behav, 12*(4), 511-520.

Wunderlich, K., Rangel, A., & O'Doherty, J. P. (2009). Neural computations underlying action-based decision making in the human brain. *Proc Natl Acad Sci U S A, 106*(40), 17199-17204.

Wunderlich, K., Rangel, A., & O'Doherty, J. P. (2010). Economic choices can be made using only stimulus values. *Proc Natl Acad Sci U S A, 107*(34), 15005-15010.

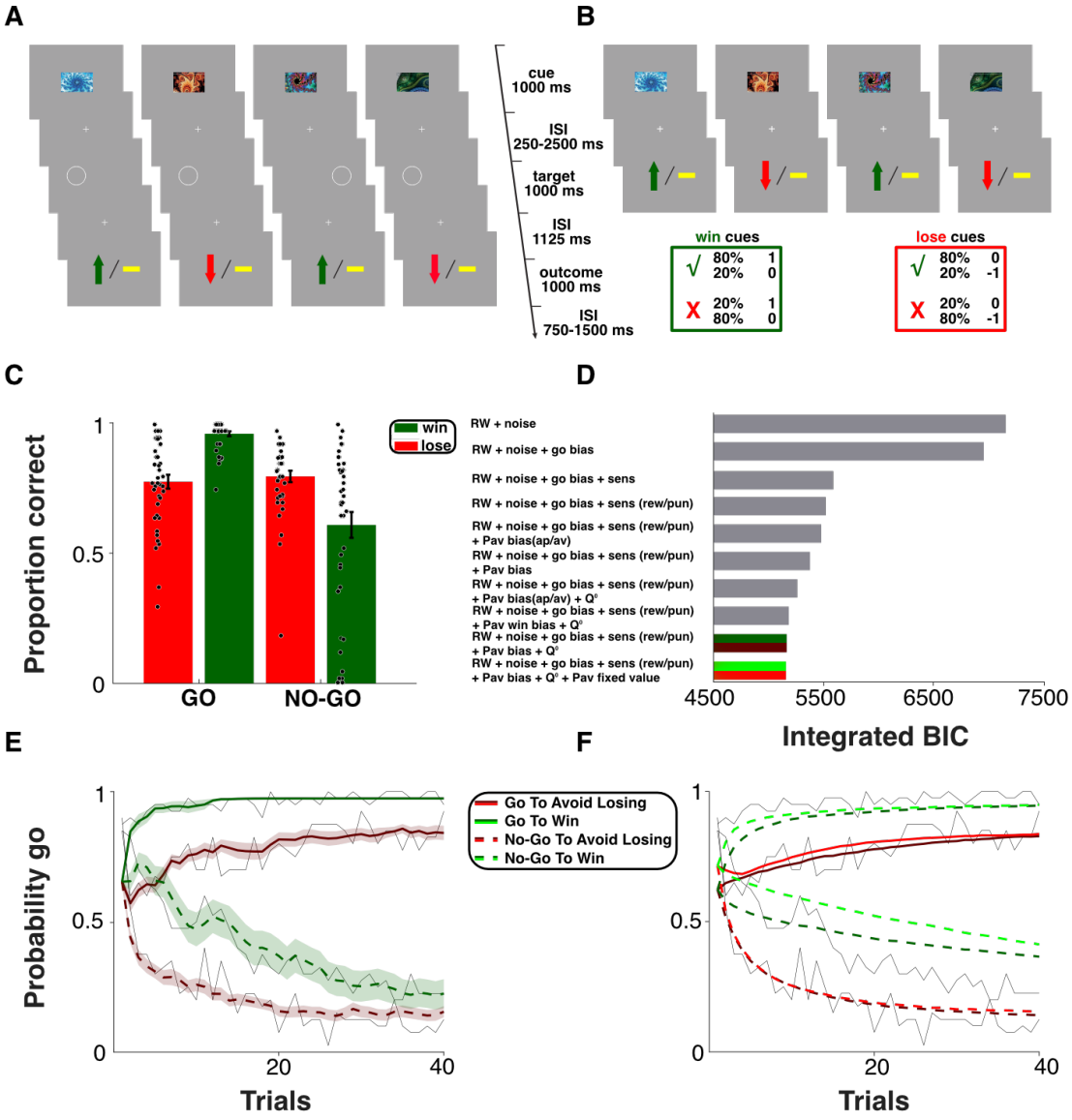

Figure 1. **Task design and behaviour**. **A-B)** Mixed (A) and Pavlovian only (B) trials. Response-outcome contingencies for win and lose cues are shown in the green and red box respectively. **C)** Proportion of correct choices as a function of action requirement and valence. Error bars denote standard error of the mean (SEM). **D)** Predictive performance of candidate models assessed using the integrated Bayesian Information Criterion (BIC). Smaller values indicate better predictive performance. **E)** Fitted behavioural data from the "dynamically learnt" Pavlovian value model. **F)** Simulated behavioural data (light colours for the "dynamically learnt" Pavlovian value model and dark colours for the "fixed" Pavlovian value model). Grey lines represent observed

1148 mean choice behaviour (i.e. trial-wise probability of choosing go action). Coloured
1149 shadings represent SEM.
1150

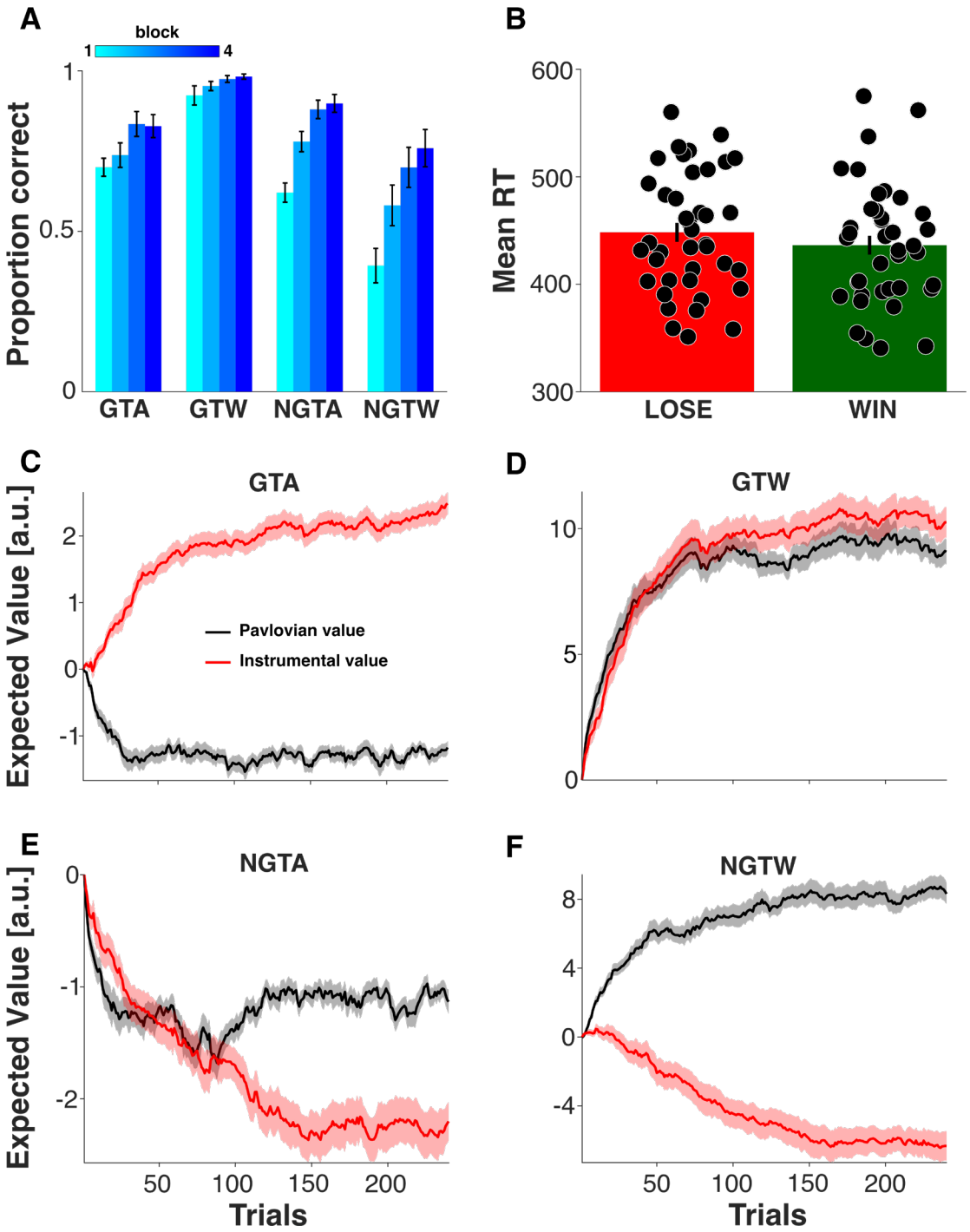

Figure 2. A) Proportion of correct choices as a function of action requirement, valence and task block (n=4). GTA: go to avoid losing. GTW: go to win. NGTA: no-go to avoid losing. NGTW: no-go to win. Error bars denote SEM. **B)** Mean of median reaction times (RT) in milliseconds for lose (red bar) and win (green bar) conditions showing valence effect on speed of responding. Error bars denote SEM. Black dots represent individual subjects. **C-F)** Trial-by-trial trace plots of cue-wise fitted

instrumental and Pavlovian expected value. Solid colour-coded lines denote mean expected value and coloured shadings represent SEM. At the population-level the Pavlovian and instrumental regressors were only partially correlated (Person's rho = 0.36).

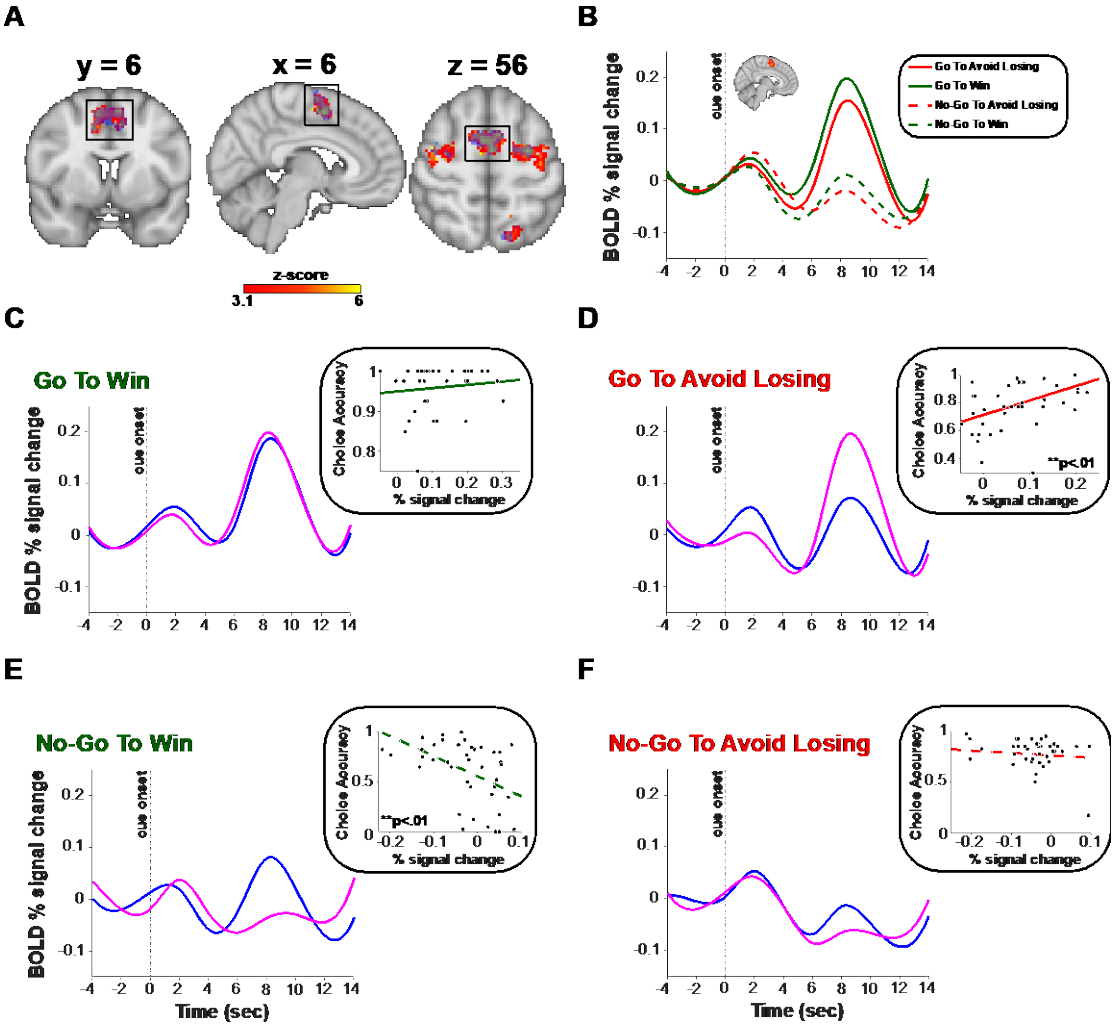

Figure 3. Neural correlates of instrumental value. A) fMRI clusters for instrumental value with (overlaid blue-light blue shading) and without (red-yellow cluster) accounting for motor confounds ($p < .05$ FWE). SMC is highlighted by black box. MNI coordinates are shown. **B)** BOLD traces extracted from the SMC and locked to cue onset as a function of action requirement and valence. Haemodynamic activity shows noticeable action effect (go > no-go cues). **C-F)** The insets show the scatterplots of cue-wise SMC BOLD activity (averaged over the transparent grey time window shown in the BOLD traces) as a function of cue-wise choice accuracy with colour-coded lines denoting 20% bend correlation fit. Black dots represent individual subjects. Notably, mean SMC haemodynamic responses to incongruent cues are significantly associated with task performance (even when accounting for motor confounds). BOLD traces

represent subjects in the top (purple solid line) and bottom (blue solid line) quartile of choice accuracy.

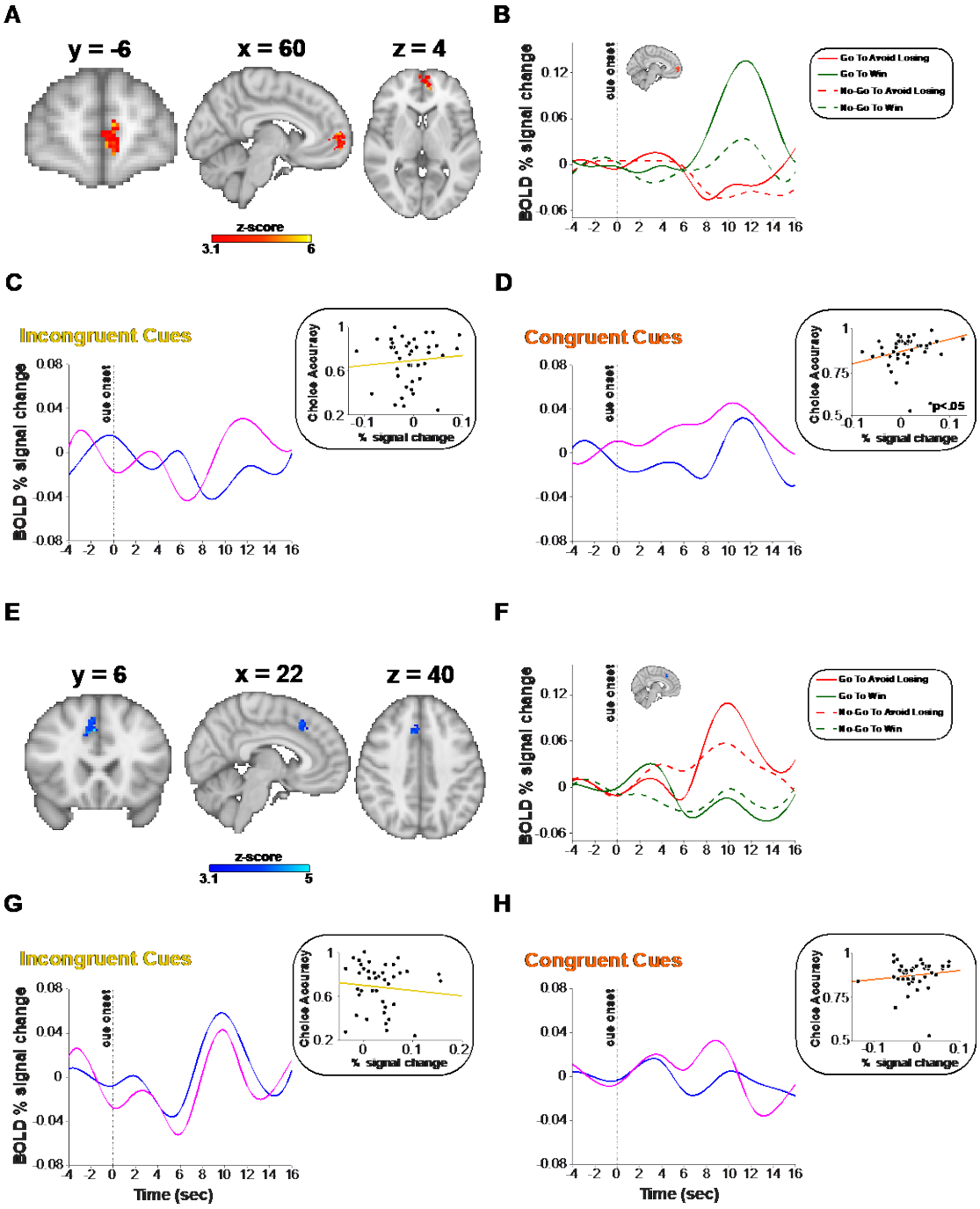

Figure 4. Neural correlates of Pavlovian value. A) fMRI cluster in the vmPFC encoding for positive Pavlovian value ($p < .05$ FWE). MNI coordinates are shown. **B)** vmPFC BOLD traces locked to cue onset as a function of action requirement and valence. Haemodynamic activity shows noticeable positive valence effect (win > lose cues). **C-D)** The insets show the scatterplots of vmPFC BOLD activity (averaged over transparent grey time window shown in the BOLD traces) as a function of choice

accuracy during incongruent (C) and congruent (D) trials. Notably, mean vmPFC haemodynamic responses to congruent (but not incongruent) cues are significantly associated with task performance. **E)** fMRI cluster in the dorsal ACC encoding for negative Pavlovian value ($p < .05$ FWE). **F)** Dorsal ACC BOLD traces locked to cue onset as a function of action requirement and valence. Haemodynamic activity shows noticeable negative valence effect (lose > win cues). **G-H)** The insets show the scatterplots of dorsal ACC BOLD activity (averaged over the transparent grey time window shown in the BOLD traces) as a function of choice accuracy during incongruent (C) and congruent (D) trials. Solid colour-coded lines denote 20% bend correlation fit. Black dots represent individual subjects. BOLD traces represent subjects in the top (purple solid line) and bottom (blue solid line) quartile of choice accuracy.

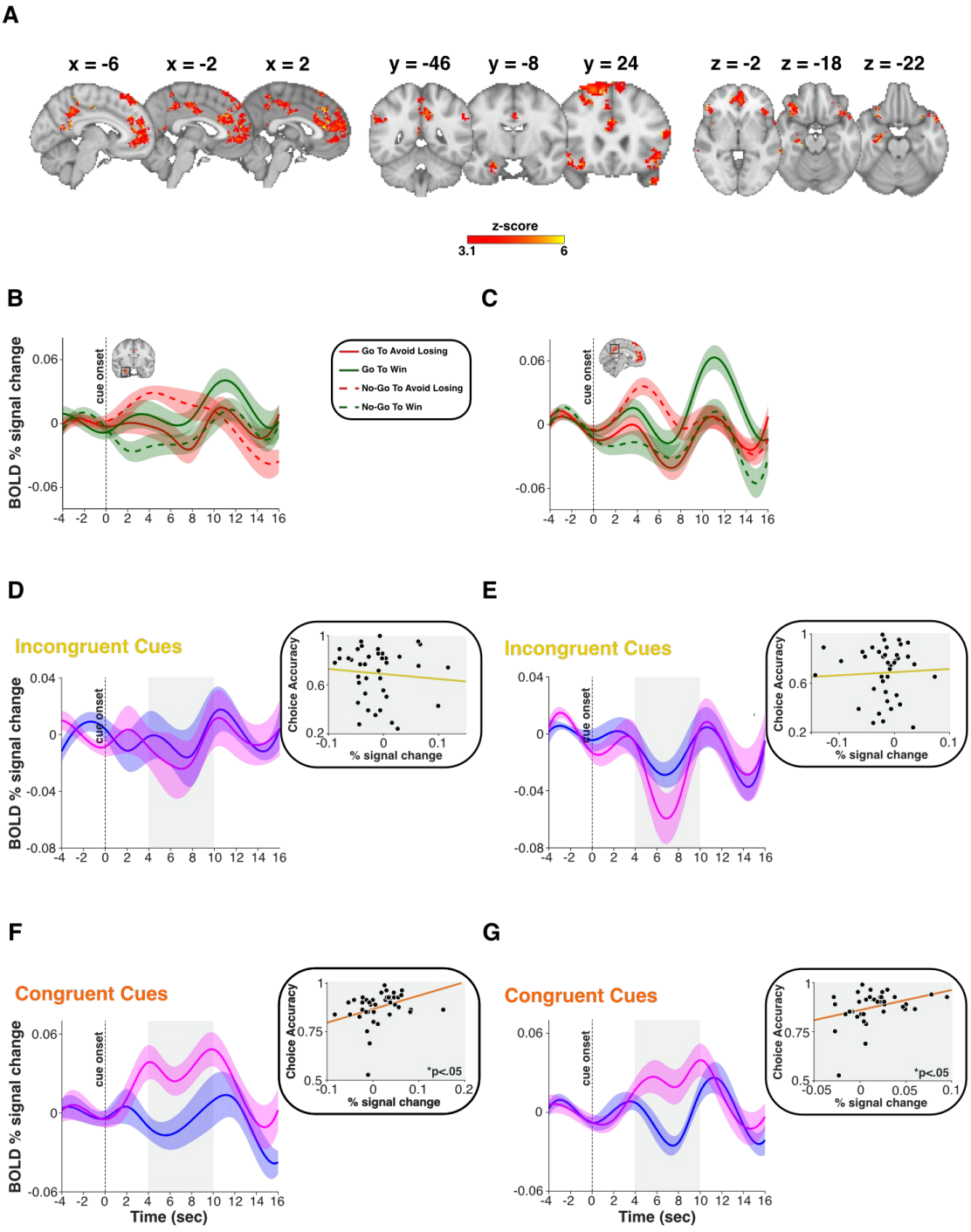

Figure 5. Neural correlates of congruent PII effects. A) fMRI clusters encoding for appetitive and aversive congruent PII effects ($p < .05$ FWE). MNI coordinates are shown. **B-C)** BOLD traces extracted from right amygdala/hippocampus complex (B) and bilateral PCC (C) locked to cue onset as a function of action requirement and valence. Haemodynamic activity shows noticeable congruence effect (congruent > incongruent cues) with earlier peak for aversive and later peak for appetitive congruent

cues. **D-G)** The insets show the scatterplots of BOLD activity (averaged over the transparent grey time window shown in the BOLD traces) in the right amygdala/hippocampus complex (D/F) and bilateral PCC (E/G) as a function of choice accuracy during incongruent (D/E) and congruent trials (F/G) with solid colour-coded lines denoting 20% bend correlation fit. Black dots represent individual subjects. Notably, mean haemodynamic responses to congruent (but not incongruent) cues in the right amygdala/hippocampus complex and bilateral PCC are significantly associated with task performance. BOLD traces represent subjects in the top (purple solid line) and bottom (blue solid line) quartile of choice accuracy.

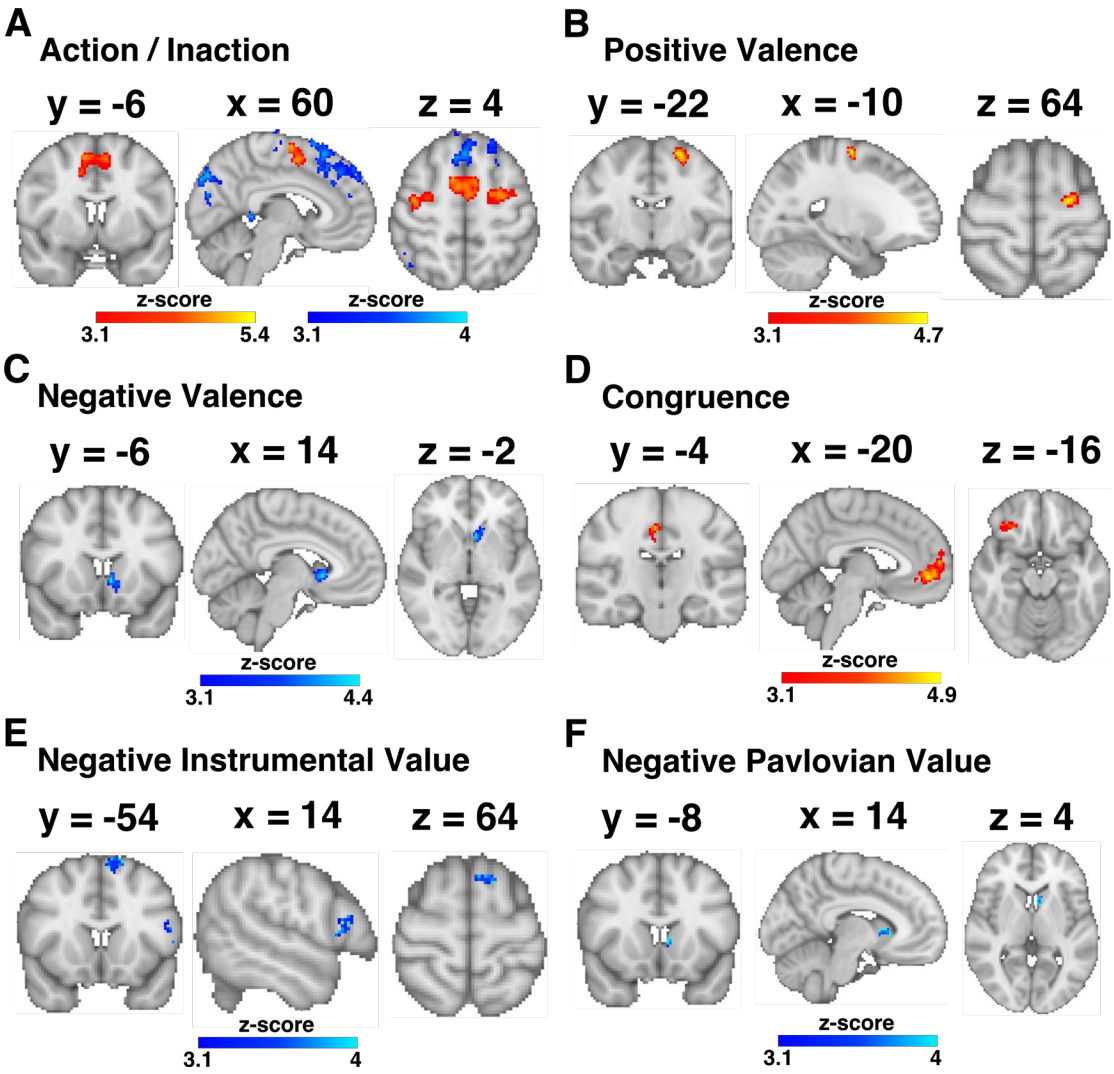

Figure 6. A-D) Results from task-informed fMRI analysis showing significant clusters associated with action (go>no-go cues), inaction (no-go>go cues), positive valence (win>lose cues), negative valence (lose>win cues) and congruence (congruent>incongruent cues). **E-F)** Results from model-informed fMRI analysis (whole-brain (E) and using an anatomical mask of the striatum (F)). A cluster in the left IFG was significantly associated with negative instrumental value (E). A cluster in the left medial caudate was significantly associated with negative Pavlovian value (F). MNI coordinates are shown.

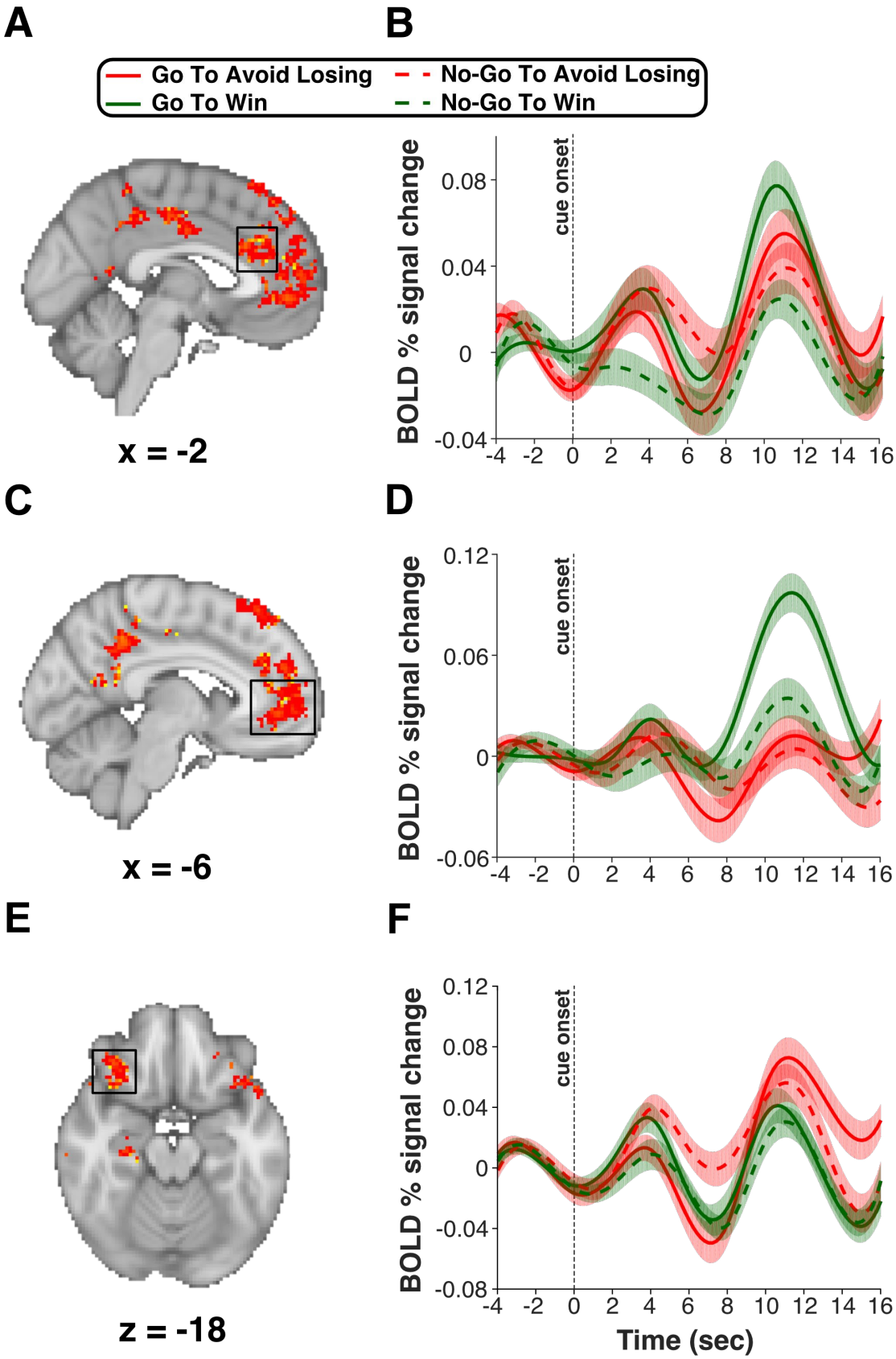

Figure 7. fMRI clusters encoding for appetitive and aversive congruent PII effects ($p < .05$ FWE) and BOLD traces locked to cue onset as a function of action requirement

1225 and valence. MNI coordinates are shown. **A-B)** Dorsal ACC. **C-D)** Bilateral medial
1226 PFC. Congruence effect is primarily driven by go to win cues. **E-F)** Right OFC.
1227

	Z-MAX	MNI x	MNI y	MNI z
Left Superior Frontal gyrus	4.68	-8	16	66
Right Frontal pole	4.42	54	40	-6
Right Cuneus	4.78	4	-80	40
Right Lingual gyrus	4.6	14	-52	2
Left Frontal pole	4.46	-24	58	24
Right Lateral Occipital cortex	4.45	46	-62	32
Left Fronto-orbital cortex/Inferior Frontal gyrus	4.41	-44	22	-10
Right Inferior Frontal gyrus (pars opercularis)	4.35	54	18	24
Left Precentral gyrus	4.06	-18	-30	60
Left Central Opercular cortex	4.22	-52	-10	10
Left Intracalcarine cortex	4.28	-14	-66	10
Right Postcentral gyrus	4.44	44	-14	36
Left Cunealt cortex	4.11	-8	-88	30
Left Angular gyrus	4.08	-46	-54	32
Right Precentral gyrus	3.77	4	-18	62
Left Precentral gyrus	4.2	-48	-12	40

Table 1. Complete list of significant fMRI clusters for no-go > go contrast ($p < .05$ FEW). MNI coordinates of maximum z statistic are shown for each cluster.

	Z-MAX	MNI x	MNI y	MNI z
Right Superior Frontal gyrus	13.6	26	18	58
Left Lateral Occipital gyrus	14	-58	-62	18
Posterior Cingulate Cortex	8.33	4	-20	34
Left Temporal Pole	11.6	-58	8	-6
Right OFC	10.3	52	22	-10
Right Angular/ Supramarginal gyrus	13.6	46	-50	38
Right Lateral Occipital gyrus	6.83	48	-68	44
Right Hyppocampus / Amygdala	7.36	28	-22	-16
Right Precuneus	7.21	4	-52	12
Right Middle Temporal Cortex	8.32	64	-34	-4
Left Middle Frontal gyrus	5.85	46	10	56

Table 2. Complete list of significant clusters for congruent Pavlovian by Instrumental interaction contrast ($p < .05$ FWE). MNI coordinates of maximum z statistic are shown for each cluster.