Joint models reveal human subcortical underpinnings of choice and learning behaviour

Miletic, S.; Stevenson, N.; Bazin, P.-L.; Alkemade, A.; Isherwood, S. J. S.; Trutti, A. C.; ... ; Forstmann, B. U.

Joint models reveal human subcortical underpinnings of choice and learning behavior

Steven Miletić[a,b,1,2], **Niek Stevenson**[b,1], **Pierre-Louis Bazin**[c], **Anneke Alkemade**[b], **Scott J. S. Isherwood**[b], **Anne C. Trutti**[b], **Desmond H. Y. Tse**[d], **Asta K. Håberg**[e], **and Birte U. Forstmann**[b]

This manuscript was compiled on August 2, 2025

Decision making and learning processes together enable adaptive strategic behavior. Animal studies demonstrated the importance of subcortical regions in these cognitive processes, but the human subcortical contributions remain poorly characterised. Here, we study choice and learning processes in the human subcortex, using a tailored ultra-high field 7 T fMRI imaging protocol combined with joint models. Joint models provide unbiased estimates of brain-behavior relations by simultaneously including behavioral and neural data at the participant and group level. Results demonstrate relations between subcortical regions and the adjustment of decision urgency. Value-related BOLD differences were found with opposite BOLD polarity in different parts of the striatum. Multiple subcortical regions showed BOLD signatures of reward prediction error processing, but contrary to expectations, these did not include the dopaminergic midbrain. Combined, this study characterises the human subcortical contributions to choice and learning, and demonstrates the feasibility and value of joint modeling in facilitating our understanding of brain-behavior relationships.

Linking propositions | Error-driven learning | Reinforcement learning evidence accumulation models (RL-EAMs) | Bayesian hierarchical estimation

Significance Statement

Strategic behavior requires learning and decision processes to interact. The role of subcortical regions in these processes has been demonstrated in animal research, but is less well known in humans. Here, we use advanced methods to study human subcortical contributions to choice and error-driven learning. We used joint brain-behavior models, simultaneously fit to neural and behavioral data, which improve interpretation and statistical power. Results demonstrate wide-ranging subcortical involvement in value processing, reward prediction errors, and urgency settings. This study paves the way for applying joint models in studying brain-behavior relations, and for further refining our understanding of the human subcortex.

Decision making and instrumental learning continuously interact (1): Error-driven learning processes refine and update the information on which value-based choices are made. In behavioral studies, recent advances have integrated insights from the traditionally separate fields of perceptual decision-making on the one hand, and error-driven learning on the other, into a singular framework (2–13). The combination of evidence accumulation to threshold (a core principle from decision-making research) and simple delta rules (a core principle in reinforcement learning) was shown to provide a precise characterisation of behavior in instrumental learning tasks: It can explain response time distributions, choice accuracy, and the learning-related changes in response time distributions and choice accuracy.

While providing a rich account of the algorithmic processes underlying choice and learning, cognitive models are agnostic about the neural implementation, which is our focus here. Both fields can lean on rich literatures on the relation between neural and behavioral data, although based largely on animal recordings. In decision making, the basal ganglia have long been implicated in action selection (14–18)Furthermore, key insights were obtained from recordings that demonstrated processes resembling evidence accumulation in a variety of brain regions including the basal ganglia (19–21), the superior colliculus (22–24), and cortical regions including parietal cortex (25–29), the frontal eye fields (30–34), and premotor and motor cortex (35–38). In parallel, studies in reinforcement learning have long focused on the role of the dopaminergic midbrain in calculating reward prediction errors, and on dopamine as a signal conveying reward prediction errors (e.g., 39–43).

Thus, both fields suggest prominent involvement of subcortical regions. Unfortunately, in humans, the role of subcortical regions in decision making and learning is less well characterised (44). This is due to various factors that make imaging the subcortex particularly difficult. Many subcortical regions suffer from signal losses when conventional functional magnetic resonance imaging (fMRI) methods are used. The underlying causes include the deep location of the subcortex, high iron concentrations, and the small sizes of individual regions (for an overview, see 45). Because of these factors, the majority of human neuroimaging studies have focused on the neocortical sheet, combined with the larger subcortical regions including the striatum and thalamus (for a meta-analysis, see 46). To achieve the signal quality necessary for investigating the typically small blood-oxygenation level dependent (BOLD) responses associated with cognitive functions in smaller

Author affiliations: [a]Institute of Psychology, Leiden University, Wassenaarseweg 52, 2333 AK, Leiden, The Netherlands; [b]Department of Psychology, University of Amsterdam, Nieuwe Achtergracht 129-B, 1018 WS, Amsterdam, The Netherlands; [c]Full Brain Picture Analytics, Lage Morsweg 73, 2332 XB, Leiden, The Netherlands; [d]Department of Neuropsychology and Psychopharmacology, Maastricht University, Minderbroedersberg 4-6, 6211 LK Maastricht, The Netherlands; [e]Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, Trondheim, 7030, Norway

Conceptualisation: SM, PLB, SJSI, ACT, BUF. Methodology: SM, NS, PLB, AA, BUF. Software: SM, NS, PLB. Validation: SM, NS, PLB, SJSI, ACT, DHYT. Formal analysis: SM, NS, PLB. Investigation: SM, PLB, SJSI, ACT, DHYT. Resources: DHYT, AKH, BUF. Writing - Original draft: SM. Writing - Review & Editing: All authors. Visualisation: SM. Supervision: PLB, AA, AKH, BUF. Project administration: AKH, BUF. Funding acquisition: AA, BUF.

PLB is the owner of Full Brain Picture Analytics. The other authors do not have any competing interests to declare.

[1]SM contributed equally to this work with NS

[2]To whom correspondence should be addressed. E-mail: s.mileticfsw.leidenuniv.nl

regions, specialised MRI protocols designed at ultra-high field strengths of 7 Tesla (T) have been developed (47–49).

Signal quality is not the only factor to consider when discussing the challenges of studying the human subcortex. Statistical considerations form a second factor hampering the characterisation of the human subcortex in cognitive processes. Model-based analysis methods offer a principled advantage in terms of bridging the algorithmic and neural levels of analysis (50, 51). Traditional model-based MRI studies, however, rely on *two-stage* approaches, in which a cognitive model is first fit to behavioral data, and the resulting parameters are used as regressors in the analysis of the neural data. While straightforward to implement, two-stage approaches do not fully take into account the reciprocity in the relation between behavior and the brain. Using this approach, the neural model is informed by the cognitive data, but the cognitive model is not informed by the neural data. Furthermore, the measurement uncertainty in the parameters of the cognitive model are ignored. When unaccounted for, this source of noise causes negatively biased effect sizes, a phenomenon known as attenuation (52, 53). It also comes at the risk of overconfidence in the effects of covariates, since the uncertainty in the estimation of the covariate is ignored (52). This is especially detrimental when studying noisy data such as fMRI timeseries obtained from the human subcortex. Joint models, which simultaneously model both the neural and behavioral modalities of data, at all levels of the hierarchy (participant and group level), are required to remedy this issue and achieve full statistical power (50, 54–57).

This study takes a joint modeling approach to studying decision-making processes and instrumental learning in the human subcortex. We bring together three contributions. Firstly, we use a single task paradigm combined with a single cognitive model, that unifies the study of decision making and reinforcement learning processes (6, 7), and allows for disentangling potential interactions between decision making and learning. In this task, participants are required to repeatedly make value-based choices between abstract symbols, and learn from the probabilistic reward associated with each symbol. Prior to each choice, participants are informed to emphasise either response speed or choice accuracy, thereby enforcing a change in choice strategy.

Secondly, we used an fMRI protocol tailored to meet the specific requirements for studying small subcortical nuclei at an ultra-high field of 7 T (47–49, 58–62). This protocol includes a short echo time to match the low T2* of iron-rich nuclei, small voxels to mitigate partial voluming effects, and a relatively high repetition time. Furthermore, we acquired multimodal quantitative anatomical MRI data, which enabled us to delineate individual subcortical nuclei with automated algorithms (63).

Finally, we analysed brain-behavior relations in the resulting data using high-powered Bayesian joint modeling techniques, in which two reciprocal links between neural and behavioral data are included: Reward prediction errors and value estimates of the reinforcement learning model are fed forward to the neural models within subjects, and simultaneously, across participants, inter-individual correlations between neural and behavioral model parameters are estimated. The simultaneous estimation of the cognitive and neural models allows for all sources of uncertainty to be modeled accurately, which leads to unbiased estimates of the brain-behavior relations.

Results

Thirty-seven participants performed an instrumental learning choice task (Figure 1A) while undergoing 7 T BOLD-fMRI. They made repeated decisions between two abstract choice symbols, each followed by choice-dependent probabilistic rewards, which they used to inform subsequent choices. In total, each participant made 342 decisions. Prior to each decision, participants were instructed to emphasise either response speed or response accuracy. The behavioral data, consisting of response times and choices, were modeled with the reinforcement learning-advantage racing diffusion (RL-ARD) model (7). This model proposes that decisions are formed through an evidence-accumulation process, where the rate of accumulation depends on the sum of an urgency signal and the internal representations of the value of each choice option (Figure 2A). The values of choice options are learned via a standard delta rule (64). The effect of the speed and accuracy instructions were modeled by allowing both the urgency and threshold parameters to vary with instructions, in line with previous work (7). Threshold refers to the overall amount of evidence that participants require to inform their decisions, whereas urgency refers to how participants become less patient as within a trial as time passes. In previous work (7), we demonstrated favorable parameter recovery properties with this exact paradigm (see their Figure 7-figure supplement 3).

We used mixed effects models (MEMs) to confirm the difficulty (defined as the difference in pay-off between the choice options) and speed-accuracy trade-off (SAT) manipulations had the intended effects on behavior. In the MEMs, fixed effects of difficulty, SAT, and their interaction were estimated, as well as random effects of difficulty and SAT. A linear MEM indicated a significant fixed effect of the SAT cues on RT ($t(66.49) = -10.413, p < 0.001$), but not of difficulty nor an interaction. A generalized MEM demonstrated an interaction between SAT cue and difficulty on choice accuracy ($z = -2.576, p = .01$), as well as a main effect of SAT cue ($z = 6.04, p < .001$), with larger SAT effects on accuracy in the easy trials compared to the hard trials (1C). Moreover, the RL-ARD provided a generally adequate account of the behavioral data, capturing the learning-dependent increase in accuracy, decrease in response time, and the differences in RT and choice accuracy between the speed-emphasised and accuracy-emphasised trials (Figure 1B-C). Note that there was some misfit in of the RTs in the early trials, which replicates an earlier finding with the same paradigm and model (7).

To ensure that the SAT manipulation did not affect reward prediction error processing, we fit a second RL-ARD specification that allowed learning rates to differ between SPD and ACC trials. The estimated learning rates (Figure 1D) show large overlap, and formal model comparisons suggested that an RL-ARD with a single learning rate provided a better trade-off between fit and model complexity (BPIC difference = 66 in favor of the simpler model; see Table S4 for participant-wise BPIC values of both models). In supporting information, we performed a simulation study which demonstrated that our sample size, trial numbers, and

fitting methods would favor a two-learning rate model if a true learning rate difference were 0.05 or larger. We also tested whether there was any between-cue difference in the effect of reward prediction error on the subsequent trial's RT (i.e., post reward prediction error slowing). A linear MEM showed evidence for a main effect of RPE on subsequent RT ($t(37.12) = 2.592, p = .014$), as well as a main effect of the previous trial's cue ($t(37.86) = -3.542, p = .001$), but no interaction between RPEs and SAT condition ($t(44.87) = -0.386, p = 0.701$; Figure 1E). Combined, the behavioral data and RL-ARDs suggest that the manipulations had the intended effects and the RL-ARD with a single learning rate provided an sufficient account of the behavioral data.

In a separate session, participants underwent high-resolution quantitative MRI scans that allowed us to derive multimodal anatomical data (T1 maps, T2* maps, and quantitative susceptibility maps), which were used to delineate 17 subcortical regions of interest using the multi-contrast anatomical subcortical structure parcellation (MASSP) algorithm at the individual level (63). The masks of the gray matter structures — the amygdala (Amg), claustrum (Cl), globus pallidus interna (GPi) and externa (GPe), periaqueductal gray (PAG), pedunculopontine nucleus (PPN), red nucleus (RN) substantia nigra (SN), subthalamic nucleus (STN), striatum (Str), thalamus (Tha), and ventral tegmental area (VTA) — were subsequently used to extract timecourses of the signal from the fMRI data. Figure 4A provides an overview of these ROIs.

These neural fMRI timecourses were modeled with a general linear model (GLM; Figure 2E) which, next to a set of nuisance regressors (see Methods), included cues (speed and accuracy), stimulus value differences, and reward prediction errors, as regressors of interest. The latter two regressors were derived from the RL-ARD model, and vary across trials within participants. We estimated their mean effect on the group level (Figure 2C). We also estimated the correlations between the speed–accuracy contrasts in the neural models (one per region of interest) and speed–accuracy difference between the urgency and threshold differences as derived from the RL-ARD (Figure 2D). Combined, this resulted in three brain-behavior relations per region of interest that were jointly informed and reciprocally constrained by the two modalities of data.

The resulting joint model is visualised in Figure 4. Figure 4B shows the inter-individual correlations between strategic adjustments in choice behavior (urgency and threshold) and the BOLD responses in the subcortical regions (see Table S1). Although the thresholds were overall higher in the ACC condition than in the SPD condition, the joint models revealed across-participant correlations between urgency and neural responses bilaterally in the Str and VTA, left Cl, and right RN and Tha. Next, we turned to brain-behavior relations of value learning. The PPN and SN showed relations with value differences, as well as the left PAG (Figure 4C). The joint model further indicated reward prediction error processing in the Amg, Cl, GPe, and Str (Figure 4D). Interestingly, we found no evidence for involvement of the VTA or SN in reward prediction error coding; if anything, results indicated a *negative* association between reward prediction errors and neural activity in the right SN.

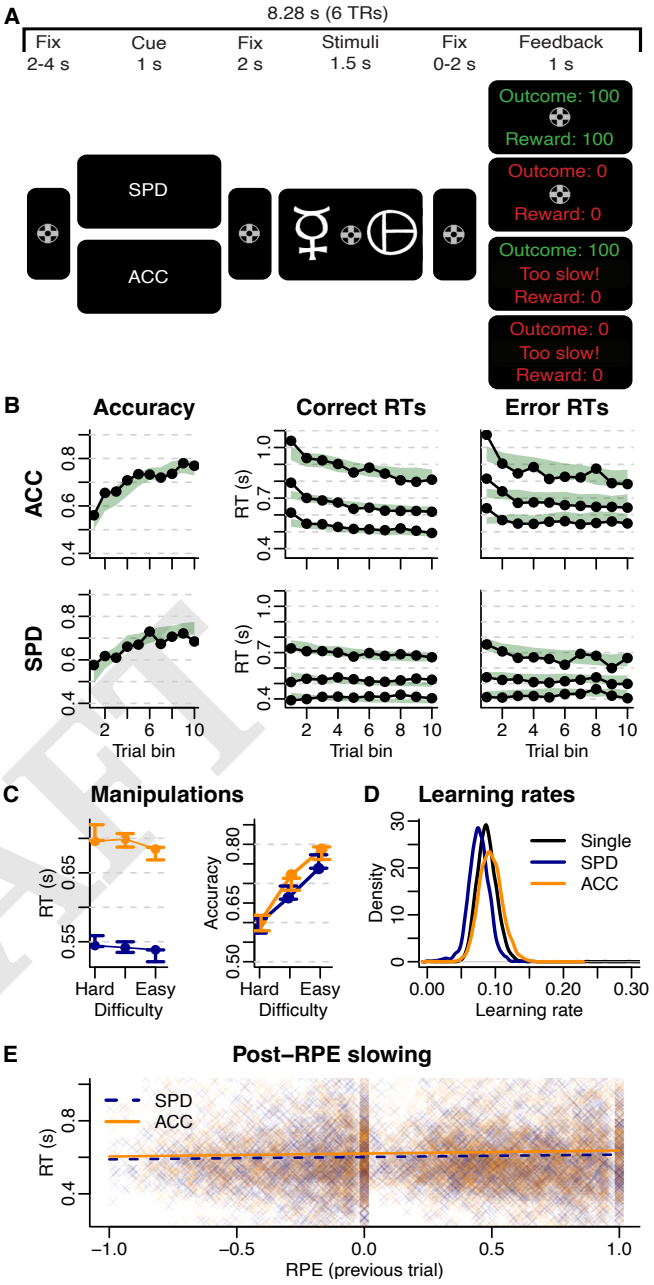

Fig. 1. A. Experimental paradigm. Each trial started with a fixation cross, followed by the speed–accuracy trade-off (SAT) cue ('SPD' or 'ACC'), another fixation cross, the stimuli representing choice options, another fixation cross, and feedback. Feedback depended both on the response time (in time or too slow) and on the outcome of the probabilistic gamble (0 or 100 points). Rewards were only given if the response was in time. Durations of the fixation crosses were jittered to decorrelate event timing. B. Data (black) and model fit (green) of the RL-ARD model in the accuracy (top) and speed (bottom) condition. Left column depicts accuracy over trials across the run. To visualize the learning effects, all trials were binned into 10 bins (approximately 17 trials per bin), and summary statistics were calculated per bin. Middle and right panel show 10th, 50th, and 90th RT percentiles for the correct (middle) and error (right) response over trial bins. Shaded areas correspond to the 95% credible interval of the model fit. C. Effects of the difficulty (x-axis) and SAT manipulations (orange = ACC, blue = SPD) on mean RT (top) and accuracy (bottom). Points are data, error bars the 95% credible interval of the model. Difficulty is defined as the difference in pay-off probability between the two choice options (smaller is harder). D. Estimated (posterior) learning rates for an RL-EAM with a single learning rate (black), and an RL-EAM with separate learning rates for SPD and ACC trials. E. Effect of reward prediction error (RPE) size on subsequent RT. Individual shaded crosses are trials, lines indicate linear mixed effects model predictions of the fixed effects of RPE size per previous trial's cue type.

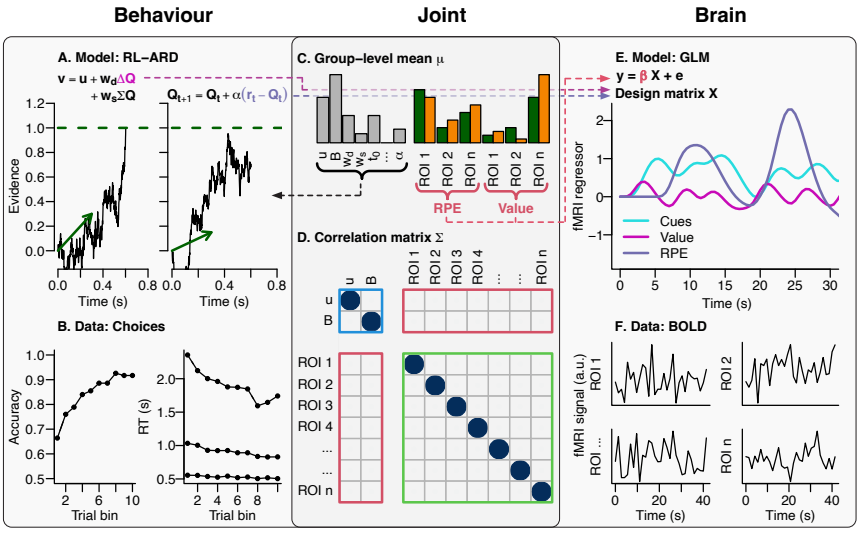

Fig. 2. Overview of joint modeling approach. The behavioral model (A) is informed by the RT and choice data (B; see Figure 1B for detail on the visualisation of the behavioral data). The trial-by-trial differences in Q-values and prediction errors are fed forward to the design matrix of the GLM (E). The GLM is informed by the neural data (F). Mutual constraint between the two modalities of data is enabled by the joint structure that uses a multivariate normal distribution at the group level. This is described by a group-level mean (C) and correlation matrix (D). All behavioral and neural parameters are estimated simultaneously on the group level and participant level. Brain-behavior associations of reward prediction errors and value differences are characterised with group-level means, while brain-behavior associations between speed-accuracy trade-off behavior and neural responses are estimated as inter-individual correlations. The correlation matrix is divided into behavior-behavior correlations (blue rectangle), brain-brain correlations (green), and brain-behavior correlations (red). For visualisation purposes, only a subset of the parameters are shown.

To investigate this further, we fit another joint model that used an RL-ARD with two separate learning rates for SPD and ACC trials. We reasoned that, although the behavioral evidence indicated no evidence for separate learning rates, the neural data might be more sensitive to such a difference. Hence, in the GLMs, we estimated separate parameters for the modulatory effect of RPEs on BOLD responses for the SPD and ACC trials. In Figure S3, we show the effect of RPE on BOLD responses in the MASSP ROIs, which lead to the same overall conclusions as the joint model that assumed no difference between SPD and ACC trials in learning rates or RPE processing.

The results so far indicated involvement of the Tha (as a single region covering all nuclei) in the speed-accuracy trade-off. In a second joint model, we zoomed in on the individual thalamic nuclei using a thalamus atlas (65). Here, we focused only on regions larger than 150 mm^3 in both hemispheres: the anteroventral (AV), centromedian (CM), lateral posterior (LP), mediodorsal (MD), pulvinar, ventral anterior (VA), ventral lateral (VL), and the ventral posterolateral (VPL) nucleus. In the atlas, the MD is split into a lateral and medial part (MDl, MDm), the pulvinar in an anterior, inferior, lateral, and medial part (PuA, PuI, PuL, PuM), and the VL in an anterior and posterior part (VLa, VLp). Figure 5A illustrates the ROIs that were included. The joint model based on thalamic nuclei highlighted that the brain-behavior correlations with speed-accuracy trade-off settings were found bilaterally within the AV, CM, MDm, PuA and PuM, as well as in the right LP, VLa, and VLp (Figure 5B, see Table S2). Again, these correlations are with urgency, and appear to dominate in the right hemisphere. In the thalamic regions, we found evidence for a relation with value difference only in the right VPL (Figure 5C). Evidence for reward prediction error processing was found in the CM, PuI, and VPL (Figure 5D).

In a third and final joint model, we zoomed in on the striatum. Unlike the thalamus, the human striatum is a relatively homogeneous structure, without clear internal cytoarchitectural or immunohistochemical boundaries between the dorsal and ventral striatum (e.g., 66). However, it has long been argued to be functionally specialised in multiple zones (e.g., 67), with distinct afferent projections (66, 68). Here, we used the recently developed second iteration of MASSP (69) to delineate the striatum into three separate parts: the nucleus accumbens (nAcc), putamen (Pu*), and caudate (Cau) (Figure 6A). We would like to point out that the nAcc in MASSP was delineated using a perpendicular line at the base of the internal capsule, which may result in the inclusion of an area that is not fully restricted to the nAcc. This approximation of the border of the nAcc is required, since visualisation of the border can only be achieved using post mortem histology. The joint model fit to the timeseries of these subregions is shown in Figure 6B-D (see Table S3) for numerical estimates). The brain-behavior association relating to speed-accuracy trade-off settings was strongest in the dorsal striatum (Pu and Cau), and only credible in the right (but not left) nAcc. As expected, reward prediction error processing was clearest in the nAcc, but also detectable in both the Pu and the Cau (Figure 6C). A positive association between the size of the BOLD responses and the size of value differences was found in the Pu, and interestingly, a *negative* association in the Cau (and no association in the nAcc) (Figure 6D).

Finally, we confirmed empirically that joint models provided more statistical power compared to a two-stage approach. To demonstrate the two-stage approach, we first estimated the behavioral model. Based on the median of the posterior parameters and the experimental paradigm, we generated trial-by-trial stimulus and reward prediction error values, which were used to generate design matrices for the neural GLMs. We then estimated the neural GLMs as well. In a second stage, we fit a multivariate Gaussian distribution on the subject-level median behavioral and neural parameters, using a Bayesian estimation routine. This way, we still estimate a distribution of correlation coefficients, but not jointly with the neural and behavioral models. Figure 3 compares the two-stage brain-behavior correlation distributions with the joint model correlation distributions for the five ROIs with largest correlations in

*Note that both the thalamic atlas and the second iteration of MASSP include 'Pu' as an abbreviation; the former referring to the Pulvinar, the latter to the Putamen. In this manuscript, Pu refers to the Putamen, and PuA, PuI, PuL and PuM to the various Pulvinar regions.

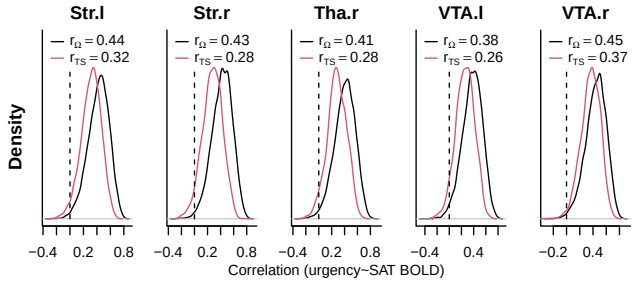

Fig. 3. Comparison of two-stage (TS, red) posterior correlation coefficients (between the behavioral and neural SAT effects) with joint (Ω, black) posterior correlation coefficients for five MASSP ROIs. Vertical dotted lines indicate correlations of 0. Correlation coefficients in the legend indicate the median of the distributions.

Figure 4. This demonstrates clearly attenuated effect sizes in the two-stage approach, with approximately 20%-40% smaller median correlation coefficients in the latter case.

Discussion

In this study, we use joint models to understand the brain-behavior relations between subcortical regions and decision-making and learning. With tailored methods, including ultra-high field 7 T fMRI, decision making and instrumental learning were jointly studied in a single paradigm and corresponding cognitive model, in a Bayesian hierarchical joint modeling framework in which brain-behavior relationships were reciprocally informed by all modalities of data. The resulting joint models revealed that the Str (and particularly the dorsal Str) was involved in choice strategy settings; however, contrary to previous reports, they demonstrated a relation with urgency, rather than response caution. Next, they revealed value-related processing, but not reward prediction error processing, in the substantia nigra. Finally, within the Str, value-related processing was demonstrated to show BOLD responses with opposite polarities in the caudate and putamen.

Our results indicate that subcortical regions may contribute to strategic control of choice behavior through urgency, rather than response caution settings, which has been argued previously (70, 71). At the group level, thresholds were higher in ACC trials compared to SPD trials, as is commonly found. However, the effect of the manipulation on urgency, not threshold, covaried with neural signals. In part, this may arise from the use of the RL-ARD, which is able to dissociate urgency from response caution adjustments, which themselves correlate (e.g., Figure 4). The implication of urgency adjustments corroborates earlier studies based on neural recordings the basal ganglia in monkeys (19, 72), as well as fMRI evidence using an expanded judgment task (73). The dominance of the right hemisphere in these relations is consistent with previous studies (70, 71, 73), and may be related to the right-lateralised response inhibition networks (74–76).

While our model-based approach is able to dissociate between urgency and threshold, the concept of urgency itself is not singular, as multiple cognitive processes may contribute to or correlate with urgency signals. Understanding these processes may help explain why we found urgency-related signals in so many different regions. For one, urgency is known to be related to arousal (77). In Supporting Information (Fig S4), we tested whether the SPD cues had a different effect compared to ACC cues on heart rate variability and respiratory volume per time (as potential correlates of arousal), but found no evidence for any difference. However, subtle arousal-related differences could have remained undetected. Future studies could include pupillometry (78) to test whether the identified urgency signals reflect pupil-link arousal in relevant subcortical areas. For example, the CM plays an important role in modulating arousal (79) (and covaries with reward prediction errors; 80, 81). Second, the MD has been implicated in various types of memory processing, including object-reward association memory (e.g. 82–86). The role of the MD may be to prepare the memory processes required for the subsequent value-based decision, and such preparations could start earlier under speed stress. Some evidence also suggests a role for the AV in modulating cortical plasticity and memory formation (87). The involvement of the RN and VTA in urgency has, to our knowledge, not been demonstrated before, but may be related to earlier studies that demonstrate these regions' involvement in conflict resolution, which is potentially caused by the conflicting instructions of the speed and accuracy requirements (60, 88). Third, urgency may cause attentional processes as well. In earlier behavioral work, we tested for effects of SAT cues on attention in this paradigm (7), but model comparisons preferred models without attention effects. It might be that the effects of attentional processes on behavior were too subtle to be picked up, but their effects on the present neural data are more marked. Fourth, it has recently been proposed that people's decision processes in accuracy-emphasised trials contain one additional phase of cognitive processing compared to speed-emphasised trials, suggesting that there may be qualitatively different decision processes in speed and accuracy trials (89). Additionally, as noted in the introduction, evidence accumulation signals have been found in a wide variety or cortical and subcortical regions before (19–38). The discovery of brain regions that correlated with urgency settings, and their function, can help us theorise about potential confounds of urgency that are difficult to derive based on behavioral studies alone. Model-based analyses should be combined with clever experimental design and manipulations to disentangle the influences of various confounding factors to estimated brain-behavior relations.

Our results further indicated value-related processing in the Str, but with opposite polarities in the Cau compared to the Pu. This striking result might reflect a gradient of functional specialisation related to value differences. Alternatively, recent research has shown that neural activity in the dorsal Str can elicit vasoconstriction and *negative* BOLD responses, implying that our finding of negative BOLD responses could nonetheless indicate increased neural activity (90). Note that value differences, in the present design, are confounded by other factors, which importantly includes difficulty: A choice based on two stimuli which differ in their value is easier compared to stimuli with similar values. Additional confounding factors include salience and arousal effects (see also 91). Disentangling the influence of these factors requires specific experimental designs in future studies.

We further found various subcortical regions in which BOLD responses covaried with reward prediction error sizes. While amygdalar and striatal involvement in reward

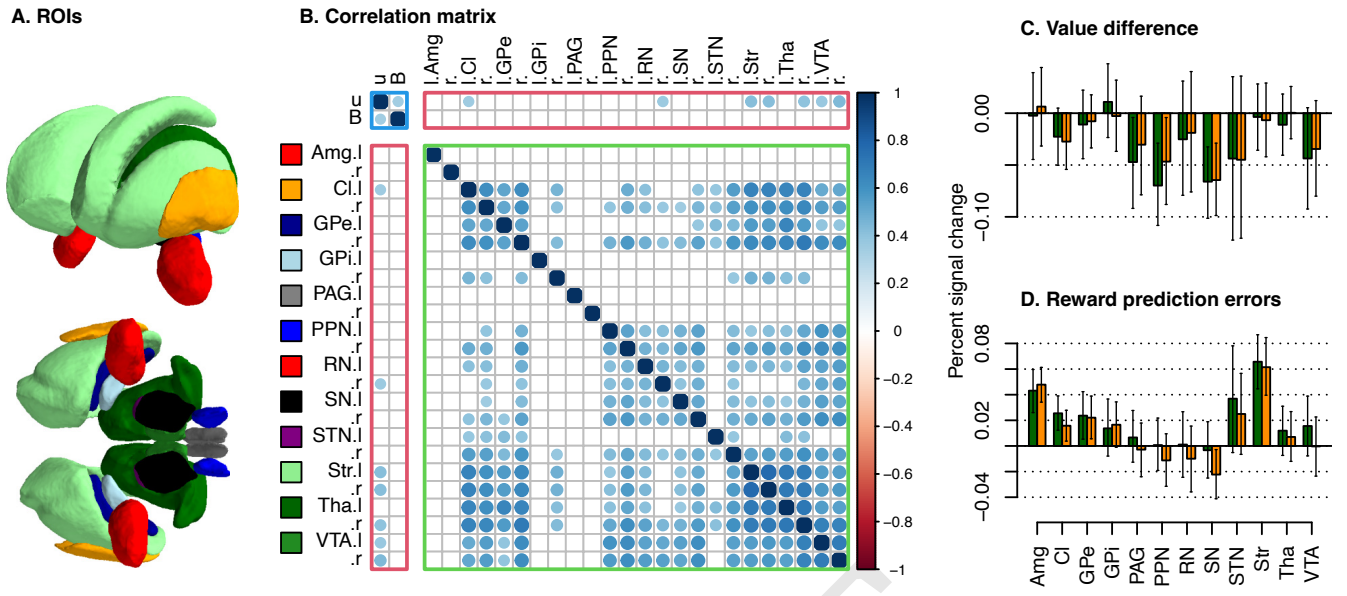

Fig. 4. Joint model fit to the MASSP ROIs. A. Illustration of the ROIs, viewed from the front-left (top) and bottom (bottom). B. Group-level correlation matrix, which is split into behavior-behavior relations (outlined by a blue rectangle), brain-brain relations (red), and brain-behavior relations (green). Only credible correlations are shown; non-credible correlations are displayed as empty squares. Relations are considered credible when the 95% credible interval of the correlation coefficient does not cover 0. All parameters are related to the speed-accuracy trade-off contrast: Its effect on urgency (u), threshold (B), and the BOLD contrast in the ROIs. C and D. Group-level estimates of within-participant brain-behavior relations of value learning and reward prediction errors. Barplots show the percentage signal change per unit change in value difference (C) and reward prediction errors (D), for each region of interest. Green and orange bars depict the left and right hemisphere, respectively. Error bars indicate 95% credible intervals.

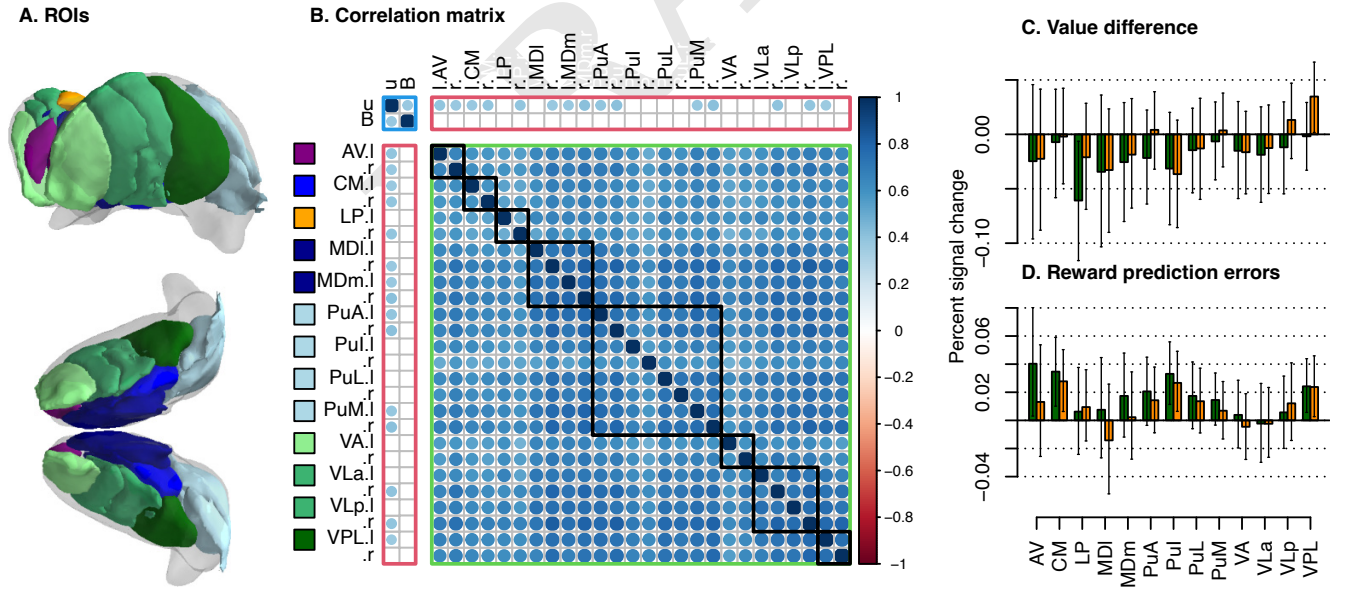

Fig. 5. Joint model fit to the thalamus ROIs. A. Illustration of the ROIs, viewed from the front-left (top) and bottom (bottom). Meshes were generated by first warping all individual-level delineations to MNI-space, and subsequently running the marching cube algorithm on the across-participant mean in MNI-space. For comparison, the MASSP delineation of the thalamus is illustrated in transparent white. B. Group-level correlation matrix, which is split into behavior-behavior relations (outlined by a blue rectangle), brain-brain relations (red), and brain-behavior relations (green). Subregions belonging to the same nuclei are clustered along the diagonal with black squares. Only credible correlations are shown; non-credible correlations are displayed as empty squares. Relations are considered credible when the 95% credible interval of the correlation coefficient does not cover 0. All parameters are related to the speed-accuracy trade-off contrast: Its effect on urgency (u), threshold (B), and the BOLD contrast in the ROIs. C and D. Group-level estimates of within-participant brain-behavior relations of value learning and reward prediction errors. Barplots show the percentage signal change per unit change in value difference (C) and reward prediction errors (D), for each region of interest. Green and orange bars depict the left and right hemisphere, respectively. Error bars indicate 95% credible intervals.

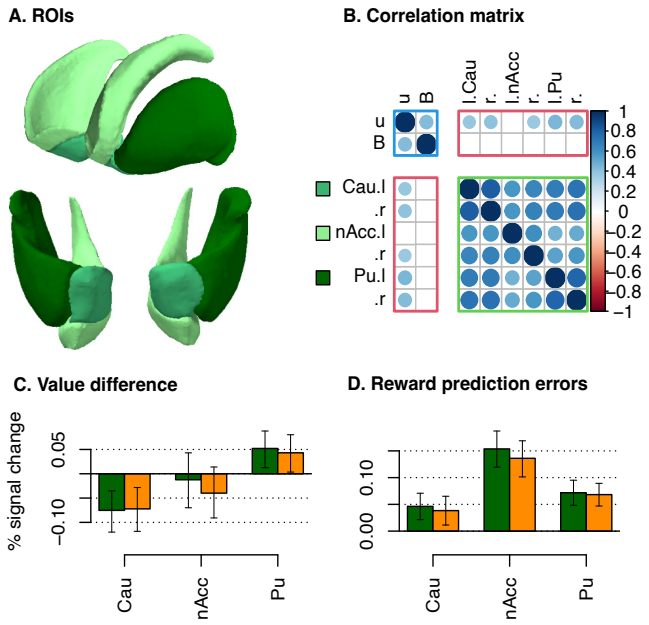

Fig. 6. Joint model fit to the striatum ROIs. A. Illustration of the ROIs, viewed from the front-left (top) and bottom (bottom). B. Group-level correlation matrix, which is split into behavior-behavior relations (outlined by a blue rectangle), brain-brain relations (red), and brain-behavior relations (green). Only credible correlations are shown; non-credible correlations are displayed as empty squares. Relations are considered credible when the 95% credible interval of the correlation coefficient does not cover 0. All parameters are related to the speed-accuracy trade-off contrast: Its effect on urgency (u), threshold (B), and the BOLD contrast in the ROIs. C and D. Group-level estimates of within-participant brain-behavior relations of value learning and reward prediction errors. Barplots show the percentage signal change per unit change in value difference (C) and reward prediction errors (D), for each region of interest. Green and orange bars depict the left and right hemisphere, respectively. Error bars indicate 95% credible intervals.

prediction error coding are well-documented (e.g. 80, 81), the Cl and GPe received less attention in the literature. Recently, electrophysiological recordings in rodents identified a neural subpopulation encoding reward prediction errors in the GPe (92). To some extent, these may arise also from covariates, such as perceived saliency (93, 94). The Cl involvement might indicate a functional role similar to the Amg in terms of arousal and salience detection (95). It is becoming increasingly clear that dopamine signals can be detected in a wider range of behaviors than classical reward prediction errors, and can also signal sensory and motor features (for review, 43). Under the generalised prediction error framework (96), they are argued to also indicate errors in the sensory world model, and are used to improve the world model. Consequently, a wide set of brain regions are likely involved in the processing of these predictions errors.

Contrary to some previous reports, we did not find evidence for dopaminergic midbrain involvement in reward prediction error encoding. A long history of animal recordings has implicated especially the VTA in reward prediction error processing (e.g. 39–42, 97), which has partially been supported in humans using fMRI (98–102), but not consistently (see 103, for a meta-analysis). A variety of factors has been argued to contribute to this discrepancy, including variability in the anatomical masks (102) and limited statistical power, as detailed in the introduction section. On the contrary, in the present study, the joint models were sufficiently powerful to identify value-related processing in the SN. Perhaps the discrepancy between the electrophysiology and BOLD findings is the result of a much more fundamental difference in methodology: While electrophysiology suggests reward prediction errors in the dopaminergic midbrain are encoded in spiking activity, BOLD responses have long been argued to correlate more strongly with synaptic activity (104–106), which could indicate local processing as well as input to a region. It has often been argued, for example, that the striatal BOLD responses are a result of dopamine release caused by dopaminergic midbrain neural spiking (107, 108). Intriguingly, since reward prediction errors are defined as the difference between obtained and expected reward, a region that calculates prediction errors needs expected reward (or value) as an input. This may explain why the SN BOLD responses were sensitive to value processing, but not reward prediction errors.

Subcortical regions play a prominent role in neurological disorders including Parkinson's disease (109) as well as psychiatric disorders like drug addiction (110) and social anxiety disorder (111–113). Parkinson's disease, for example, is associated with a specific loss of dopaminergic cells in the substantia nigra. Our results indicate a role for the substantia nigra in value processing. Earlier work suggests that the loss of dopaminergic cells in PD can lead to an increased propensity to learn from positive compared to negative outcomes, which can be reversed with dopaminergic medication (114). Learning biases are also crucial in addiction (115) and anxiety (e.g., 116, 117). Abnormal value computation may lead to an over-reliance on positive or negative outcomes. Task paradigms that disentangle reward and punishment learning can be used in future applications to test whether maladaptive value computation in disorders is associated with BOLD responses in the substantia nigra. Additionally, many of the subcortical regions we studied are (potential) targets for deep brain stimulation (DBS) in a variety of neurological and psychiatric disorders (e.g., 118). Other regions are also of potential interest, including the bed nucleus of the stria terminalis as a potential target for obsessive-compulsive disorder and the lateral habenula for major depression. Joint modeling approaches with specialised task designs can also be used to further understand these regions' functions in health and disease, especially in light of their ability to capture interindividual differences.

Especially in the context of translation to the clinic, it is important to consider the emotional and social components in tasks and models. In our current approach, we only relied on cognitive processes such as evidence accumulation and reward learning, but disorders such as social anxiety and autism include social and affective components, which can manifest as altered processing of social rewards (119). Combined with more complex paradigms, RL-EAMs and joint models can be further extended to better understand the brain-behavior relations in such disorders.

Despite a generally good fit of the RL-ARD, some misfit remains in the first trials. Factors such as increased uncertainty (120) could cause the relatively slow responses in the initial trials of each block. Additionally, it could be that the additional time participants take in the initial trials (relative to model predictions) reflect extra cognitive

processes involved in interpreting the abstract stimuli and forming memory traces. These memory traces are likely necessary for stimulus identification in later trials, where RTs are primarily governed by evidence accumulation based on Q-values. This hypothesis can be tested in a future experiment where the same stimulus sets are used across multiple blocks with new reward contingencies. The additional time should then be observed only in the first block in which a stimulus appears. It would then also be possible to assess whether the observed RT increase is better explained by heightened response caution or by an increase in non-decision time.

In conclusion, this study revealed various human subcortical underpinnings of decision making and learning. It uncovered new brain-behavior relations (e.g., thalamic nuclei in urgency settings, GPe in reward prediction error processing), and refined previous work (e.g., functionally specialised zones along the anterior-posterior axis of the Str in value processing). It also demonstrates feasibility and value of the combination of joint modeling and tailored fMRI methods in progressing our understanding of the human subcortex in cognitive processes.

Materials and Methods

Participants. Thirty-seven healthy volunteers (mean age 27 years old [SD 6 years, range 19–39 years old], 20 females) were recruited via local advertisement. The study was approved by the Ethics Review Board of the Faculty of Social and Behavioral Sciences of the University of Amsterdam (reference: 2021-BC-13146) and the Regional Committees for Medical and Health Related Research Ethics of Central Norway (reference: 116630). All participants gave written informed consent prior to the onset of the study. All participants were screened for MRI safety, had normal or corrected-to-normal vision, and no history of psychiatric and neurological illness. All participants participated in five scanning sessions as part of a larger project; here, we report and analyse two of these sessions.

Paradigm. The experimental paradigm made use of an instrumental learning task (114) with a cue-based speed–accuracy trade-off manipulation (7, see Figure 1). In every trial of the task, the participants made a decision between two abstract symbols, each associated with a fixed reward probability that is unknown to the participants. One choice option always had a higher probability of being rewarded than the alternative option. The participants received feedback in the form of points after each choice, which the participants can use to learn which symbols have the highest reward probabilities.

Prior to each trial, the participants were presented with a cue to instruct them to emphasise either response speed ('SPD') or accuracy ('ACC') on the upcoming trial. Speed and accuracy cues were randomly interleaved. On speed trials, participants had to respond within 700 ms to be eligible for a reward; on accuracy trials, participants had to respond within 1.5 s. After each choice, participants received two types of feedback: Firstly, the outcome of the choice (+0 or +100 points), and secondly, the actually obtained reward. If the participants responded in time (1.5 s in 'ACC' trials, 0.7 s in 'SPD' trials), their reward was equal to the outcome of the choice. If they responded too late, the participants were penalised with −100 points, irrespective of the outcome of the choice. The presentation of both the outcome of the choice and the actual reward allowed participants to both learn from the outcome of their choice as well as from their response timing.

In total, participants performed 342 trials divided over three runs. Each trial took 8.28 s (corresponding to 6 volumes; see below). Each run consisted of 3 new stimuli sets, that differed in their reward probabilities (80%/20%, 70%/30%, 60%/40%, respectively, for the three stimuli sets within each block) and therefore difficulty. Event timing was jittered to decorrelate the BOLD response design matrix, by pseudo-randomly sampling the duration of each fixation cross from 0.5 s, 1 s, 1.5 s and 2 s. Additionally, 10% null trials were included, during which the screen remained empty for 8.28 s.

Mixed effects models. We first tested for the effects of the SAT and difficulty manipulations on RT and accuracy using mixed effects models (e.g., 121). Linear models were used for RT, and generalised models with a binomial distribution for accuracy. In both models, difficulty (continuous) and cue (SAT, two levels) were included as both random and fixed effects. Their interaction was included as a fixed, but not random effect, since the maximal model did not converge. Degrees of freedom for the linear mixed effects model were estimated using Satterthwaite's method. We used the implementation in R packages 'lme4' and 'lmerTest' (122, 123).

Cognitive model specification. The behavioral data were modeled with the reinforcement learning advantage racing diffusion (RL-ARD) model (7), which is an instance of the broader class of combined reinforcement learning evidence accumulation models (RL-EAMs; 6). The RL-ARD conceptualises decision making as a race between accumulators, each accumulating evidence for one choice option. The first accumulator to reach a common threshold-level of evidence a triggers the motor processes that execute the decision. The time to respond equals the time to reach the threshold, plus an intercept $t0$ that corresponds to the time required for early perception encoding and response execution.

In the RL-ARD, each accumulator accumulates the *advantage* of one choice option over another option. Specifically, the rate of evidence accumulation (the drift rate v) of each accumulator depends on three terms: an evidence-independent base rate u (urgency); the advantage of one choice option of the other option, weighted by free parameter w_d; and the total amount of evidence, weighted by free parameters w_s. 'Evidence' in this model is based on Q-values, which represent the participant's internal belief about how rewarding each choice option is. For two-choice tasks such as in the present study, the drift rates for the two accumulators are:

$$\begin{aligned} v_{1-2} &= u + w_d(Q_1 - Q_2) + w_s(Q_1 + Q_2) \\ v_{2-1} &= u + w_d(Q_2 - Q_1) + w_s(Q_1 + Q_2) \end{aligned} \quad [1]$$

where Q_i is the Q-value for choice alternative i, which are updated after every trial according to a simple delta rule:

$$Q_{i,t+1} = Q_{i,t} + \alpha(r - Q_{i,t}) \quad [2]$$

where t is the trial number, r is the obtained reward (in this specific experimental paradigm, the 'outcome'), and α a free parameter known as the learning rate.

To model the effect of the SAT manipulation, we allowed both the V_0 and a parameters to vary freely between the speed and accuracy conditions, based on our earlier work (7).

In total, the RL-ARD has eight free parameters: two evidence-independent base rate u_{acc} and u_{spd}, weights on the difference and sum of the evidence w_d and w_s, non-decision time $t0$, learning rate α, and two thresholds a_{spd} and a_{acc}. Instead of estimating a u parameter for each condition separately, we estimated the across-condition mean u and difference $u_{spd-acc}$ parameters (hence, $u_{spd} = u + u_{spd-acc}$ and $u_{acc} = u - u_{spc-acc}$), and similarly, we estimated an across-condition mean a and difference $a_{spd-acc}$ parameter. The direct estimation of the between-condition differences in these parameters facilitates estimation of covariance with neural model parameters, which are detailed below.

MRI data acquisition. In multiple sessions, participants were scanned in a MAGNETOM 'Terra' 7 T MRI system (Siemens Healthineers, Germany) with a 32-channel phased array head coil (Nova Medical Inc, USA). The first session contained two anatomical scans: A multi-echo gradient recalled echo (GRE) and an MP2RAGE, both at 0.75 mm isotropic resolution. For the MP2RAGE, we used the following parameters: repetition time (TR) = 4.3 s, inversion times ($TI_{1,2}$) = 840 ms and 2370 ms, flip angles ($FA_{1,2}$) = 5° and 6°, echo time (TE) = 1.99 ms, field of view (FOV) = 240 × 240 × 168 mm, bandwidth = 250 Hz/px. For the GRE, the following parameters were used: TR = 31.0 ms, $TE_{1\text{-}4}$ = 2.51, 7.22, 14.44 and 23.23 ms, FA = 12°, FOV = 240 × 240 × 168 mm. In the remainder of this anatomical session, resting state data was collected that is not of interest for the current study.

The second session contained three functional runs with the task paradigm. A single echo echo planar imaging (EPI) sequence was used designed by the CMRR (https://www.cmrr.umn.edu/multiband/), with parameters based on our previous studies (48, 124) to tailor the sequence for the subcortex: 1.5 mm isotropic resolution, TE 14 ms, TR 1.38 s, partial Fourier 6/8, in-plane acceleration (GRAPPA) 3, multiband 2, bandwidth 1446 Hz/px, phase encoding direction A>P, FOV = 192 × 192 × 132 mm. In contrast to our previous work, we included a multiband factor of 2 in the protocol. Pilot testing indicated that on this MRI system, the increase in statistical power obtained through the increase in number of volumes (due to the lower TR with multiband acquisition) outweighed the loss in SNR (even in subcortical areas) for statistical testing purposes.

Each run consisted of 754 volumes (17 minutes 56 seconds). Immediately after each run, we collected 5 volumes of the same protocol with opposite phase encoding direction (P>A), which was used for susceptibility distortion correction purposes. Finally, at the end of the functional session, a low-resolution 1 mm MP2RAGE scan was acquired for co-registration purposes, using the same parameters as in the anatomical session.

During functional runs, physiological data on the participant's heart rate and respiration were acquired using a photoplethysmograph (with sampling frequency 200 Hz) and respiratory belt (with sampling frequency 50 Hz), respectively. In six runs (two in one participant, one in another participant, and three in a third participant), recording of physiological data failed due to technical reasons.

Anatomical masks. We used the multi-contrast anatomical subcortical structure parcellation (MASSP) algorithm (63) to obtain participant-specific anatomical masks of 17 subcortical structures. MASSP relies on multiple contrasts; here, we used quantified susceptibility (QSM) values, the longitudinal relaxation rates (R1), and effective transverse relaxation rates ($R2^*$). R1 values were computed based on the MP2RAGE data using a look-up table (125). $R2^*$ values were computed by least squares fitting of a mono-exponential decay function to the four echoes of the GRE data. QSM values were obtained using the phase maps of the last three echoes of the GRE data (126) with TGV-QSM (127). In both cases, LCPCA denoising (128) was performed beforehand on the 8 images of the GRE (4 magnitude and 4 phase). Prior to estimating $R2^*$ and QSM, the GRE data were brought into MP2RAGE-space by co-registration of the first GRE echo (magnitude image) to the second inversion of the MP2RAGE, using a rigid transformation in ANTs.

The MASSP algorithm combines shape, location, and R1, $R2^*$, QSM value priors to delineate the following 17 subcortical structures in an individual's data: Amygdala (Amg), claustrum (Cl), fornix (fx), the external and internal segments of the globus pallidus (GPe, GPi), internal capsule (ic), periaqueductal grey (PAG), pedunculopontine nucleus (PPN), red nucleus (RN), substantia nigra (SN), subthalamic nucleus (STN), striatum (Str), thalamus (Tha), ventral tegmental area (VTA), and the lateral, third, and fourth ventricles (LV, 3V, 4V). For all regions except fx, 3V and 4V, separate masks were obtained for both hemispheres. Here, we only focus on the gray matter structures, and thus excluded the internal capsule, fornix, and ventricles from the ROI analyses below; totalling 12 ROIs bilaterally.

Like in (129), we trained the MASSP algorithm on renormalised versions of the quantitative contrasts using a fuzzy C-means clustering of intensities, and linearly interpolating between cluster centroids. We also registered the data to the MASSP atlas in two successive steps. These alterations compared to the original MASSP implementation (63) led to small parcellation improvements for some structures.

To segment the thalamus into individual nuclei, we used the thalamic segmentation tool *segmentThalamicNuclei.sh* as part of *freesurfer* 7.2.0. The segmentation applies a probabilistic atlas that was built using a combination of *in vivo* and *ex vivo* data (65). The segmentation is performed in subject space with the T1w contrast after running the freesurfer pipeline (recon-all) as part of *fmriprep* (see below). The tool outputs discrete segmentations at a resolution of 0.5 mm, which were resampled to 1.5 mm resolution with linear interpolation.

fMRI preprocessing. Results included in this manuscript come from preprocessing performed using *fMRIPrep* 20.2.0 ((130, 131); RRID:SCR_016216), which is based on *Nipype* 1.5.1 ((132, 133); RRID:SCR_002502). For brevity in the main article, full details are included in the Supporting Information.

Neural model specification: Whole-brain generalised linear models (GLMs). The timeseries of the neural data were modeled using GLMs. In these GLMs, we modeled each voxel's timeseries y as:

$$\begin{aligned} y = {} & \beta_0 + \beta_{trial} x_{cue} + \beta_{spd-acc} x_{cue,spd-acc} + \\ & \beta_{trial} x_{response} + \beta_{left-right} x_{left-right} + \\ & \beta_{trial} x_{stimulus} + \beta_{\Delta value} x_{\Delta value} + \\ & \beta_{trial} x_{feedback} + \beta_{RPE} x_{RPE} + e \\ & e \sim \mathcal{N}(0, \sigma) \end{aligned} \quad [3]$$

where every β is a parameter to be estimated, x are the timeseries of the experiment events convolved with the canonical double-gamma haemodynamic response function (HRF; 134), and σ the residual variance. Note that we estimated a single β_{trial} parameter to account for the shared effects of the presentations of cues, stimuli, and feedback, as well as the effects of motor responses (e.g., the effects of visual processing and overall motor preparation). In experimental paradigm, the effects of these event types cannot be disentangled from one another due to their rapid succession within a trial. Note, however, that the contrasts of interest are orthogonal to these events and can be estimated well.

Mirroring the cognitive model, we estimated a between-cue difference for the BOLD responses relating to the cue. Specifically, the regressor $x_{cue,spd-acc}$ was also modeled on the onset of the cue but shows a negative deflection for 'ACC' cues and a positive deflection for 'SPD' cues. As such, the corresponding $\beta_{cue,spd-acc}$ reflects the difference in 'SPD' over 'ACC' cues. Similarly, the $\beta_{left-right}$ parameter reflects the BOLD-contrast resulting from left compared to right motor responses. The corresponding $x_{left-right}$ regressor was modeled on the onsets of the button presses.

The regressors $x_{stimulus}$ and $x_{\Delta value}$ relate to the stimulus and stimulus value differences, respectively. The amplitude of the stimulus value regressor varied parametrically across trials, with the trial-by-trial amplitude determined based on the difference in Q-values (internal value representations) as estimated by the RL-ARD model. Similarly, the regressors $x_{feedback}$ and x_{RPE} relate to the effects of the feedback and the reward prediction error, respectively, which were simulated obtained from the RL-ARD. Both the value difference and reward prediction error regressors were demeaned per run, to orthogonalise them with respect to the stimulus and feedback regressors. We included the temporal derivatives of all task regressors (note that these are not shown in Equation 3, but are included in Equation S1 in the Supporting Information).

As control analyses, we first fit the GLM using a traditional two-stage mass-univariate approach, where a GLM is fit per voxel. In this approach, we first fit the RL-ARD to the behavioral data, and extracted trial-by-trial regressors per subject by simulating them from the RL-ARD model. Specifically, the model was used to simulate the task paradigm for 100 times, each time with a different set of RL-ARD parameters (randomly sampled from the posterior distributions). On each trial of the simulation, the difference in values of the two stimuli was calculated, and the mean of the stimulus value differences at each trial across the 100 simulations was used to determine the regressor's amplitude. These stimulus value differences were then demeaned per run. The trial-by-trial height of the parametrically varying reward prediction error regressor was determined based on the same simulation of the RL-ARD (except now using the reward prediction error instead of the value differences), and this regressor was also demeaned per run.

To model physiological noise, we included a set of 18 regressors obtained using RETROICOR (135): 3th order phase Fourier expansion of the heart rate signal, 4nd order phase expansion of the respiration signal, and a 2nd order phase Fourier expansion of the interaction between heart rate and respiration (136). Two additional regressors were used to model heart rate variability (HRV; 137), and respiratory volume per time unit (RVT; 138, 139).

These physiological regressors were estimated using the PhysIO toolbox (140) implemented in the TAPAS software package (141). For six runs (one in a single participant, two in another participant, and three in a third participant), collection of the physiological data failed due to technical reasons. For these runs, the first 20 aCompCor components (142) were instead included in the design matrix. Additionally, for all participants 7 motion-related regressors were included (translation and rotation in three dimensions, plus the framewise displacement), and a set of discrete cosines to model low-frequency drifts. To model residual physiological noise, we also included a regressor with the mean signal within CSF, estimated by *fMRIprep*. Finally, we included a nuisance regressor to model the effect of response times using the *RTDur* approach (143). This regressor is generated by convolving a boxcar function (starting at the onset of each stimulus, with the response time on that trial as duration) with the same haemodynamic response function as was used for the task-related regressors.

Prior to fitting the whole-brain GLM, the data were minimally smoothed using SUSAN (144, kernel size FWHM = 1.5 mm). Run-level GLMs were estimated using FSL FEAT (145), and afterwards the three run-level GLMs per participant were combined with a fixed effects analysis. Group-level models were estimated using FSL FLAME1+2 (146). For the speed - accuracy cue contrast, the design matrix included both an intercept and two model-based parametrically varying parameters: the between-condition differences in the threshold parameter (speed - accuracy) and in the urgency parameter, which were z-scored. All group-level statistical parametric maps (SPMs) were corrected for the false discovery rate with the Benjamini-Hochberg procedure (FDR; $q < 0.05$). SPMs of the whole-brain results can be found in the Supporting Information.

Joint models. The main analysis used joint models. In joint models, the cognitive model (RL-ARD) and the neural model (GLM) are estimated simultaneously (50, 54–57). Furthermore, the joint models we employ assumed that the cognitive and neural parameters are multivariate normally distributed across subjects: $\theta, \beta \sim MVN(\delta, \Sigma)$. This assumption allows for estimation group-level mean parameters δ as well as correlations between parameters through the variance-covariance matrix Σ, and thereby allows for estimating which cognitive processes correlate with BOLD responses in which regions of interest (ROIs).

The variance-covariance matrix of a multivariate normal grows quadratically with the number of cognitive and neural parameters estimated. Therefore, we applied multiple restrictions to the participant-level models to retain feasibility of parameter estimation (147). Specifically, we made a distinction between estimating parameters *jointly* (i.e., both the group-level mean and the correlations between parameters of neural and cognitive models across individual) or *non-jointly* in which the group-level mean was estimated, but the no correlations were estimated.

Of the cognitive model, we estimated all parameters jointly. Figures 4-6 focus on only the parameters related to the SAT manipulation. Of the neural model, we estimated the $\beta_{cue,spd-acc}$, $\beta_{\Delta value}$, and β_{RPE} parameters of interest jointly, as well as the β_{CSF} and β_{RT} nuisance parameters. We estimated only these latter nuisance parameters jointly as we hypothesised these could most strongly correlate with parameters of interest. All other neural parameters (including the temporal derivatives and the standard deviation of the errors) were estimated non-jointly.

Joint models were fit to neural data from the ROIs defined by MASSP and by the thalamus atlas. To obtain the signal per ROI, first, the mean timeseries within each ROI defined by MASSP was extracted from the unsmoothed functional data. The mean timeseries were rescaled to percent signal change, through division by the mean signal, multiplying by 100 and subtracting 100. To reduce the total number of parameters in the joint models, we first filtered the timeseries and design matrix by least square regression of the same set of confounds as used in the whole-brain GLMs (except for the CSF and RT regressors, which were estimated in the joint model), to reduce physiological noise and remove low-frequency drifts from the signal.

Bayesian estimation. To allow for estimation of whole-brain general linear models (GLM) of the neural data, we first fit the cognitive model to the behavioral data only. All model estimations were performed using a Bayesian particle Metropolis-within-Gibbs (PMwG) sampler (148, 149). The PMwG sampler strictly adheres to a hierarchical model in which group-level parameters and participant-level parameters are estimated simultaneously. The group level is modeled with a multivariate Gaussian distribution, which is updated using Gibbs sampling. At the participant level, chains are updated using a combination of particle sampling and Metropolis-Hastings. We followed earlier work (149) by using four sampling stages. The first, pre-burn stage, was used to approximate the participant-level likelihood landscape for proposal distributions. The burn stage was run until the mean Gelman's diagnostic (150) was below 1.1. The adaptation stage was used to collect samples to generate a distribution that allows for efficient proposal samples in the last stage, the sampling stage. This sampling stage was run until convergence (assessed using Gelman's diagnostics and visual inspection of the chains). Samplers were run with three chains.

The priors on the group-level mean were Gaussian distributions. The mean and standard deviation of these priors of the cognitive model parameters were based on the posterior distributions described in (7), which used the same task and model (experiment 2). The prior for u was set to $N(2.5, 1)$, B to $N(1.5, 1)$, t_0 to $N(0.15, 1)$, w_d to $N(2.25, 1)$, w_s to $N(0.5, 1)$, and α to $N(0.12, 1)$. The $t0$, w_s, and w_d parameters were estimated on the log scale, and α on the probit scale. The prior for the contrasts of interest, $B_{spd-acc}$ and $u_{spd-acc}$, were set to $N(-0.5, 1)$ and $N(0.5, 1)$, respectively (note that threshold and urgency should have opposite signs to allow for faster responding under speed stress: thresholds should *decrease*, but urgency should *increase*). Visual comparisons confirmed that the posteriors were not strongly influenced by the priors for the parameters of interest.

For the group-level (co-)variance matrix we used an inverse-Gamma — inverse-Wishart mixture prior with 2 degrees of freedom and a scale parameter of 0.3. These settings give rise to uniform priors on the correlations (151), for parts of the group-level covariance matrix that were allowed to covary.

To visualise the quality of model fit, we randomly sampled 100 parameter sets from the posterior distributions, and used these to simulate the experimental design. These posterior predictive distributions were then used to calculate the credible intervals by taking the range between the 2.5% and 97.5% quantile of the averages across participants for each behavioral measure (RT quantiles and accuracy).

Next, we fit the joint models, in which we used the same priors for the cognitive models, except we decreased the variance of the group-level means to 0.7 for $B_{spd-acc}$ and $u_{spd-acc}$, and to 0.5 for the other parameters. The priors for the neural parameters were set to $N(0, 0.1)$, except for the RT nuisance parameter, which was set to $N(0, 0.001)$. Note that the amplitude of the RT nuisance regressor is much larger than the amplitudes of the other neural regressors, due to its duration being modeled (as opposed to using a stick function of 0.001 s). This also entails that the absolute parameter estimates are much smaller, hence, we also used a smaller variance for this parameter to stabilise estimation.

Joint models were implemented in a customised version of the EMC2 software package for R (152). The analysis scripts and data underlying this manuscript can be found at https://osf.io/pc5bm. A practical tutorial on joint modelling in this framework can be found in (153).

ACKNOWLEDGMENTS. We thank Pål Erik Goa for supporting this study by facilitating data acquisition. We thank Sarah Habli, Lisbeth Røe, and Daniel R. Sokołowski for their help collecting the data.

1. RS Sutton, AG Barto, *Reinforcement Learning: An Introduction.* (MIT press, Cambridge, MA), 2nd edition, (2018).
2. L Fontanesi, S Palminteri, M Lebreton, Decomposing the effects of context valence and feedback information on speed and accuracy during reinforcement learning: a meta-analytical approach using diffusion decision modeling. *Cogn. Affect. Behav. Neurosci.* **19**, 490–502 (2019).
3. L Fontanesi, S Gluth, MS Spektor, J Rieskamp, A reinforcement learning diffusion decision model for value-based decisions. *Psychon. Bull. & Rev.* (2019).
4. ML Pedersen, MJ Frank, G Biele, The drift diffusion model as the choice rule in reinforcement learning. *Psychon. Bull. & Rev.* **24**, 1234–1251 (2017).
5. ML Pedersen, MJ Frank, Simultaneous Hierarchical Bayesian Parameter Estimation for Reinforcement Learning and Drift Diffusion Models: a Tutorial and Links to Neural Data. *Comput. Brain & Behav.* (2020).
6. S Miletić, RJ Boag, BU Forstmann, Mutual benefits: Combining reinforcement learning with sequential sampling models. *Neuropsychologia* **136**, 107261 (2020).
7. S Miletić, et al., A new model of decision processing in instrumental learning tasks. *eLife* **10**, 1–33 (2021).
8. SD McDougle, AG Collins, Modeling the influence of working memory, reinforcement, and action uncertainty on reaction time and choice during instrumental learning. *Psychon. Bull. & Rev.* (2020).
9. BM Turner, Toward a Common Representational Framework for Adaptation. *Psychol. Rev.* (2019).
10. DK Sewell, HK Jach, RJ Boag, CA Van Heer, Combining error-driven models of associative learning with evidence accumulation models of decision-making. *Psychon. Bull. Rev.* (2019).
11. DK Sewell, A Stallman, Modeling the Effect of Speed Emphasis in Probabilistic Category Learning. *Comput. Brain & Behav.* **3**, 129–152 (2020).
12. B Wagner, D Mathar, J Peters, Gambling Environment Exposure Increases Temporal Discounting but Improves Model-Based Control in Regular Slot-Machine Gamblers. *Comput. Psychiatry* **6**, 142–165 (2022).
13. N Shahar, et al., Improving the reliability of model-based decision-making estimates in the two-stage decision task with reaction-times and drift-diffusion modeling. *PLoS Comput. Biol.* **15**, 1–25 (2019).
14. P Redgrave, T Prescott, K Gurney, The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience* **89**, 1009–1023 (1999).
15. J Deniau, G Chevalier, Disinhibition as a basic process in the expression of striatal functions. II. The striato-nigral influence on thalamocortical cells of the ventromedial thalamic nucleus. *Brain Res.* **334**, 227–233 (1985).
16. G Chevalier, S Vacher, J Deniau, M Desban, Disinhibition as a basic process in the expression of striatal functions. I. The striato-nigral influence on tecto-spinal/tecto-diencephalic neurons. *Brain Res.* **334**, 215–226 (1985).
17. A Nambu, et al., Excitatory Cortical Inputs to Pallidal Neurons Via the Subthalamic Nucleus in the Monkey. *J. Neurophysiol.* **84**, 289–300 (2000).
18. JW Mink, WT Thach, Basal ganglia intrinsic circuits and their role in behavior. *Curr. Opin. Neurobiol.* **3**, 950–957 (1993).
19. D Thura, P Cisek, The basal ganglia do not select reach targets but control the urgency of commitment. *Neuron* **95**, 1160–1170 (2017).
20. J Lauwereyns, K Watanabe, B Coe, O Hikosaka, A neural correlate of response bias in monkey caudate nucleus. *Nature* **418**, 413–417 (2002).
21. L Ding, JI Gold, Caudate encodes multiple computations for perceptual decisions. *J. Neurosci.* **30**, 15747–15759 (2010).
22. DP Munoz, RH Wurtz, Saccade-related activity in monkey superior colliculus I. Characteristics of burst and buildup cells. *J. Neurophysiol.* **73**, 2313–2333 (1995).
23. P Grimaldi, SH Cho, H Lau, MA Basso, Superior colliculus signals decisions rather than confidence: analysis of single neurons. *J. Neurophysiol.* **120**, 2614–2629 (2018).
24. EJ Jun, et al., Causal role for the primate superior colliculus in the computation of evidence for perceptual decisions. *Nat. Neurosci.* **24**, 1121–1131 (2021).
25. MN Shadlen, WT Newsome, Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. neurophysiology* **86**, 1916–36 (2001).
26. ME Mazurek, JD Roitman, J Ditterich, MN Shadlen, A Role for Neural Integrators in Perceptual Decision Making. *Cereb. Cortex* **13**, 1257–1269 (2003).
27. AC Huk, MN Shadlen, Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *J. Neurosci.* **25**, 10420–10436 (2005).
28. R Kiani, TD Hanks, MN Shadlen, Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J. Neurosci.* **28**, 3017–3029 (2008).
29. NA Steinemann, et al., Direct observation of the neural computations underlying a single decision (2024).
30. DP Hanes, JD Schall, Neural Control of Voluntary Movement Initiation. *Science* **274**, 427–430 (1996).
31. JD Schall, The neural selection and control of saccades by the frontal eye field. *Philos. Transactions Royal Soc. B: Biol. Sci.* **357**, 1073–1082 (2002).
32. BA Purcell, et al., Neurally constrained modeling of perceptual decision making. *Psychol. Rev.* **117**, 1113–1143 (2010).
33. JN Kim, MN Shadlen, Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat. neuroscience* **2**, 176–185 (1999).
34. L Ding, JI Gold, Neural correlates of perceptual decision making before, during, and after decision commitment in monkey frontal eye field. *Cereb. Cortex* **22**, 1052–1067 (2012).
35. P Cisek, JF Kalaska, Neural correlates of reaching decisions in dorsal premotor cortex: Specification of multiple direction choices and final selection of action. *Neuron* **45**, 801–814 (2005).
36. D Thura, P Cisek, Modulation of Premotor and Primary Motor Cortical Activity during Volitional Adjustments of Speed-Accuracy Trade-Offs. *J. Neurosci.* **36**, 938–956 (2016).
37. R Romo, A Hernández, A Zainos, Neuronal Correlates of a Perceptual Decision in Ventral Premotor Cortex. *Neuron* **41**, 165–173 (2004).
38. D Peixoto, et al., Decoding and perturbing decision states in real time. *Nature* **591**, 604–609 (2021).
39. W Schultz, Responses of midbrain dopamine neurons to behavioral trigger stimuli in the monkey. *J. Neurophysiol.* **56**, 1439–1461 (1986).
40. W Schultz, P Dayan, PR Montague, A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
41. PR Montague, P Dayan, TJ Sejnowski, A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* **16**, 1936–1947 (1996).
42. W Schultz, P Apicella, T Ljungberg, Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J. Neurosci.* **13**, 900–913 (1993).
43. SJ Gershman, et al., Explaining dopamine through prediction errors and beyond. *Nat. Neurosci.* **27**, 1645–1655 (2024).
44. S Miletic, "Modelling Structure and Function of the Human Subcortex," PhD thesis, University of Amsterdam (2023).
45. BU Forstmann, G De Hollander, L Van Maanen, A Alkemade, MC Keuken, Towards a mechanistic understanding of the human subcortex. *Nat. Rev. Neurosci.* **18**, 57–65 (2017).
46. MC Keuken, L Van Maanen, M Boswijk, BU Forstmann, M Steyvers, Large scale structure-function mappings of the human subcortex. *Sci. Reports* **8**, 15854 (2018).
47. G De Hollander, MC Keuken, W van der Zwaag, BU Forstmann, R Trampel, Comparing functional MRI protocols for small, iron-rich basal ganglia nuclei such as the subthalamic nucleus at 7 T and 3 T. *Hum. Brain Mapp.* **38**, 3226–3248 (2017).
48. S Miletić, et al., fMRI protocol optimization for simultaneously studying small subcortical and cortical areas at 7 T. *NeuroImage* **219** (2020).
49. S Miletić, et al., 7T functional MRI finds no evidence for distinct functional subregions in the subthalamic nucleus during a speeded decision-making task. *Cortex* (2022).
50. BM Turner, JJ Palestro, S Miletić, BU Forstmann, Advances in techniques for imposing reciprocity in brain-behavior relations. *Neurosci. Biobehav. Rev.* **102**, 327–336 (2019).
51. DY Teller, Linking propositions. *Vis. Res.* **24**, 1233–1246 (1984).
52. A Gelman, J Hill, *Data Analysis Using Regression and Multilevel/Hierarchical Models.* (Cambridge University Press, Cambridge), (2006).
53. C Spearman, The Proof and Measurement of Association between Two Things. *The Am. J. Psychol.* **15**, 72–101 (1904).
54. BM Turner, BU Forstmann, M Steyvers, *Joint Models of Neural and Behavioral Data*, Computational Approaches to Cognition and Perception. (Springer International Publishing), (2019).
55. BM Turner, T Wang, EC Merkle, Factor analysis linking functions for simultaneously modeling neural and behavioral data. *NeuroImage* **153**, 28–48 (2017).
56. BM Turner, CA Rodriguez, TM Norcia, SM McClure, M Steyvers, Why more is better: Simultaneous modeling of EEG, fMRI, and behavioral data. *NeuroImage* **128**, 96–115 (2016).
57. BM Turner, BU Forstmann, BC Love, TJ Palmeri, L Van Maanen, Approaches to analysis in model-based cognitive neuroscience. *J. Math. Psychol.* **76**, 65–79 (2017).
58. JM Groot, et al., A high-resolution 7 Tesla resting-state fMRI dataset optimized for studying the subcortex. *Data Brief* **55**, 110668 (2024).
59. JM Groot, et al., Echoes from Intrinsic Connectivity Networks in the Subcortex. *J. Neurosci.* **43**, 6609–6618 (2023).
60. SJS Isherwood, et al., Investigating Intra-Individual Networks of Response Inhibition and Interference Resolution using 7T MRI. *NeuroImage* p. 119988 (2023).
61. B Lloyd, et al., Subcortical nuclei of the human ascending arousal system encode anticipated reward but do not predict subsequent memory. *Cereb. Cortex* **35**, bhaf101 (2025).
62. AC Trutti, et al., Investigating working memory updating processes of the human subcortex using 7 Tesla fMRI. *eLife* **13** (2024).
63. PL Bazin, A Alkemade, MJ Mulder, AG Henry, BU Forstmann, Multi-contrast anatomical subcortical structures parcellation. *eLife* **9**, 1–23 (2020).
64. RA Rescorla, AR Wagner, A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Class. Cond. II Curr. Res. Theory* **21**, 64–99 (1972).
65. JE Iglesias, et al., A probabilistic atlas of the human thalamic nuclei combining ex vivo MRI and histology. *NeuroImage* **183**, 314–326 (2018).
66. SN Haber, B Knutson, The reward circuit: Linking primate anatomy and human imaging. *Neuropsychopharmacology* **35**, 4–26 (2010).
67. WM Pauli, RC O'Reilly, T Yarkoni, TD Wager, Regional specialization within the human striatum for diverse psychological functions. *Proc. Natl. Acad. Sci.* **113**, 1907–1912 (2016).
68. NLG del Rey, MÁ García-Cabezas, Cytology, architecture, development, and connections of the primate striatum: Hints for human pathology. *Neurobiol. Dis.* **176**, 105945 (2023).
69. PL Bazin, et al., Automated parcellation and atlasing of the human subcortex with ultra high-resolution quantitative mri. *Imaging Neurosci.* (2025).
70. L Van Maanen, et al., Neural Correlates of Trial-to-Trial Fluctuations in Response Caution. *J. Neurosci.* **31**, 17488–17495 (2011).
71. BU Forstmann, et al., Striatum and pre-SMA facilitate decision-making under time pressure. *Proc. Natl. Acad. Sci. United States Am.* **105**, 17538–17542 (2008).
72. D Thura, JF Cabana, A Feghaly, P Cisek, Integrated neural dynamics of sensorimotor decisions and actions. *PLOS Biol.* **20**, e3001861 (2022).
73. L Van Maanen, L Fontanesi, GE Hawkins, BU Forstmann, Striatal activation reflects urgency in perceptual decision making. *NeuroImage* **139**, 294–303 (2016).
74. Y Hung, SL Gaillard, P Yarmak, M Arsalidou, Dissociations of cognitive inhibition, response inhibition, and emotional interference: Voxelwise ALE meta-analyses of fMRI studies. *Hum. Brain Mapp.* **39**, 4065–4082 (2018).
75. EC Cieslik, VI Mueller, CR Eickhoff, R Langner, SB Eickhoff, Three key regions for supervisory attentional control: Evidence from neuroimaging meta-analyses. *Neurosci. &*

Biobehav. Rev. **48**, 22–34 (2015).
76. SJS Isherwood, MC Keuken, PL Bazin, BU Forstmann, Cortical and subcortical contributions to interference resolution and inhibition – An fMRI ALE meta-analysis. *Neurosci. & Biobehav. Rev.* **129**, 245–260 (2021).
77. PR Murphy, E Boonstra, S Nieuwenhuis, Global gain modulation generates time-dependent urgency during perceptual choice in humans. *Nat. Commun.* **7**, 13526 (2016).
78. PR Murphy, J Vandekerckhove, S Nieuwenhuis, Pupil-Linked Arousal Determines Variability in Perceptual Decision Making. *PLoS Comput. Biol.* **10**, e1003854 (2014).
79. J Motelow, H Blumenfeld, *Consciousness and Subcortical Arousal Systems*. (Elsevier Inc.), pp. 277–298 (2014).
80. AK Anderson, et al., Dissociated neural representations of intensity and valence in human olfaction. *Nat. Neurosci.* **6**, 196–202 (2003).
81. DM Small, et al., Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron* **39**, 701–711 (2003).
82. XB Li, T Inoue, S Nakagawa, T Koyama, Effect of mediodorsal thalamic nucleus lesion on contextual fear conditioning in rats. *Brain Res.* **1008**, 261–272 (2004).
83. S Zola-Morgan, LR Squire, Amnesia in monkeys after lesions of the mediodorsal nucleus of the thalamus. *Annals Neurol.* **17**, 558–564 (1985).
84. JP Aggleton, M Mishkin, Visual recognition impairment following medial thalamic lesions in monkeys. *Neuropsychologia* **21**, 189–197 (1983).
85. JP Aggleton, M Mishkin, Memory impairments following restricted medial thalamic lesions in monkeys. *Exp. Brain Res.* **52**, 199–209 (1983).
86. D Gaffan, A Parker, Mediodorsal thalamic function in scene memory in rhesus monkeys. *Brain* **123**, 816–827 (2000).
87. ND Child, EE Benarroch, Anterior nucleus of the thalamus. *Neurology* **81**, 1869–1876 (2013).
88. SJS Isherwood, et al., Multi-study fMRI outlooks on subcortical BOLD responses in the stop-signal paradigm. *eLife* **12** (2024).
89. G Weindel, L van Maanen, JP Borst, Trial-by-trial detection of cognitive events in neural time-series. *Imaging Neurosci.* **2**, 1–28 (2024).
90. DH Cerri, et al., Distinct neurochemical influences on fMRI response polarity in the striatum. *Nat. Commun.* **15**, 1916 (2024).
91. JP O'Doherty, The problem with value. *Neurosci. Biobehav. Rev.* **43**, 259–268 (2014).
92. MA Farries, TW Faust, A Mohebi, JD Berke, Selective encoding of reward predictions and prediction errors by globus pallidus subpopulations. *Curr. Biol.* **33**, 4124–4135.e5 (2023).
93. MG Kutlu, et al., Dopamine release in the nucleus accumbens core signals perceived saliency. *Curr. Biol.* **31**, 4748–4761.e8 (2021).
94. MG Kutlu, et al., Dopamine signaling in the nucleus accumbens core mediates latent inhibition. *Nat. Neurosci.* **25**, 1071–1081 (2022).
95. MB Madden, et al., A role for the claustrum in cognitive control. *Trends Cogn. Sci.* **26**, 1133–1152 (2022).
96. MPH Gardner, G Schoenbaum, SJ Gershman, Rethinking dopamine as generalized prediction error. *Proc. Royal Soc. B: Biol. Sci.* **285**, 20181645 (2018).
97. M Watabe-Uchida, N Eshel, N Uchida, Neural Circuitry of Reward Prediction Error. *Annu. Rev. Neurosci.* **40**, 373–394 (2017).
98. K D'Ardenne, SM McClure, LE Nystrom, JD Cohen, BOLD Responses Reflecting Dopaminergic Signals in the Human Ventral Tegmental Area. *Science* **319**, 1264–1267 (2008).
99. WM Pauli, et al., Distinct contributions of ventromedial and dorsolateral subregions of the human substantia nigra to appetitive and aversive learning. *J. Neurosci.* **35**, 14220–14233 (2015).
100. Y Zhang, KMH Larcher, B Misic, A Dagher, Anatomical and functional organization of the human substantia nigra and its connections. *eLife* **6**, 1–6 (2017).
101. TU Hauser, E Eldar, RJ Dolan, Separate mesocortical and mesolimbic pathways encode effort and reward learning signals. *Proc. Natl. Acad. Sci. United States Am.* **114**, E7395–E7404 (2017).
102. L Fontanesi, S Gluth, J Rieskamp, BU Forstmann, The role of dopaminergic nuclei in predicting and experiencing gains and losses: A 7T human fMRI study (2019).
103. J Garrison, B Erdeniz, J Done, Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies. *Neurosci. & Biobehav. Rev.* **37**, 1297–1310 (2013).
104. JB Goense, NK Logothetis, Neurophysiology of the BOLD fMRI Signal in Awake Monkeys. *Curr. Biol.* **18**, 631–640 (2008).
105. NK Logothetis, J Pauls, M Augath, T Trinath, A Oeltermann, Neurophysiological investigation of the basis of the fMRI signal. *Nature* **412**, 150–157 (2001).
106. CN Hall, C Howarth, Z Kurth-Nelson, A Mishra, Interpreting BOLD: Towards a dialogue between cognitive and cellular neuroscience. *Philos. Transactions Royal Soc. B: Biol. Sci.* **371**, 20150348 (2016).
107. T Lohrenz, KT Kishida, PR Montague, BOLD and its connection to dopamine release in human striatum: A cross-cohort comparison. *Philos. Transactions Royal Soc. B: Biol. Sci.* **371**, 20150352 (2016).
108. EA Ferenczi, et al., Prefrontal cortical regulation of brainwide circuit dynamics and reward-related behavior. *Science* **351**, aac9698 (2016).
109. EC Hirsch, AM Graybiel, Y Agid, Melanized dopaminergic neurons are differentially affected in Parkinson's disease. *Nature* **334**, 345—-348. (1988).
110. GF Koob, ND Volkow, Neurobiology of addiction: A neurocircuitry analysis. *The lancet. Psychiatry* **3**, 760–773 (2016).
111. AB Brühl, A Delsignore, K Komossa, S Weidt, Neuroimaging in social anxiety disorder—A meta-analytic review resulting in a new neurofunctional model. *Neurosci. & Biobehav. Rev.* **47**, 260–280 (2014).
112. NA Groenewold, et al., Volume of subcortical brain regions in social anxiety disorder: Mega-analytic results from 37 samples in the ENIGMA-Anxiety Working Group. *Mol. Psychiatry* **28**, 1079–1089 (2023).
113. JE LeDoux, DS Pine, Using Neuroscience to Help Understand Fear and Anxiety: A Two-System Framework. *Am. J. Psychiatry* **173**, 1083–1093 (2016).
114. MJ Frank, LC Seeberger, RC O'Reilly, By Carrot or by Stick: Cognitive Reinforcement Learning in Parkinsonism. *Science* **306**, 1940–1943 (2004).
115. AD Redish, Addiction as a Computational Process Gone Awry. *Science* **306**, 1944–1947 (2004).
116. DG Dillon, et al., Peril and Pleasure: An Rdoc-Inspired Examination of Threat Responses and Reward Processing in Anxiety and Depression. *Depress. Anxiety* **31**, 233–249 (2014).
117. C Luckhardt, et al., Reward processing in adolescents with social phobia and depression. *Clin. Neurophysiol.* **150**, 205–215 (2023).
118. AM Lozano, et al., Deep brain stimulation: current challenges and future directions. *Nat. Rev. Neurol.* **15**, 148–160 (2019).
119. JA Richey, et al., Common and distinct neural features of social and non-social reward processing in autism and social anxiety disorder. *Soc. Cogn. Affect. Neurosci.* **9**, 367–377 (2014).
120. A Ez-zizi, S Farrell, D Leslie, G Malhotra, CJ Ludwig, Reinforcement Learning Under Uncertainty: Expected Versus Unexpected Uncertainty and State Versus Reward Uncertainty. *Comput. Brain & Behav.* **6**, 626–650 (2023).
121. DJ Barr, R Levy, C Scheepers, HJ Tily, Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J. Mem. Lang.* **68**, 255–278 (2013).
122. A Kuznetsova, PB Brockhoff, RHB Christensen, lmerTest Package: Tests in Linear Mixed Effects Models. *J. Stat. Softw.* **82** (2017).
123. D Bates, M Mächler, B Bolker, S Walker, Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* **67** (2015).
124. G De Hollander, "Understanding the Human Subcortex Using Ultra-High Field MRI and Computational Cognitive Models," Doctoral dissertation (2017).
125. JP Marques, et al., MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. *NeuroImage* **49**, 1271–1281 (2010).
126. MWA Caan, et al., MP2RAGEME: T1, T2*, and QSM mapping in one sequence at 7 tesla. *Hum. Brain Mapp.* **40**, 1786–1798 (2019).
127. C Langkammer, et al., Fast quantitative susceptibility mapping using 3D EPI and total generalized variation. *NeuroImage* **111**, 622–630 (2015).
128. PL Bazin, et al., Denoising High-Field Multi-Dimensional MRI With Local Complex PCA. *Front. Neurosci.* **13**, 1–10 (2019).
129. S Miletić, et al., Charting human subcortical maturation across the adult lifespan with in vivo 7 T MRI. *NeuroImage* **249**, 118872 (2022).
130. O Esteban, et al., fMRIPrep (2018).
131. O Esteban, et al., fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nat. Methods* **16**, 111–116 (2019).
132. KJ Gorgolewski, et al., Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. *Front. Neuroinformatics* **5** (2011).
133. KJ Gorgolewski, et al., Nipype (2018).
134. GH Glover, Deconvolution of Impulse Response in Event-Related BOLD fMRI. *NeuroImage* **9**, 416–429 (1999).
135. GH Glover, Tq Li, D Ress, Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn. Reson. Medicine* **44**, 162–167 (2000).
136. AK Harvey, et al., Brainstem functional magnetic resonance imaging: Disentangling signal from physiological noise. *J. Magn. Reson. Imaging* **28**, 1337–1344 (2008).
137. C Chang, JP Cunningham, GH Glover, Influence of heart rate on the BOLD signal: The cardiac response function. *NeuroImage* **44**, 857–869 (2009).
138. SJ Harrison, et al., A Hilbert-based method for processing respiratory timeseries. *NeuroImage* **230**, 117787 (2021).
139. RM Birn, MA Smith, TB Jones, PA Bandettini, The respiration response function: The temporal dynamics of fMRI signal fluctuations related to changes in respiration. *NeuroImage* **40**, 644–654 (2008).
140. L Kasper, et al., The PhysIO Toolbox for Modeling Physiological Noise in fMRI Data. *J. Neurosci. Methods* **276**, 56–72 (2017).
141. S Frässle, et al., TAPAS: An Open-Source Software Package for Translational Neuromodeling and Computational Psychiatry. *Front. Psychiatry* **12**, 1–25 (2021).
142. Y Behzadi, K Restom, J Liau, TT Liu, A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage* **37**, 90–101 (2007).
143. JA Mumford, et al., The response time paradox in functional magnetic resonance imaging analyses. *Nat. Hum. Behav.* **8**, 349–360 (2024).
144. S Smith, J Brady, SUSAN—A New Approach to Low Level Image Processing. *Int. J. Comput. Vis.* **23**, 45–78 (1997).
145. MW Woolrich, BD Ripley, M Brady, SM Smith, Temporal autocorrelation in univariate linear modeling of FMRI data. *NeuroImage* **14**, 1370–1386 (2001).
146. MW Woolrich, TE Behrens, CF Beckmann, M Jenkinson, SM Smith, Multilevel linear modelling for FMRI group analysis using Bayesian inference. *NeuroImage* **21**, 1732–1747 (2004).
147. N Stevenson, et al., Using group level factor models to resolve high dimensionality in model-based sampling. *Psychol. Methods* (2024).
148. D Gunawan, GE Hawkins, MN Tran, R Kohn, SD Brown, New estimation approaches for the hierarchical Linear Ballistic Accumulator model. *J. Math. Psychol.* **96**, 102368 (2020).
149. N Stevenson, et al., Joint modelling of latent cognitive mechanisms shared across decision-making domains. *Comput. Brain & Behav.* (2024).
150. A Gelman, DB Rubin, Inference from Iterative Simulation Using Multiple Sequences. *Stat. Sci.* **7**, 457–472 (1992).
151. A Huang, MP Wand, Simple Marginally Noninformative Prior Distributions for Covariance Matrices. *Bayesian Analysis* **8**, 439–452 (2013).
152. N Stevenson, et al., EMC2: An R Package for cognitive models of choice (2025).
153. N Stevenson, S Miletic, B Forstmann, Bridging Brain and Behavior: A Step-by-Step Tutorial to Joint Modeling with fMRI (2025).