UNIVERSITY OF BIRMINGHAM
University of Birmingham Research at Birmingham

Multimodal imaging of brain connectivity reveals predictors of individual decision strategy in statistical learning

Karlaftis, Vasilis M.; Giorgio, Joseph; Vértes, Petra E; Wang, Rui; Shen, Yuan; Tino, Peter; Welchman, Andrew; Kourtzi, Zoe

DOI:
10.1038/s41562-018-0503-4

License:
None: All rights reserved

Document Version
Peer reviewed version

Citation for published version (Harvard):
Karlaftis, VM, Giorgio, J, Vértes, PE, Wang, R, Shen, Y, Tino, P, Welchman, A & Kourtzi, Z 2019, 'Multimodal imaging of brain connectivity reveals predictors of individual decision strategy in statistical learning', *Nature Human Behaviour*, vol. 3, no. 3, pp. 297-307. https://doi.org/10.1038/s41562-018-0503-4

Link to publication on Research at Birmingham portal

Publisher Rights Statement:
Checked for eligibility: 19/12/2018
Karlaftis, Vasileios Misak, et al. "Multimodal imaging of brain connectivity reveals predictors of individual decision strategy in statistical learning." Nature Human Behaviour (2018).
https://doi.org/10.1038/s41562-018-0503-4

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.
•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
•User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 15. May. 2024

Multimodal imaging of brain connectivity reveals predictors of individual decision strategy in statistical learning

Vasilis M Karlaftis[a], Joseph Giorgio[a], Petra E Vértes[b], Rui Wang[c], Yuan Shen[d], Peter Tino[e], Andrew Welchman[a], Zoe Kourtzi[a]*

[a] Department of Psychology, University of Cambridge, Cambridge, United Kingdom

[b] Department of Psychiatry, Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge, United Kingdom

[c] Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

[d] School of Science and Technology, Nottingham Trent University, Nottingham, United Kingdom

[e] School of Computer Science, University of Birmingham, Birmingham, United Kingdom

Correspondence:
Zoe Kourtzi
Department of Psychology
University of Cambridge
Cambridge, UK
Email: zk240@cam.ac.uk

Text: 6,939 words

Methods: 1,050 words

Figure legends: less than 300 words

References: 52

Display items: 7

Abstract

Successful human behavior depends on the brain's ability to extract meaningful structure from information streams and make predictions about future events. Individuals can differ markedly in the decision strategies they use to learn the environment's statistics, yet we have little idea why. Here, we investigate whether the brain networks involved in learning temporal sequences without explicit reward differ depending on the decision strategy that individuals adopt. We demonstrate that individuals alter their decision strategy in response to changes in temporal statistics and engage dissociable circuits: extracting the exact sequence statistics relates to plasticity in motor cortico-striatal circuits, while selecting the most probable outcomes relates to plasticity in visual, motivational and executive cortico-striatal circuits. Combining graph metrics of functional and structural connectivity, we provide evidence that learning-dependent changes in these circuits predict individual decision strategy. Our findings propose brain plasticity mechanisms that mediate individual ability for interpreting the structure of variable environments.

Learning and experience are known to facilitate our ability to extract meaningful structure from streams of information and interpret complex environments. Despite the general consensus that 'practice makes perfect', there is striking variability among individuals in the extent to which they take advantage of past experience. In the laboratory, this variability has been demonstrated in tasks such as perceptual decision making[1,2] or statistical learning of regularities (i.e. learning of probabilistic spatial or temporal structures) through mere exposure to the environment[3,4]. Previous work examining individual variability in decision making and probabilistic learning tasks, has highlighted the role of individual decision strategies[5–10]. In particular, humans and animals have been shown to engage in probability matching or maximization when making choices in probabilistic environments (e.g.[9,11,12]). Probability matching involves making choices stochastically to match the probabilistic distribution of all possible outcomes, while probability maximization involves choosing the most probable or frequently rewarded outcome in a given context.

Individual variability in these decision strategies has mainly been investigated in the context of reward learning (e.g.[9,11,12]). Yet, reward-based learning captures only one aspect of human flexibility in natural environments, as feedback and rewards are often not explicit. Here, we test the role of decision strategies in statistical learning. In particular, we designed a statistical learning task that tests whether individuals learn to extract temporal structure from mere exposure to unfamiliar sequences without explicit reward (i.e. trial-by-trial feedback). We changed the temporal sequence statistics unbeknownst to the participants, to simulate structure in natural environments that may vary from simple regularities to more complex probabilistic combinations. That is, participants were first exposed to sequences determined by frequency statistics (i.e. one item in the sequence occurred more frequently than others) and then sequences that were determined by context-based statistics (i.e. some item combinations were more frequent than others). Participants predicted which item would

appear next in the sequence. We modeled the participant responses to interrogate the decision strategy that individuals adopt during learning (i.e. how individuals extract temporal structure). We reasoned that individuals would adapt their decision strategies in response to changes in the temporal sequence statistics and the learning goal (i.e. learning frequency vs. context-based statistics).

Previous work has implicated cortico-striatal circuits in sequence and probabilistic learning[13–16]. Here, we sought to determine whether these circuits are involved in statistical learning of temporal structures without explicit reward. We ask whether individual decision strategies (from matching to maximization) involve distinct cortico-striatal circuits and whether learning-dependent plasticity in these circuits can account for individual variability in learning to extract the environment's statistics. We reasoned that brain plasticity, as expressed by learning-dependent connectivity changes in cortico-striatal circuits, would predict changes in decision strategy when learning frequency vs. context-based statistics.

To test these hypotheses, we combined our statistical learning task with multi-session (before vs. after training) measurements of functional (resting-state fMRI: rs-fMRI) and structural (Diffusion Tensor Imaging: DTI) connectivity. rs-fMRI has been shown to reveal functional connectivity within and across brain networks that subserve task performance[17,18]. Moreover, there is accumulating evidence for changes in both functional and structural brain connectivity due to training (e.g. for reviews[19,20]), suggesting learning-dependent plasticity in human brain networks that mediate adaptive behavior. To map cortico-striatal circuits at fine scale we employed DTI-based segmentation analysis[21] of the striatum into finer sub-regions and computed the functional connectivity between these striatal regions and cortical networks, as revealed by analysis of the rs-fMRI data. Our results show that individuals adapt their decision strategies (from matching towards maximization) in response to changes in the temporal statistics. These adaptive decision strategies relate to distinct cortico-striatal circuits

for learning temporal statistics. That is, adopting a strategy closer to matching when learning frequency statistics relates to learning-dependent connectivity changes in the motor circuit. In contrast, deviating from matching towards maximization when learning context-based statistics relates to functional connectivity changes in the visual cortico-striatal circuit.

We next combined graph theory analysis with a multivariate statistical analysis (Partial Least Squares-PLS regression) to determine multimodal predictors of decision strategy. This approach allows us to a) combine information from multivariate signals (rs-fMRI, DTI)– rather than using data from each MRI modality alone, b) test whether plasticity in functional and/or structural connectivity in cortico-striatal circuits predicts– rather than simply relates to– individual decision strategy. In particular, we employed graph theory to extract metrics of brain connectivity that are comparable across brain imaging modalities and have been suggested to relate to learning and brain plasticity[22,23]. We then used PLS modeling to combine these multimodal graph metrics and identify brain connectivity predictors (rs-fMRI, DTI) of individual decision strategy when learning temporal statistics. Our results demonstrate that learning-dependent changes in resting cortico-striatal connectivity (functional and structural) that predict individual decision strategy for statistical learning. In particular, we discern distinct brain plasticity mechanisms that predict: a) changes in individual decision strategy in response to changes in the environment's statistics, b) individual variability in decision strategy independent of temporal statistics. Our findings provide evidence for adaptive decision strategies that involve distinct brain routes for statistical learning, proposing a strong link between learning-dependent plasticity in brain connectivity and individual learning ability.

Results

Behavioral improvement with statistical learning

To investigate learning of temporal structures, we generated temporal sequences of different Markov orders (i.e. level-0, level-1 and level-2: context lengths of 0, 1 or 2 previous items, respectively) (**Figure 1a, 1b**). We simulated event structures that typically vary in their complexity in natural environments by exposing participants to sequences of unfamiliar symbols that increased in context length unbeknownst to the participants. That is, participants were first trained on sequences determined by frequency statistics (i.e. level-0: occurrence probability per symbol) and then on sequences determined by context-based statistics (i.e. level-1 and level-2: the probability of the next symbol depends on the preceding symbol(s)). Participants were asked to predict which symbol they expected to appear next in the sequence. Participants were not given trial-by-trial feedback, consistent with statistical learning paradigms.

Figure 1

We quantified participants' performance in this prediction task by measuring how closely the probability distribution of the participant responses matched the distribution of the presented symbols[10]. This performance index (PI, see Supplementary Information) is preferable to a simple measure of accuracy as the probabilistic nature of the sequences means that the 'correct' upcoming symbol is not uniquely specified.

We then computed a normalized performance index by subtracting performance for random guessing. Comparing normalized PI across sessions and levels (two-way repeated measures ANOVA with Session (Pre, Post) and Level (level-0, level-1, level-2)) showed a significant main effect of Session ($F(1,20)=117.9$, $p<0.001$, $\eta_p^2=0.855$) and Level ($F(2,40)=17.9$, $p<0.001$, $\eta_p^2=0.473$), but no significant interaction between Session and Level ($F(1.44,28.71)=2.7$, $p=0.098$, $\eta_p^2=0.120$, Greenhouse-Geisser corrected), suggesting that participants improved significantly after training and showed similar improvement across levels (**Figure 2a**).

Decision strategies for learning: from matching to maximization

Previous work on probabilistic learning[8–10] and decision-making in the context of sensorimotor tasks[5–7] has shown that individuals adopt decision strategies (from matching to maximization) when making probabilistic choice. Here, we test the role of these decision strategies in statistical learning (i.e. without explicit feedback or reward). In our statistical learning task, participants were exposed to stochastic sequences and therefore needed to learn the probabilities of different outcomes. Modeling the participants' responses allows us to quantify their decision strategy, reflecting how the participants extract and respond to context-target contingencies in probabilistic sequences. In particular, participants may adopt: a) probability matching; that is, match their choices to the relative probabilities of the context-target contingencies presented in the sequences, or b) deviate from matching towards maximization; that is, choose the most probable outcome in a given context.

We quantified participant's decision strategy during training by comparing individual participant responses to two models: (i) a probability matching model, where probabilistic distributions of possible outcomes were derived from the Markov models that generated the presented sequences, and (ii) a probability maximization model, where only the most likely outcome is allowed for each context. We quantified each participant's strategy choice during training based on the distance of the participant response distribution from the matching and maximization model. We then computed a single measure of strategy index as the integral between the participant's strategy choice and the matching model across trials and training blocks. Therefore, strategy index is a continuous measure that captures the strategy that individuals adopt over time (i.e. during training) on a continuous scale between matching and maximization (**Figure 2b, Supplementary Figure 1, Supplementary Figure 2**). Zero strategy index indicates that the participant response distribution matches the probability distribution of the presented sequence (i.e. exact matching). Participant's performance

deviating from the matching model may result to a positive or negative strategy index. Overestimating the probability of the most probable context-target contingency in the sequence results in a positive strategy index indicating that the participant's strategy ranges between matching and maximization. In contrast, underestimating the probability of the most probable context-target contingency in the sequence results in a negative strategy index indicating that the participant's strategy ranges between matching and a random model of response (i.e. participants choose all context-target contingencies with equal probability). Thus, we interpret strategy index values close to zero as strategy closer to matching; while higher positive values as strategy deviating from matching towards maximization.

Figure 2b, c shows differences in strategy index across sequence levels and individual participants. A one-way repeated measures ANOVA with Level (level-0, level-1, level-2) showed a significant main effect of Level ($F(1.44,28.79)=8.0$, $p=0.004$, $\eta_p^2=0.286$, Greenhouse-Geisser corrected), indicating higher strategy index for increasing context length. In particular, strategy index for level-1 was higher than strategy index for level-0 ($t(19)=2.5$, $p=0.020$, CI=[0.03, 0.30], Cohen's $d=0.567$), but not for level-2 compared to level-1 ($t(19)=1.9$, $p=0.066$, CI=[-0.01, 0.13], Cohen's $d=0.435$). Further, the strategy indexes for level-1 and level-2 were highly correlated ($r(19)=0.72$, $p<0.001$, CI=[0.42, 0.89]), while no significant correlations were found for level-0 (level-0 vs. level-1: $r(19)=-0.21$, $p=0.35$, CI=[-0.71, 0.28]; level-0 vs. level-2: $r(19)=-0.15$, $p=0.52$, CI=[-0.55, 0.34]). To avoid collinearity[24], we computed a mean strategy index for level-1 and level-2 to generate a single predictor of learning context-based statistics for further regression analyses. This mean strategy index for context-based statistics was significantly higher than the strategy index for frequency statistics ($t(19)=3.2$, $p=0.005$, CI=[0.07, 0.32], Cohen's $d=0.711$). Further, the strategy index for frequency statistics was not significantly different from matching (i.e. zero strategy index; one sample t-test: $t(20)=-0.23$, $p=0.82$, CI=[-0.08, 0.07], Cohen's $d=-0.050$).

In contrast, the strategy index for context-based statistics was significantly higher than zero (one sample t-test: $t(20)=4.01$, $p<0.001$, CI=[0.08, 0.26], Cohen's $d=0.874$). Taken together, these results provide evidence that participants adapted their decision strategy in response to changes in temporal statistics across sequence levels; that is, individuals adopted a strategy that deviated from matching towards maximization for learning first frequency and then context-based statistics.

These differences in decision strategy across sequence levels could not be simply explained by changes in reward processing, cognitive strategy training or differences in performance improvement across sequence levels. Specifically, the participants were not given explicit reward (i.e. no trial-by-trial feedback) or explicitly trained on effective cognitive strategies to boost task performance. Further, there were no significant differences in performance index across levels after training (see *Learning frequency and context-based statistics*) and participant performance after training did not correlate significantly with decision strategy (level-0: $r(19)=0.21$, $p=0.36$, CI=[-0.21, 0.58]; level-1: $r(19)=0.06$, $p=0.81$, CI=[-0.37, 0.42]; level-2: $r(19)=0.15$, $p=0.52$, CI=[-0.37, 0.52]). In contrast, we have previously shown that individual decision strategy is positively correlated with learning rate (i.e. how fast participants extract the correct sequence structure) in our statistical learning task[10]. Taken together, these results suggest that the adaptive decision strategies we observed in response to changes in temporal statistics reflect changes in the learning process (i.e. how individuals extract temporal sequence structure) rather than overall changes in task training.

Figure 2

Learning-dependent changes in DTI-informed resting-state connectivity

Previous work has established distinct cortico-striatal circuits with dissociable functions[25] that have been implicated in a range of learning tasks, including sequence and probabilistic learning[13–15]. Here, we investigated whether brain plasticity in these cortico-striatal circuits

relate to individual decision strategy in statistical learning (i.e. without trial-by-trial feedback). In particular, to determine functional connectivity at rest we used: a) DTI-based segmentation to define striatal regions and b) ICA-based decomposition of the rs-fMRI timecourse to define functional cortical networks.

First, we used DTI data to segment the striatum into finer sub-regions that will then serve as regions of interest for the functional connectivity analysis of the rs-fMRI data (see Supplementary Information). In particular, we defined striatum (i.e. caudate and putamen) anatomically from the Automated Anatomical Labeling (AAL) atlas[26] and segmented it into sub-regions based on their structural connectivity profile (**Supplementary Figure 3**). We derived four segments per hemisphere that corresponded to a) ventral striatum, b) head of caudate and anterior putamen, c) body and tail of caudate, and d) posterior putamen (**Figure 3a, Supplementary Table 1**). This segmentation is in agreement with previous histological studies[25].

We then identified functional brain networks during rest by decomposing the rs-fMRI timecourse into functionally connected components (i.e. components comprising voxel clusters with correlated timecourse) using Group Independent Component Analysis (GICA, see Supplementary Information). We followed the standard pipeline to perform the pre-processing on the rs-fMRI data for GICA (see Supplementary Information). Following GICA, we selected components associated with known cortico-striatal circuits that have been implicated in learning[25] (**Figure 3b, Supplementary Table 2**): a) Right Central Executive (CP_9, peak activations in right middle frontal gyrus and right inferior parietal lobule), b) Left Central Executive (CP_14, peak activations in left inferior frontal gyrus and left inferior parietal lobule), c) Sensorimotor (CP_4, peak activations in bilateral supplementary motor area), d) Lateral Motor (CP_5, peak activations in bilateral postcentral gyrus), e) Secondary Visual (CP_2, peak activations in bilateral middle occipital gyrus), f) Early Visual (CP_12,

peak activations in bilateral calcarine sulcus), and g) Anterior Cingulate (CP_15, peak activations in bilateral anterior cingulate).

We next tested whether learning-dependent changes in intrinsic and extrinsic functional connectivity within cortico-striatal circuits (i.e. between DTI-defined striatal segments and ICA-defined cortical components) relate to individual decision strategy. As strategy index is a continuous measure of decision strategy, we correlated changes in functional connectivity with individual strategy index rather than comparing between separate groups of participants (i.e. matchers vs. maximizers). Positive correlations indicate that higher increase in connectivity after training relates to maximization (top-right quadrant of the correlation plots), whereas negative correlations indicate that higher increase in connectivity relates to matching (top-left quadrant of the correlation plots).

Figure 3

Correlating intrinsic connectivity with strategy

Intrinsic connectivity is a measure of signal coherence within a local network and quantifies activity correlation across voxels within the network. Previous work has shown that functional networks during task and rest are highly similar[27], suggesting that task-related BOLD activity relates to intrinsic connectivity at rest. Further, variability in intrinsic connectivity has been suggested to explain task performance[28]. Here, we ask whether learning-dependent changes in intrinsic connectivity within each cortical network relate to individual decision strategy when learning temporal statistics.

We calculated an intrinsic connectivity measure for each cortical network indicating its local connectivity strength (N=7). We then correlated intrinsic connectivity change (Post minus Pre) with strategy for frequency and context-based statistics (**Supplementary Table 3a**). For frequency statistics, learning-dependent changes in connectivity in the Lateral Motor network correlated positively with strategy index ($r(19)=0.77$, $p<0.001$, CI=[0.60, 0.89],

surviving False Coverage Rate-FCR correction) (**Figure 4a**). For context-based statistics, learning-dependent changes in connectivity in the Secondary Visual network correlated negatively with strategy index (r(19)=-0.49, p=0.025, CI=[-0.74, -0.10]) (**Figure 4a**). In contrast, we observed positive (marginally significant) correlations of learning-dependent changes in connectivity in the Left Central Executive (LCEN) and Anterior Cingulate (ACC) networks with strategy index (LCEN: r(19)=0.42, p=0.059, CI=[0.01, 0.68]; ACC: r(19)=0.35, p=0.121, CI=[0.04, 0.63]) (**Supplementary Figure 4**).

Correlating extrinsic connectivity with strategy

Extrinsic connectivity is a measure of functional connectivity between brain regions. In particular, extrinsic connectivity is computed as the correlation of the brain signals in– typically distant– regions across time and quantifies the coherence of their activity[17,29]. Previous work suggests that extrinsic connectivity changes with training and relates to behavioral performance[19]. Here, we test whether learning-dependent changes in cortico-striatal extrinsic connectivity relate to individual decision strategy.

We selected pairs of striatal (**Figure 3a, Supplementary Table 1**) and cortical areas (**Figure 3b**, **Supplementary Table 2**) based on known cortico-striatal circuits[25] (N=14): a) motivational: ventral striatum to ACC, b) executive: caudate head and anterior putamen to RCEN and LCEN (i.e. dorsolateral prefrontal and parietal cortex), c) visual: caudate body and tail to Secondary Visual and Early Visual networks, and d) motor: posterior putamen to Sensorimotor and Lateral Motor networks (**Supplementary Table 3b**). These pathways have been identified by previous functional[30,31] and structural connectivity[32,33] studies. We calculated the Pearson correlation between the timecourses in these cortico-striatal areas, as a measure of extrinsic functional connectivity. We then correlated connectivity change (Post minus Pre, after Fisher z-transform) with the strategy index for frequency and context-based statistics. For learning frequency statistics, learning-dependent changes in connectivity

between the right posterior putamen and the Lateral Motor network (r(19)=0.51, p=0.018, CI=[0.20, 0.74], surviving FCR correction) correlated positively with strategy index (**Figure 4b**). In contrast, for context-based statistics, learning-dependent changes in connectivity between the left body/tail of caudate and the Early Visual network (r(19)=-0.46, p=0.034, CI=[-0.83, -0.13], surviving FCR correction) correlated negatively with strategy index (**Figure 4b**).

Figure 4

Relating adaptive decision strategies to brain plasticity

Taken together, our results provide evidence that plasticity in distinct cortico-striatal circuits– as expressed by changes in intrinsic and extrinsic connectivity– relates to adaptive decision strategies when learning temporal statistics. We interpret this brain plasticity in the context of our behavioral findings showing that participants adapted their strategy from matching towards maximization when learning first frequency and then context-based statistics.

Our results showed that matching when learning frequency statistics relates to decreased intrinsic connectivity within the Lateral Motor network and decreased extrinsic connectivity between this network and posterior putamen. Previous work has implicated the motor circuit in habitual learning[34,35] and stimulus-response associations[36]. Thus, decreased connectivity in this circuit may facilitate matching that involves learning the exact sequence statistics rather than reinforcing habitual responses.

In contrast, deviating from matching towards maximization when learning context-based statistics relates to decreased connectivity within the visual cortico-striatal circuit (intrinsic connectivity in Secondary Visual network, extrinsic connectivity between body/tail of caudate and the Early Visual network). Previous work has implicated the visual cortico-striatal circuit in learning predictive associations[16] and decision making[37,38], highlighting its role in higher cognitive functions rather than simply processing of low-level sensory

information. Thus, decreased connectivity in this circuit may facilitate selecting the most probable outcome when learning complex context-target contingencies rather than learning the exact probability distributions.

Multimodal predictors of decision strategy

Our results so far provide evidence that learning-dependent changes in resting functional connectivity relate to adaptive changes in decision strategies. Next, we test whether learning-dependent plasticity in both functional and structural connectivity in these circuits predicts individual decision strategy, extending beyond the univariate and correlational approach we followed for our rs-fMRI connectivity analysis.

To combine data from rs-fMRI and DTI, we employed graph theory that allows us to extract comparable metrics across participants and brain imaging modalities using the same topological brain structure (e.g. AAL parcellation). In particular, we constructed participant-specific whole-brain binary graphs for each brain imaging modality (rs-fMRI, DTI). We then selected twelve nodes from these graphs per imaging modality corresponding to the cortico-striatal circuits in the rs-fMRI analysis (**Figure 3b, Figure 4**): a) striatum: bilateral caudate, bilateral putamen; b) RCEN network: right middle frontal gyrus (MFG); c) LCEN network: triangular part of left inferior frontal gyrus (IFG); d) Lateral Motor network: bilateral postcentral gyrus; e) Early Visual network: bilateral calcarine sulcus; and f) ACC network: bilateral anterior cingulate gyrus (ACC) (**Figure 5a, b**).

For each selected node, we computed a measure of global and local integration. In networks, global integration describes the extent to which nodes integrate information from the whole graph. Different metrics have been used to quantify global integration; for example, regions with high global integration may have many connections to the rest of the brain (i.e. high degree) or have fast routes to all other brain regions (i.e. low path length). Here, we focus on nodal degree (i.e. number of a node's connections to the whole brain), as high

degree nodes (also known as hubs) have been shown to play a key role in learning (e.g. for review[39]). In contrast, local integration quantifies the regional organization of a graph; for example, modules are defined as brain nodes that are highly connected with each other but less strongly to the rest of the brain, therefore forming a community[40]. Here, we focus on clustering coefficient which measures the proportion of a node's first neighbors that are also connected to one another[41]. Both degree and clustering coefficient have been previously shown to relate to learning and brain plasticity[22,23].

Figure 5

We next asked whether learning-dependent changes in the local and global integration of cortico-striatal networks predict variability in decision strategy across sequence levels (i.e. frequency vs. context-based statistics) and individuals. To identify the linear combinations of regional metrics of functional and structural brain connectivity that best predict individual strategy, we entered into a PLS regression model the difference in rs-fMRI and DTI graph metrics (degree, clustering coefficient) before vs. after training (i.e. post- minus pre-training values for degree and clustering coefficient). PLS regression[42] is a statistical method that is used to relate a set of predictors to a set of response variables. That is, PLS identifies a set of independent components from the predictors (i.e. linear combinations of the rs-fMRI and DTI graph metrics) that show strongest association (i.e. maximum covariance) with the response variables of interest (i.e. strategy index for frequency and context-based statistics)[42]. This statistical method has been previously used in neuroimaging studies[43,44] with multi-collinear predictors or high data dimensionality (i.e. the number of predictors exceeds the number of samples). We followed this methodology to combine nodal graph metrics derived from rs-fMRI and DTI data and identify predictors of strategy, as the number of predictors exceeds our sample size (i.e. 48 predictors, 21 participants).

We found that the first three PLS components (PLS-1, PLS-2, PLS-3) predicted significantly the strategy index for frequency and context-based statistics compared to a null model ($p=0.024$ for 10,000 permutations). These three components together explained 85% of the variance in strategy index (**Supplementary Figure 5**). For further analysis, we focused on the first two components (**Supplementary Table 4**), as they were robustly estimated across a range of density levels (10% to 30% density; **Supplementary Figure 6**) and two additional atlases (Shen and Brainnetome atlases) (see Supplementary Information). **Figure 6a, b** summarizes the weights (combinations of nodes and metrics) for PLS-1 and PLS-2 at 20% density ($|z|>2.576$ indicates significant predictors ($p=0.01$)[42]).

Figure 6

Our analyses showed that these PLS components predict: a) differences in decision strategy across sequence levels (i.e. frequency vs. context-based statistics) and b) differences in decision strategy across individuals independent of sequence statistics. **Figure 7a** shows that PLS-1 dissociates strategy across sequence levels; that is, a negative weight is assigned for frequency statistics vs. a positive weight for context-based statistics (i.e. the two strategies are separated by the $y=0$ axis). In contrast, PLS-2 predicts individual variability in strategy independent of the sequence statistics; that is, positive weights are assigned for both frequency and context-based statistics (**Figure 7a**).

To further quantify these findings, we computed two complementary indexes. First, we calculated a strategy difference index, by subtracting strategy index for frequency statistics from the strategy index for context-based statistics (i.e. higher values indicate strategy closer to maximization for context-based than frequency statistics). Second, we calculated a mean strategy index, by averaging the strategy index for frequency and context-based statistics (i.e. higher values indicate strategy closer to maximization across sequence levels). We found that PLS-1 correlates positively with the strategy difference index

(r(19)=0.89, p<0.001, CI=[0.68, 0.96]) but not with the mean strategy index (r(19)=0.18, p=0.44, CI=[-0.27, 0.51]), suggesting that this component captures learning-dependent changes in brain connectivity that predict changes in strategy in response to changes in the sequence statistics (**Figure 7b**). In contrast, PLS-2 correlates positively with the mean strategy index (r(19)=0.79, p<0.001, CI=[0.49, 0.92]) but not with the strategy difference index (r(19)=0.13, p=0.58, CI=[-0.25, 0.48]), suggesting that this component captures learning-dependent changes in brain connectivity that predict variability in decision strategy across individuals independent of the sequence structure (**Figure 7b**). **Supplementary Figure 7** provides a complementary illustration of the relationship between each PLS component (PLS-1, PLS-2) and decision strategy for frequency vs. context-based statistics.

Figure 7c summarizes the brain nodes that correspond to significant predictors (|z|>2.576, p=0.01[42]) for PLS-1 and PLS-2 across imaging modalities (rs-fMRI, DTI) and graph metrics (degree change, clustering coefficient change). For PLS-1, the brain metrics that significantly predict change in decision strategy in response to changes in the sequence statistics include: a) degree change in left putamen (DTI), right calcarine (DTI) and left IFG (rs-fMRI); b) clustering change in left postcentral (DTI) and right ACC (DTI) (**Figure 7c, Supplementary Table 4a**). That is, global integration in the visual and left executive circuits, while local integration within the motor and motivational circuits predict changes in decision strategy in response to changes in sequence structure (i.e. learning frequency vs. context-based statistics), as indicated by the positive correlation of PLS-1 with the strategy difference index **(Figure 7b).** In contrast, for PLS-2, the brain metrics that significantly predict individual variability in decision strategy independent of the temporal statistics include: a) degree change in left ACC (DTI), bilateral caudate (DTI) and right MFG (DTI); b) clustering change in left caudate (DTI) and left ACC (rs-fMRI) (**Figure 7c, Supplementary Table 4a**). Therefore, global integration in the motivational and right executive circuits, while local

integration within the motivational circuit support learning by maximizing, as indicated by the positive correlation of PLS-2 with the mean strategy index **(Figure 7b)**.

These results showing that graph metrics in the visual and motor cortico-striatal circuits predict decision strategy are consistent with our previous correlational analyses (**Figure 4**), suggesting that learning-dependent plasticity in these circuits may facilitate switching from matching towards maximization for learning more complex context-based statistics. Further, the multivariate treatment of the data afforded by the PLS analysis supports the role of regions in motivational and executive cortico-striatal circuits in decision strategy, corroborating our correlational analyses that showed marginal effects for these regions (**Supplementary Figure 4**). These findings are consistent with previous work implicating the motivational circuit in goal-directed actions[34,45] and individual strategy choice[35], while the executive circuit in updating task rules[46,47].

Figure 7

Finally, our findings generalized to other graph metrics that relate to global and local integration (see Supplementary Information). In particular, we tested: a) the average shortest path length and betweenness centrality as measures of global integration, b) the local efficiency as measure of local integration. The first two components of models including these measures were highly correlated with the components of the main model we tested that included degree and clustering coefficient (**Supplementary Table 5**).

Comparing training vs. no-training control groups

We conducted a no-training control experiment to investigate whether the brain connectivity changes we observed were training-specific rather than due to repeated exposure to the task. Participants in this group were tested with structured sequences in two test sessions (26.1 ±5.2 days apart) but did not receive training in between sessions.

Comparing behavioral performance in the two test sessions for the no-training control group, we found no significant main effect of Session ($F(1,20)=0.1$, $p=0.740$, $\eta_p^2=0.006$) nor a significant interaction between Session and Level ($F(1.33,26.56)=0.2$, $p=0.695$, $\eta_p^2=0.012$, Greenhouse-Geisser corrected). Further, comparing performance between the two groups (training, no-training control) showed a significant main effect of Group ($F(1,40)=39.0$, $p<0.001$, $\eta_p^2=0.493$) and a significant interaction between Group and Session ($F(1,40)=73.0$, $p<0.001$, $\eta_p^2=0.646$). Taken together, these results suggest that behavioral improvement was specific to the trained group rather than the result of repeated exposure during the two test sessions.

Further, we tested whether the learning-dependent changes we observed in the intrinsic and extrinsic connectivity analyses were specific to training. We conducted these analyses for the no-training control group and for the areas that showed significant correlations of brain connectivity changes with strategy for the training group (**Figure 4**). We computed strategy index for the control group from the post-training session, as there were no training data for this group. None of the correlations observed for the training group were significant for the no-training control group for either the intrinsic or extrinsic connectivity analysis. To compare these correlations of intrinsic and extrinsic connectivity with strategy index directly between groups, we performed a linear regression analysis with an interaction term (Group x Strategy). We observed significant differences between groups in key networks: a) intrinsic connectivity change in the Lateral Motor network (Group x Strategy interaction: $F(2,35)=8.0$, $p=0.001$, $\eta_p^2=0.316$) and in the Secondary Visual network (Group x Strategy interaction: $F(2,34)=5.6$, $p=0.008$, $\eta_p^2=0.249$); b) extrinsic connectivity change between the right posterior putamen and the Lateral Motor network (Group x Strategy interaction: $F(2,34)=3.8$, $p=0.031$, $\eta_p^2=0.184$).

Finally, we conducted a PLS regression analysis to test whether changes in degree and clustering predict individual strategy for the no-training control group. This analysis did not show any significant model compared to the null model (10,000 permutations) for any number of PLS components. Further, we found no significant correlations when correlating each of the first two PLS components from the training group with the corresponding PLS components from the no-training control group (PLS-1: $r(19)=-0.22$, $p=0.34$, CI=[-0.48, 0.11]; PLS-2: $r(19)=-0.10$, $p=0.66$, CI=[-0.50, 0.19]). Taken together, these results suggest that predicting individual strategy from changes in graph metrics of brain connectivity (degree, clustering coefficient) is specific to the training group.

Discussion

Here, we sought to identify the human brain plasticity mechanisms that mediate individual ability to learn probabilistic temporal structures and make predictions in variable environments. Linking multimodal brain imaging measures (rs-fMRI, DTI) to individual behavior, we demonstrate that these task-free measures of plasticity in brain connectivity predict individual decision strategy when learning temporal statistics. Our findings advance our understanding of the brain plasticity mechanisms that mediate our ability to learn temporal statistics in variable environments.

First, modeling the participants' predictions in our statistical learning task provides a window into the mental processes that support learning (i.e. how participants extract temporal statistics and make choices in variable environments). Learning studies typically test changes in overall task performance (i.e. accuracy, learning rate) due to training. In contrast, characterizing individual decision strategy provides insight into the learning process (i.e. what information participants learn and how they make choices), extending beyond measures of overall behavioral improvement due to task training. We demonstrate that individuals adapt

their decision strategy in response to changes in the environment's statistics (i.e. changes in the sequence structure). In particular, participants deviate from matching towards maximization when learning more complex structures (i.e. context-based statistics). Our results could not be simply explained by task difficulty, as participants reached similar performance after training when learning frequency or context-based statistics. In contrast, our results reveal that individuals alter their choices to meet the learning goal in different contexts (i.e. learning frequency vs. context-based statistics). Although our experimental design does not allow us to dissociate sequence structure from decision strategy, considering variability in decision strategy across participants allows us to test the case where sequence structure remains the same but decision strategy differs across participants. The complementary case of the same decision strategy for different sequence structures could be tested by providing the participants with trial-by-trial feedback that has been shown to encourage maximization irrespective of sequence level[9].

Second, previous work has investigated these decision strategies in the context of reward learning (e.g.[9,11,12]). Here, we test the role of decision strategy in statistical learning; that is, without explicit feedback or reward. Our results demonstrate that learning predictive statistics proceeds without explicit trial-by-trial feedback and reveal adaptive decision strategies that cannot be simply explained by changes in reward processing or training on explicit cognitive strategies that aim to boost task performance, as we did not provide trial-by-trial feedback nor instructed the participants to adopt a given strategy. Consistent with previous studies, we show that when making choices in stochastic environments individuals adopt a decision strategy (matching, maximizing) without having been explicitly instructed to follow one or the other (e.g.[11]). Further, previous work has shown that training results in changes in resting functional connectivity in a range of tasks (e.g. for review[19]); for example, perceptual[48,49] and motor learning[50,51]. Yet, most of the previous work examining learning-

dependent changes in functional connectivity has focused on reward-based rather than statistical learning (i.e. training without trial-by-trial feedback). Here, we demonstrate that statistical learning by mere exposure to temporal sequences involves cortico-striatal circuits that have been previously implicated in probabilistic[13–15] and reward-based learning[34,52]. We provide evidence that these circuits support adaptive decision strategies and learning even when the reward structure is uncertain.

Third, combining modeling of individual behavior with functional brain connectivity analysis (i.e. DTI-informed analysis of rs-fMRI data), we investigate the brain plasticity mechanisms that relate to adaptive decision strategies. Using this approach, we extend beyond previous brain imaging studies that have typically investigated whether changes in task performance (i.e. accuracy, learning rate) due to training relate to learning-dependent changes in brain function. Our results demonstrate that changes in individual decision strategies in response to changes in the environment's statistics relate to learning-dependent plasticity in distinct cortico-striatal circuits. That is, decreased connectivity in the motor circuit that is known to be involved in associative and habitual learning[34–36] may facilitate matching for learning the exact frequency statistics rather than reinforcing habitual responses. In contrast, decreased connectivity in the visual cortico-striatal circuit that has been implicated in learning predictive associations[16] may facilitate learning complex context-target contingencies by selecting the most probable outcome rather than learning the exact probability distributions.

Fourth, we provide evidence that plasticity in these cortico-striatal circuits—as indicated by learning-dependent changes in functional and structural connectivity at rest—predicts individual decision strategy when learning temporal statistics. To identify multimodal imaging predictors of individual decision strategy, we extracted graph metrics from each imaging modality (rs-fMRI, DTI) and combined them in a multivariate analysis

method (PLS regression). Our results demonstrate that graph metrics reflecting interactions within (as indicated by local integration metrics) and between (as indicated by global integration metrics) cortico-striatal circuits predict 85% of individual variability in decision strategy. In particular, this analysis reveals distinct brain plasticity mechanisms that predict: 1) changes in the decision strategy from matching to maximization in response to changes in the environment's statistics, 2) variability in decision strategy across participants independent of the sequence statistics. These mechanisms involve both functional and structural connectivity changes in motor and visual cortico-striatal circuits, in line with our rs-fMRI connectivity findings, as well as executive and motivational circuits, consistent with the role of these circuits in flexible rule learning (e.g. for review[52]).

In sum, by interrogating individual decision strategy, we provide insights into individual variability in statistical learning. Our results provide evidence for distinct brain plasticity mechanisms that predict adaptive decision strategies to flexibly solve the same learning problem (i.e. learn temporal statistics). Importantly, brain plasticity in functional and structural connectivity accounts for variability in individual strategy when learning temporal statistics. This evidence for a strong link between plasticity in brain connectivity and behavioral choice demonstrates the brain's capacity to adapt in variable environments and solve problems flexibly that could be harnessed to optimize adaptive human behavior.

Methods

Observers and Study Design

Forty-four healthy volunteers (gender: 15 females, 29 males; age: 23.54 +/-3years) took part in the experiment; half in the training group and half in the no-training control group. The sample size was determined based on previous rs-fMRI studies of learning-dependent plasticity that employed similar data analysis methods[49,50,53]. Data collection and analysis

were not performed blind to the experimental groups. Participants were randomly allocated into the two experimental groups and recruited by advertising to University students. The only exclusion criterion during recruitment was MRI safety. Data from one participant per group were excluded from further analyses due to excessive head movement, resulting in twenty-one participants in each group. All participants were naive to the study, had normal or corrected-to-normal vision and signed an informed consent. Experiments were approved by the University of Birmingham Ethics Committee.

Participants in the training group took part in multiple behavioral training and test sessions that were conducted on different days. In addition, they participated in two MRI sessions, one before the first and one after the last training session. During the training sessions participants were presented with structured sequences of unfamiliar symbols that were determined by three different Markov order models. To test whether the training was specific to the trained sequences participants were presented with both structured and random sequences during the test sessions (see Supplementary Information).

MRI data analysis

Intrinsic connectivity analysis

Following GICA (see Supplementary Information), we assessed the temporal coherence of cortical components by calculating intrinsic functional connectivity[54]. That is, intrinsic connectivity quantifies how correlated the activity across voxels within a network is. Therefore, we correlated the filtered timecourse of each voxel with every other voxel in the participant-specific component. We then applied Fisher z-transform to the correlation matrix and averaged the z-values across voxels; resulting in one component connectivity value for each participant and run. Lastly, we averaged the intrinsic connectivity values across runs to derive a single value for each participant and session.

We then tested whether changes in intrinsic connectivity with training (Post minus Pre) relate to individual decision strategy. In particular, we performed a semipartial correlation of intrinsic connectivity change with strategy index for frequency and context-based statistics.. We computed skipped Pearson correlations using the Robust Correlation Toolbox[55]. This method accounts for potential outliers and determines statistical significance using bootstrapped confidence intervals (CI) for 1,000 permutations.

To correct for multiple comparisons, we used False Coverage Rate (FCR)[56]. FCR is equivalent to the False Discovery Rate (FDR) correction for multiple comparisons when significance is determined by CI rather than p-values. In particular, for N number of tests we sorted the p-values for all statistical tests in ascending order (i.e. $p(1)\leq\ldots\leq p(N)$). We then computed the parameter R for significance level at a=0.05: $R=\max\{i: p(i)\leq i*a/N\}$. Finally, we assessed significance after multiple comparison correction based on the adjusted CI at 1-R*a/N percent[56]. In particular, we found R=1 for the N=7 tests; therefore, FCR-corrected significance for intrinsic connectivity correlations was determined at 99.3% CI.

Extrinsic connectivity analysis

To investigate changes in cortico-striatal functional connectivity due to training, we correlated the resting-state timecourse of striatal segments (as determined by the DTI-based segmentation) with the timecourse of cortical components (as determined by the ICA of the rs-fMRI signals). We then standardized the correlation coefficients (Fisher z-transform) and averaged the z-values across runs to derive a single extrinsic connectivity value for each participant and session.

We followed the same semipartial correlation method as before (see *Intrinsic connectivity analysis*) to test for learning-dependent changes in cortico-striatal functional connectivity that relate to individual decision strategy. We used the Robust Correlation Toolbox[55] to test for correlations between extrinsic connectivity change (Post minus Pre) and

strategy index for frequency and context-based statistics. We tested whether these correlations were significant after FCR correction. FCR-corrected significance for extrinsic connectivity correlations was determined at 99.3% CI (R=2 for N=14 tests).

Partial Least Squares regression analysis

To test for significant predictors of decision strategy, we used PLS regression. PLS regression applies a decomposition on a set of predictors to create orthogonal latent variables that show the maximum covariance with the response variables[42,57]. In particular, we selected twelve (12) graph nodes (i.e. AAL areas): a) striatum: bilateral caudate, bilateral putamen; b) RCEN network: right MFG; c) LCEN network: triangular part of left IFG; d) Lateral Motor network: bilateral postcentral gyrus; e) Early Visual network: bilateral calcarine sulcus; and f) ACC network: bilateral ACC. For each selected node, we computed degree as measure of global integration and clustering coefficient as measure of local integration , respectively[58]. We then entered the change in degree and clustering (Post minus Pre) of the selected nodes as predictors in the PLS model and strategy index for learning frequency and context-based statistics as response variables. Predictors and response variables were standardized (z-scored) before entered in the PLS model.

To test the significance of the model, we permutated the response variables 10,000 times and performed a PLS regression for each permutation to generate a null distribution from our data[42]. We then tested whether our sample explains more variance in the response variables than the 95 percentile of the permutated samples. We computed the significance as a function of the number of latent variables (i.e. PLS components) to select significant components for further analysis.

Next, we assessed the stability of the predictor loadings (i.e. weights) to determine the significant predictors of the response variables. We generated 1,000 bootstrap samples from our data by sampling with replacement. We then performed a PLS regression for each

bootstrap sample to generate a distribution per weight. To generate these distributions, we first corrected the estimated components for axis rotation and reflection across bootstrap samples using Procrustes rotation[59]. We normalized the weights of the observed sample (i.e. original data) to the standard deviation of the bootstrapped weights; resulting in z-score-like weights. We accepted as significant the predictors showing $|z|>2.576$ ($p=0.01$)[42], for each component independently.

Statistical analysis

The sample size for all statistical tests was n=21 (i.e. number of participants per group) unless stated otherwise. All statistical tests were two-tailed and tested for normality. Correlational analyses were also tested for heteroscedasticity within the Robust Correlation Toolbox[55] and validated by bootstrapping (1,000 permutations), as nonparametric testing is more appropriate than standard Pearson correlation (parametric test) under heteroscedasticity conditions[55]. All confidence intervals are reported at 95%.

Data availability: Behavioral and imaging data in raw and pre-processed format are available upon request from the corresponding author.

Code availability: Custom code used for data analyses is available upon request from the corresponding author.

References

1. Saarinen, J. & Levi, D. M. Perceptual learning in vernier acuity: What is learned? *Vision Res.* **35,** 519–527 (1995).
2. Christian, J. *et al.* Socio-cognitive profiles for visual learning in young and older adults. *Front. Aging Neurosci.* **7,** 1–11 (2015).
3. Siegelman, N., Bogaerts, L., Christiansen, M. H. & Frost, R. Towards a theory of individual differences in statistical learning. *Philos. Trans. R. Soc. B Biol. Sci.* **372,** 20160059 (2017).
4. Aslin, R. N. & Newport, E. L. Statistical Learning: From Acquiring Specific Items to Forming General Rules. *Curr. Dir. Psychol. Sci.* **21,** 170–176 (2012).
5. Acerbi, L., Vijayakumar, S. & Wolpert, D. M. On the Origins of Suboptimality in Human Probabilistic Inference. *PLoS Comput. Biol.* **10,** e1003661 (2014).
6. Eckstein, M. P. *et al.* Rethinking human visual attention: Spatial cueing effects and optimality of decisions by honeybees, monkeys and humans. *Vision Res.* **85,** 5–9 (2013).
7. Murray, R. F., Patel, K. & Yee, A. Posterior Probability Matching and Human Perceptual Decision Making. *PLOS Comput. Biol.* **11,** e1004342 (2015).
8. Erev, I. & Barron, G. On Adaptation, Maximization, and Reinforcement Learning Among Cognitive Strategies. *Psychol. Rev.* **112,** 912–931 (2005).
9. Shanks, D. R., Tunney, R. J. & McCarthy, J. D. A re-examination of probability matching and rational choice. *J. Behav. Decis. Mak.* **15,** 233–250 (2002).
10. Wang, R., Shen, Y., Tino, P., Welchman, A. E. & Kourtzi, Z. Learning predictive statistics from temporal sequences: Dynamics and strategies. *J. Vis.* **17,** 1 (2017).
11. Schulze, C., van Ravenzwaaij, D. & Newell, B. R. Of matchers and maximizers: How competition shapes choice under risk and uncertainty. *Cogn. Psychol.* **78,** 78–98 (2015).
12. Herrnstein, R. J. Relative and absolute strength of response as a function of frequency of reinforcement. *J. Exp. Anal. Behav.* **4,** 267–272 (1961).
13. Wang, R., Shen, Y., Tino, P., Welchman, A. & Kourtzi, Z. Learning predictive statistics: strategies and brain mechanisms. *J. Neurosci.* 0144-17 (2017). doi:10.1523/JNEUROSCI.0144-17.2017
14. Gheysen, F., Van Opstal, F., Roggeman, C., Van Waelvelde, H. & Fias, W. The Neural Basis of Implicit Perceptual Sequence Learning. *Front. Hum. Neurosci.* **5,** (2011).
15. Stillman, C. M. *et al.* Caudate Resting Connectivity Predicts Implicit Probabilistic Sequence Learning. *Brain Connect.* **3,** 601–610 (2013).
16. Turk-Browne, N. B., Scholl, B. J., Chun, M. M. & Johnson, M. K. Neural Evidence of Statistical Learning: Efficient Detection of Visual Regularities Without Awareness. *J. Cogn. Neurosci.* **21,** 1934–1945 (2009).
17. Fox, M. D. & Raichle, M. E. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci* **8,** 700–711 (2007).
18. Deco, G. & Corbetta, M. The dynamical balance of the brain at rest. *Neurosci.* **17,** 107–123 (2011).
19. Kelly, C. & Castellanos, F. X. Strengthening connections: Functional connectivity and brain plasticity. *Neuropsychol. Rev.* **24,** 63–76 (2014).
20. Sampaio-Baptista, C. & Johansen-Berg, H. White Matter Plasticity in the Adult Brain. *Neuron* **96,** 1239–1251 (2017).
21. Behrens, T. E. J. *et al.* Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nat. Neurosci.* **6,** 750–757 (2003).
22. Román, F. J. *et al.* Enhanced structural connectivity within a brain sub-network

supporting working memory and engagement processes after cognitive training. *Neurobiol. Learn. Mem.* **141,** 33–43 (2017).
23. Heitger, M. H. *et al.* Motor learning-induced changes in functional brain connectivity as revealed by means of graph-theoretical network analysis. *Neuroimage* **61,** 633–650 (2012).
24. Farrar, D. & Glauber, R. Multicollinearity in Regression Analysis: The Problem Revisited. **49,** 92–107 (1967).
25. Seger, C. A. The Involvement of Corticostriatal Loops in Learning Across Tasks, Species, and Methodologies. in *The basal ganglia IX* 25–39 (Springer, 2009). doi:10.1007/978-1-4419-0340-2_2
26. Tzourio-Mazoyer, N. *et al.* Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. *Neuroimage* **15,** 273–289 (2002).
27. Smith, S. M. *et al.* Correspondence of the brain's functional architecture during activation and rest. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 13040–5 (2009).
28. Van Dijk, K. R. a *et al.* Intrinsic Functional Connectivity As a Tool For Human Connectomics: Theory, Properties, and Optimization. *J. Neurophysiol.* **103,** 297–321 (2010).
29. van den Heuvel, M. P. & Hulshoff Pol, H. E. Exploring the brain network: A review on resting-state fMRI functional connectivity. *Eur. Neuropsychopharmacol.* **20,** 519–534 (2010).
30. Di Martino, A. *et al.* Functional Connectivity of Human Striatum: A Resting State fMRI Study. *Cereb. Cortex* **18,** 2735–2747 (2008).
31. Pauli, W. M., O'Reilly, R. C., Yarkoni, T. & Wager, T. D. Regional specialization within the human striatum for diverse psychological functions. *Proc. Natl. Acad. Sci.* **113,** 1907–1912 (2016).
32. Lehéricy, S. *et al.* Diffusion tensor fiber tracking shows distinct corticostriatal circuits in humans. *Ann. Neurol.* **55,** 522–529 (2004).
33. Draganski, B. *et al.* Evidence for Segregated and Integrative Connectivity Patterns in the Human Basal Ganglia. *J. Neurosci.* **28,** 7143–7152 (2008).
34. Balleine, B. W. & O'Doherty, J. P. Human and Rodent Homologies in Action Control: Corticostriatal Determinants of Goal-Directed and Habitual Action. *Neuropsychopharmacology* **35,** 48–69 (2010).
35. Piray, P., Toni, I. & Cools, R. Human Choice Strategy Varies with Anatomical Projections from Ventromedial Prefrontal Cortex to Medial Striatum. *J. Neurosci.* **36,** 2857–2867 (2016).
36. McNamee, D., Liljeholm, M., Zika, O. & O'Doherty, J. P. Characterizing the associative content of brain structures involved in habitual and goal-directed actions in humans: a multivariate FMRI study. *J. Neurosci.* **35,** 3764–3771 (2015).
37. Heekeren, H. R., Marrett, S. & Ungerleider, L. G. The neural systems that mediate human perceptual decision making. *Nat. Rev. Neurosci.* **9,** 467–479 (2008).
38. Ahissar, M. & Hochstein, S. The reverse hierarchy theory of visual perceptual learning. *Trends Cogn. Sci.* **8,** 457–464 (2004).
39. van den Heuvel, M. P. & Sporns, O. Network hubs in the human brain. *Trends Cogn. Sci.* **17,** 683–696 (2013).
40. Blondel, V. D., Guillaume, J. L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008,** (2008).
41. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature* **393,** 440–442 (1998).
42. McIntosh, A. R. & Lobaugh, N. J. Partial least squares analysis of neuroimaging data:

Applications and advances. *Neuroimage* **23,** 250–263 (2004).
43. Whitaker, K. J. *et al.* Adolescence is associated with genomically patterned consolidation of the hubs of the human brain connectome. *Proc. Natl. Acad. Sci.* **113,** 201601745 (2016).
44. Vértes, P. E. *et al.* Gene Transcription Profiles Associated with Inter-modular Hubs and Connection Distance in Human Functional Magnetic Resonance Imaging Networks. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **371,** 735–769 (2016).
45. Levy, D. J. & Glimcher, P. W. The root of all value: a neural common currency for choice. *Curr. Opin. Neurobiol.* **22,** 1027–1038 (2012).
46. Ridderinkhof, K. R., Van Den Wildenberg, W. P. M., Segalowitz, S. J. & Carter, C. S. Neurocognitive mechanisms of cognitive control: The role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain Cogn.* **56,** 129–140 (2004).
47. D'Ardenne, K. *et al.* Role of prefrontal cortex and the midbrain dopamine system in working memory updating. *Proc. Natl. Acad. Sci.* **109,** 19900–19909 (2012).
48. Lewis, C. M., Baldassarre, A., Committeri, G., Romani, G. L. & Corbetta, M. Learning sculpts the spontaneous activity of the resting human brain. *Proc. Natl. Acad. Sci.* **106,** 17558–17563 (2009).
49. Ventura-Campos, N. *et al.* Spontaneous Brain Activity Predicts Learning Ability of Foreign Sounds. *J. Neurosci.* **33,** 9295–9305 (2013).
50. Ma, L., Narayana, S., Robin, D. A., Fox, P. T. & Xiong, J. Changes occur in resting state network of motor system during 4weeks of motor skill learning. *Neuroimage* **58,** 226–233 (2011).
51. Albert, N. B., Robertson, E. M. & Miall, R. C. The Resting Human Brain and Motor Learning. *Curr. Biol.* **19,** 1023–1027 (2009).
52. Robbins, T. . Shifting and stopping: fronto-striatal substrates, neurochemical modulation and clinical implications. *Philos. Trans. R. Soc. B Biol. Sci.* **362,** 917–932 (2007).
53. Sami, S. & Miall, R. C. Graph network analysis of immediate motor-learning induced changes in resting state BOLD. *Front. Hum. Neurosci.* **7,** 1–14 (2013).
54. Campbell, K. L. *et al.* Robust Resilience of the Frontotemporal Syntax System to Aging. *J. Neurosci.* **36,** 5214–5227 (2016).
55. Pernet, C. R., Wilcox, R. & Rousselet, G. A. Robust Correlation Analyses: False Positive and Power Validation Using a New Open Source Matlab Toolbox. *Front. Psychol.* **3,** (2013).
56. Benjamini, Y. & Yekutieli, D. False discovery rate-adjusted multiple confidence intervals for selected parameters. *J. Am. Stat. Assoc.* **100,** 71–93 (2005).
57. Krishnan, A., Williams, L. J., McIntosh, A. R. & Abdi, H. Partial Least Squares (PLS) methods for neuroimaging: A tutorial and review. *Neuroimage* **56,** 455–475 (2011).
58. Sporns, O. Network attributes for segregation and integration in the human brain. *Curr. Opin. Neurobiol.* **23,** 162–171 (2013).
59. Milan, L. & Whittaker, J. Application of the Parametric Bootstrap to Models that Incorporate a Singular Value Decomposition. *Appl. Stat.* **44,** 31 (1995).

Acknowledgements: We would like to thank Caroline di Bernardi Luft for helping with data collection; the CamGrid team; Morten L. Kringelbach, Henrique M. Fernandes and Tim J. Van Hartevelt for help with DTI analyses; Gustavo Deco for helpful discussions; Heidi Johansen-Berg and Guy Williams for help with optimizing the DTI sequences and helpful discussions. This work was supported by grants to ZK from the Biotechnology and Biological Sciences Research Council (H012508 and BB/P021255/1), the Leverhulme Trust (RF-2011-378), the Alan Turing Institute (TU/B/000095), the Wellcome Trust (205067/Z/16/Z) and the [European Community's] Seventh Framework Programme [FP7/2007-2013] under agreement PITN-GA-2011-290011, AEW from the Wellcome Trust (095183/Z/10/Z) and the [European Community's] Seventh Framework Programme [FP7/2007-2013] under agreement PITN-GA-2012-316746, PT from Engineering and Physical Sciences Research Council (EP/L000296/1), PEV from the MRC (MR/K020706/1). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Competing interests: The authors declare no competing interests.

Author contributions: VMK: Performed research, Contributed analytic tools, Analyzed data, Wrote the paper; JG: Performed research, Contributed analytic tools, Analyzed data, Wrote the paper; PEV: Contributed analytic tools, Wrote the paper; RW: Performed research, Contributed analytic tools, Wrote the paper; YS: Contributed analytic tools, Wrote the paper; PT: Designed research, Contributed analytic tools, Wrote the paper; AW: Designed research, Wrote the paper; ZK: Designed research, Wrote the paper.

Figures

Figure 1: Trial and sequence design. (a) Trial design: Stimuli comprised four symbols chosen from Ndjuká syllabary. A temporal sequence of 8-14 symbols was presented followed by a cue and the test display. (b) Sequence design: the three Markov models used in the study. Zero-order model (level-0): each of the four symbols constitutes a different state (A, B, C, D) that occurred with a different probability . First- (level-1) and second- (level-2) order models: each state (indicated by circles) is associated with two transitional probabilities; one high (solid arrow) and one low probability (dashed arrow). Rows in the conditional probability matrix represent the temporal context, whereas columns the corresponding target.

Figure 2: Behavioral performance. (a) Normalized performance index for the training group (n=21) is shown per level and test session (pre-training: grey bars, post-training: black bars). Error bars indicate standard error of the mean across participants. (b) Boxplots of strategy index show individual variability for each level (level-0, level-1, level-2). The upper and lower error bars display the minimum and maximum data values and the central boxes represent the interquartile range (25th to 75th percentiles). The thick line in the central boxes represents the median. Open circles denote outliers. The strategy index for frequency statistics was not significantly different from matching (i.e. zero strategy index; $t(20)=-0.23$, $p=0.82$, CI=[-0.08, 0.07], Cohen's $d=-0.050$). Note that the variability across participants around zero could be due to fact that the task is probabilistic and the participants were not given trial-by-trial feedback. In contrast, the strategy index for context-based statistics (mean strategy index for level-1 and level-2) was significantly higher than zero ($t(20)=4.01$, $p<0.001$, CI=[0.08, 0.26], Cohen's $d=0.874$). (c) Scatterplot of strategy index for frequency and context-based statistics. Individual participant data are shown with open circles (n=21).

852 Points below the diagonal indicate participants that showed higher strategy index for context-
853 based compared to frequency statistics.
854
855 **Figure 3: Striatal segments and ICA components.** (a) Four striatal segments as estimated
856 by a DTI connectivity-based and hypothesis-free classification method. Segments are
857 displayed in neurological convention (left is left) and overlaid on the MNI template (green:
858 ventral striatum, blue: caudate head and anterior putamen, yellow: caudate body/tail, red:
859 posterior putamen). (b) The 7 selected ICA components are depicted organized into known
860 cortical networks. Group spatial maps are thresholded at z=1.96 for visualization
861 purposes and displayed in neurological convention on the MNI template. The x,y,z
862 coordinates denote the location of the sagittal, coronal and axial slices, respectively.
863
864 **Figure 4: Intrinsic and extrinsic connectivity analysis.** Significant skipped Pearson
865 correlations (two-sided, n=21) of (a) intrinsic connectivity change (post- minus pre-training)
866 and (b) extrinsic connectivity change with strategy index for frequency and context-based
867 statistics. Open circles in the correlation plots denote outliers as detected by the Robust
868 Correlation Toolbox.
869
870 **Figure 5: Resting-state fMRI and DTI graphs.** Whole brain graphs for (a) resting-state (rs-
871 fMRI) data and (b) DTI data. Graphs were generated based on the AAL parcellation (90 areas
872 excluding Cerebellum and Vermis) and displayed at 5% density for visualization. The
873 thickness of the edges is proportional to the average functional and structural connectivity,
874 respectively. The selected nodes are colored to represent regions within known cortico-striatal
875 circuits: caudate and putamen (magenta), right MFG and left IFG (red), postcentral gyrus
876 (cyan), calcarine sulcus (blue), and ACC (yellow). Graphs are displayed in neurological

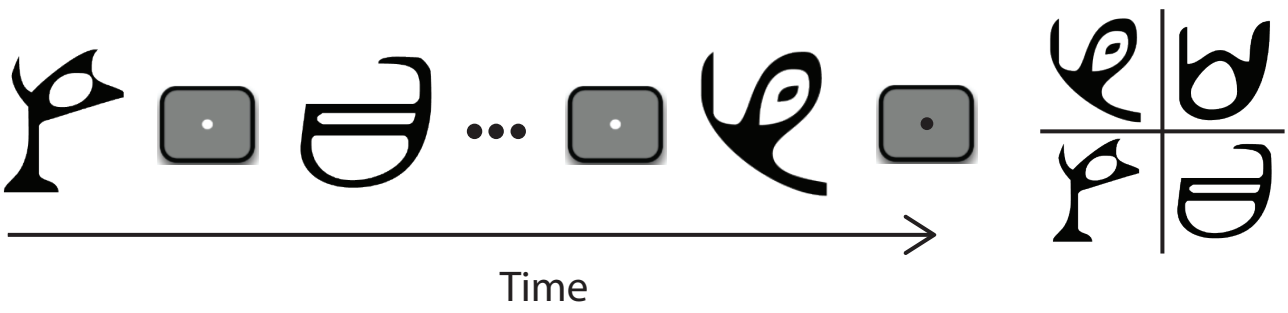
convention (left is left) in axial and sagittal views. 3D movies illustrating the rs-fMRI and DTI graphs are included in the Supplementary Information.

Figure 6: PLS weights for degree and clustering coefficient. Scatterplot of PLS-1 and PLS-2 weights for change (i.e. post- minus pre-training) in (a) degree and (b) clustering coefficient. PLS predictor weights for each selected node are indicated by symbols separately for DTI (circles) and rs-fMRI (squares) data. The color of the symbols corresponds to nodes (**Figure 5**) in cortico-striatal circuits: caudate and putamen (magenta), right MFG and left IFG (red), postcentral gyrus (cyan), calcarine sulcus (blue), and ACC (yellow). PLS predictor weights with $|z|>2.576$ ($p=0.01$) are marked by an asterisk to denote significant predictors for the respective PLS component. **Supplementary Table 4a** shows the numerical values of the PLS weights for each predictor.

Figure 7: PLS components predicting decision strategy. (a) Scatterplot of PLS-1 and PLS-2 weights (values akin to z-score) for the response variables (i.e. strategy index for frequency vs. context-based statistics). **Supplementary Table 4b** shows the numerical values of the PLS weights for each response variable. PLS-1 separates decision strategies for frequency vs. context-based statistics (i.e. negative vs. positive weight), capturing changes in decision strategy across sequence levels. PLS-2 weights equally the strategy for frequency and context-based statistics, capturing variability in decision strategy across participants independent of the sequence levels. (b) Pearson correlations (two-sided, n=21) of PLS-1 score with difference in strategy index for frequency and context-based statistics ($r(19)=0.89$, $p<0.001$, CI=[0.68, 0.96]) and PLS-2 score with mean strategy index ($r(19)=0.79$, $p<0.001$, CI=[0.49, 0.92]). (c) Significant predictors ($|z|>2.576$, $p=0.01$) for the first two PLS components are shown on the DTI graph for illustration purposes only (neurological

convention: left is left). Red nodes indicate the significant predictors for PLS-1 and blue nodes for PLS-2, irrespective of imaging modality (i.e. rs-fMRI, DTI) or graph metric (i.e. degree change, clustering coefficient change).

a

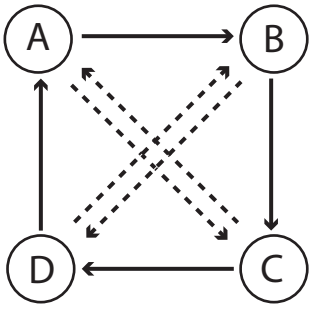


b

Level-0: Zero-order model

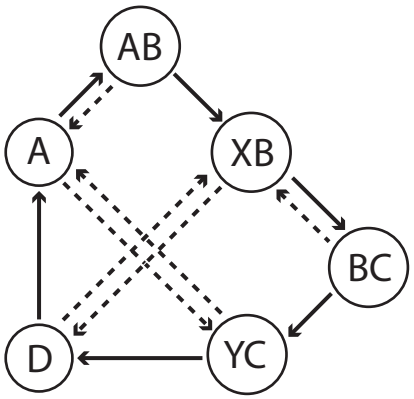
A	B	C	D
0.18	0.72	0.05	0.05

Level-1: First-order model

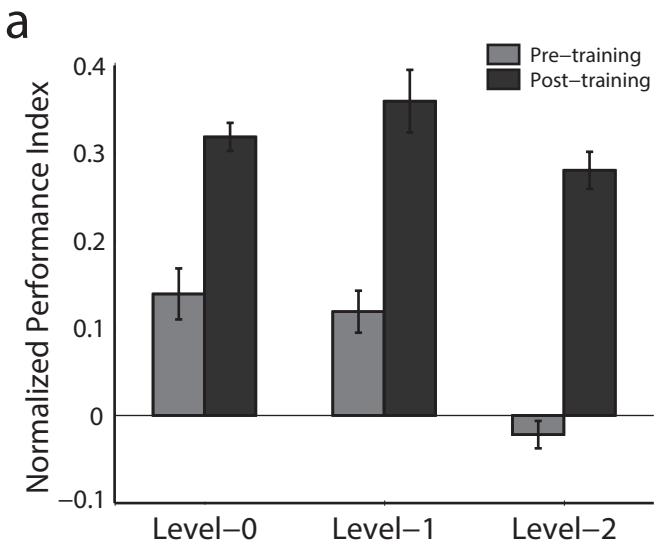


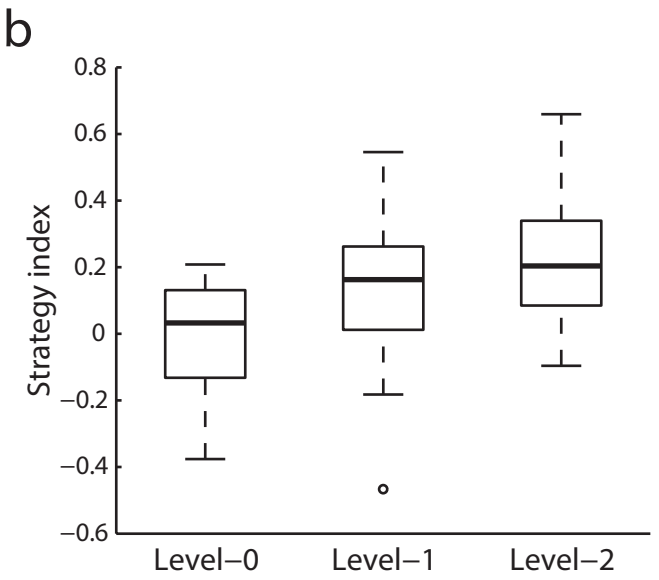
Level-1		Target			
		A	B	C	D
Context	A		0.8	0.2	
	B			0.8	0.2
	C	0.2			0.8
	D	0.8	0.2		

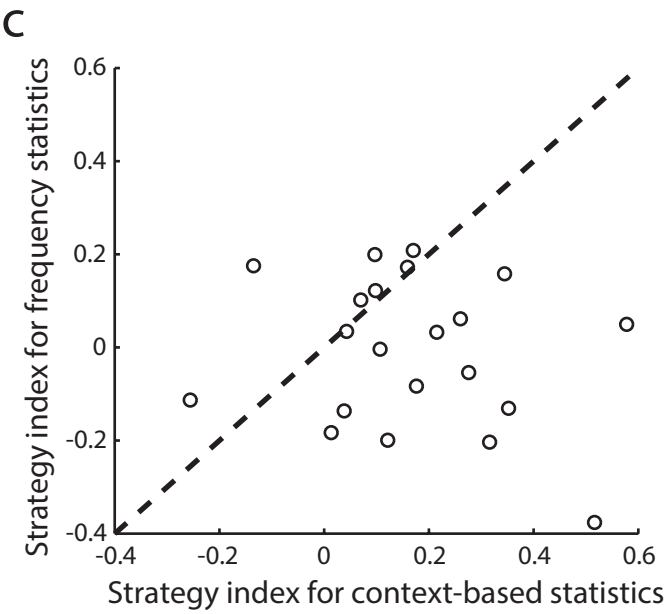
Level-2: Second-order model



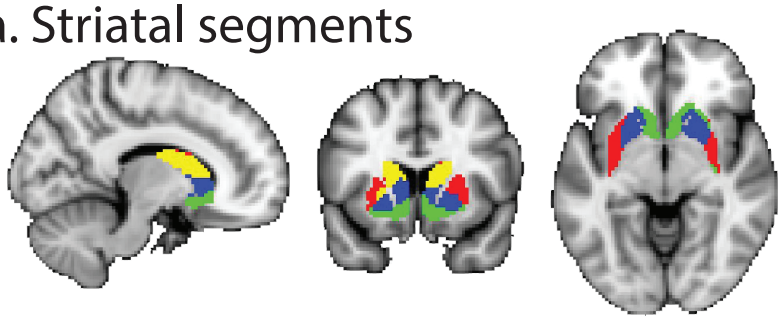
Level-2		Target			
		A	B	C	D
Context	A		0.8	0.2	
	B			0.8	0.2
	C	0.2			0.8
	D	0.8	0.2		
	AB	0.2	0.8		
	BC		0.2	0.8	


a
Pre−training
Post−training
Normalized Performance Index
0.4
0.3
0.2
0.1
0
−0.1
Level−0
Level−1
Level−2


b
Strategy index
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
Level−0
Level−1
Level−2


c
Strategy index for frequency statistics
0.6
0.4
0.2
0
-0.2
-0.4
-0.4
-0.2
0
0.2
0.4
0.6
Strategy index for context-based statistics

a. Striatal segments



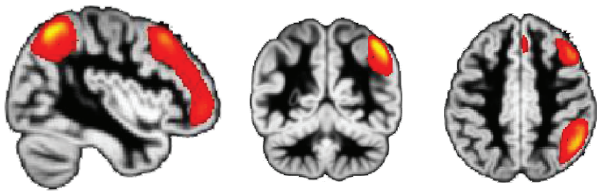
x=12 y=12 z=-6

ventral striatum
caudate head, anterior putamen
caudate body / tail
posterior putamen

b. ICA components

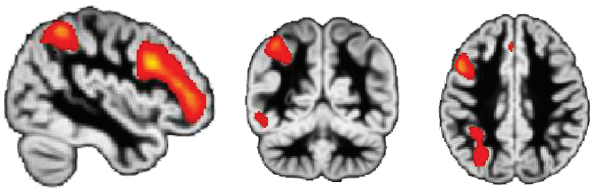
Executive networks

CP_9 (Right Central Executive)



x=42 y=-54 z=43.5

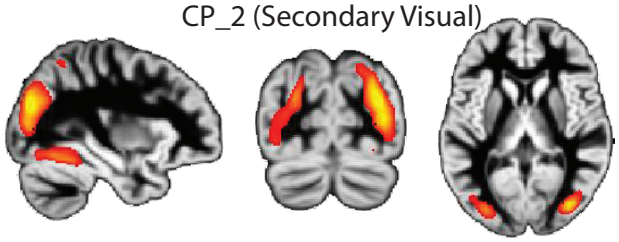
CP_14 (Left Central Executive)



x=-45 y=-55.5 z=37.5

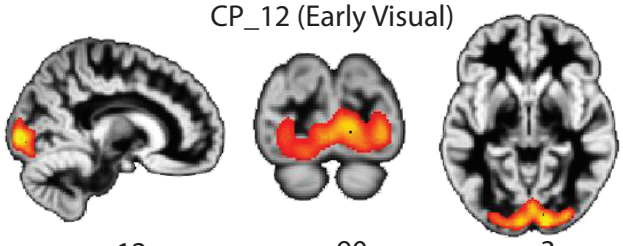
Visual networks

CP_2 (Secondary Visual)



x=31.5 y=-79.5 z=6

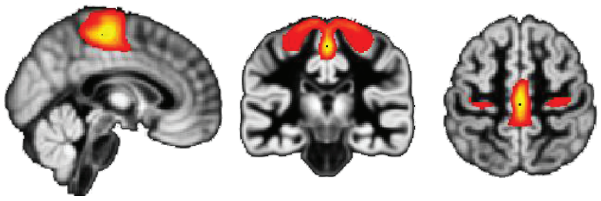
CP_12 (Early Visual)



x=12 y=-90 z=-3

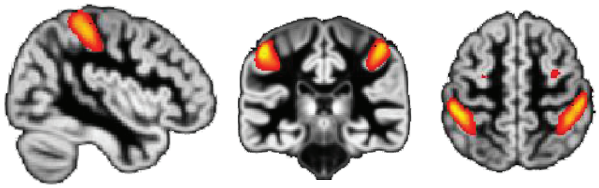
Motor networks

CP_4 (Sensorimotor)



x=0 y=-25.5 z=57

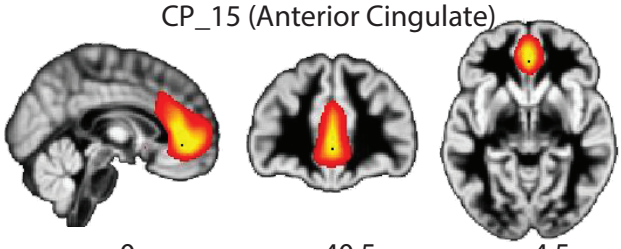
CP_5 (Lateral Motor)



x=43.5 y=-30 z=57

Motivational networks

CP_15 (Anterior Cingulate)

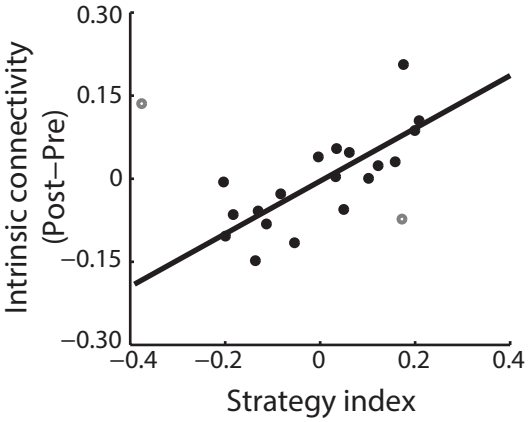


x=0 y=40.5 z=-4.5

a. Intrinsic connectivity

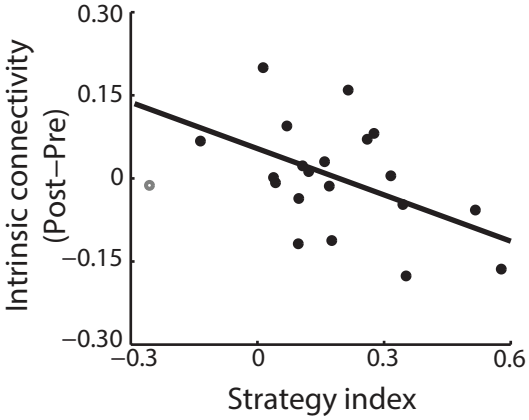
Frequency statistics

Lateral Motor



Context-based statistics

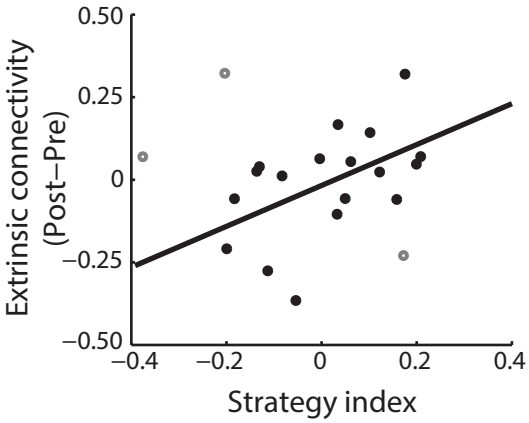
Secondary Visual



b. Extrinsic connectivity

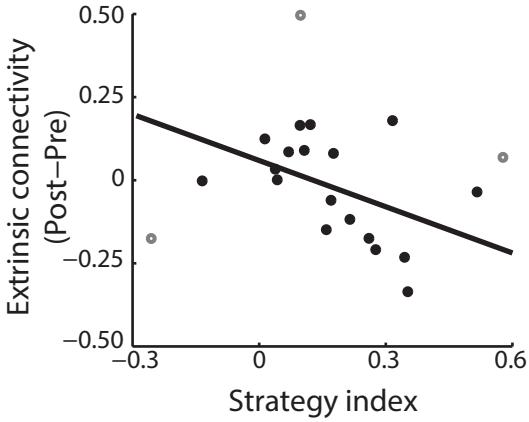
Frequency statistics

right posterior Putamen - Lateral Motor

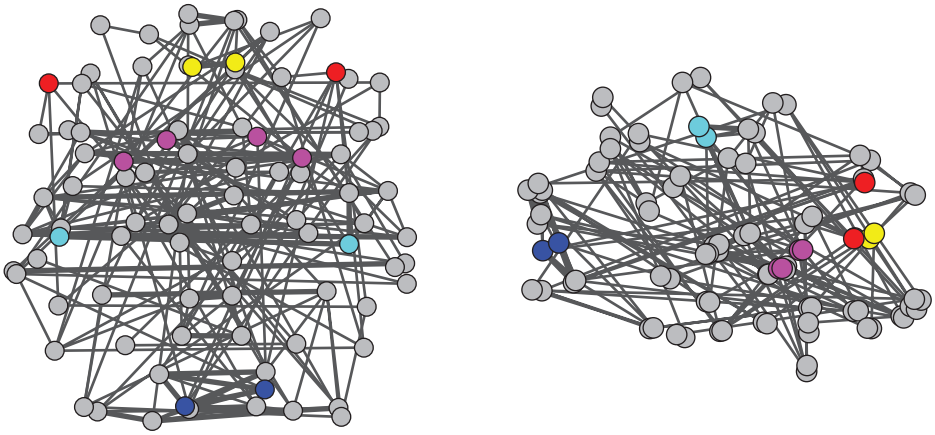


Context-based statistics

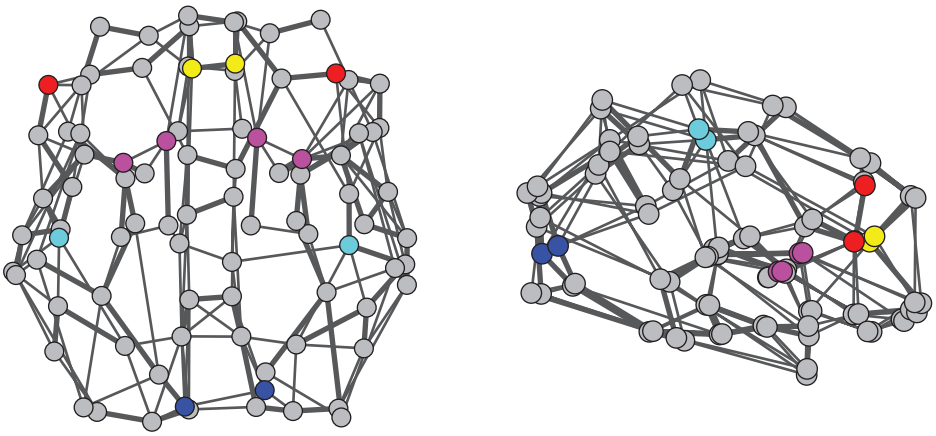
left Caudate body/tail - Early Visual



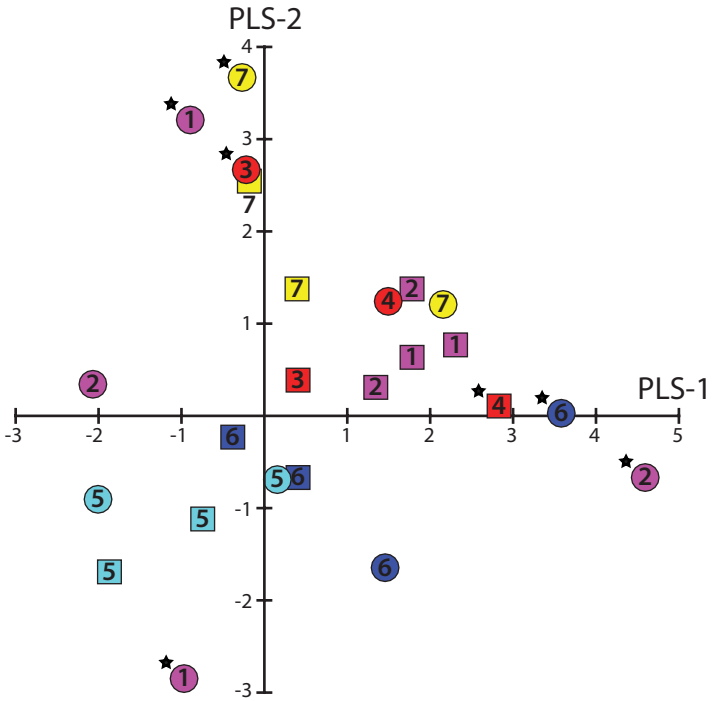
a. rs-fMRI graph



b. DTI graph

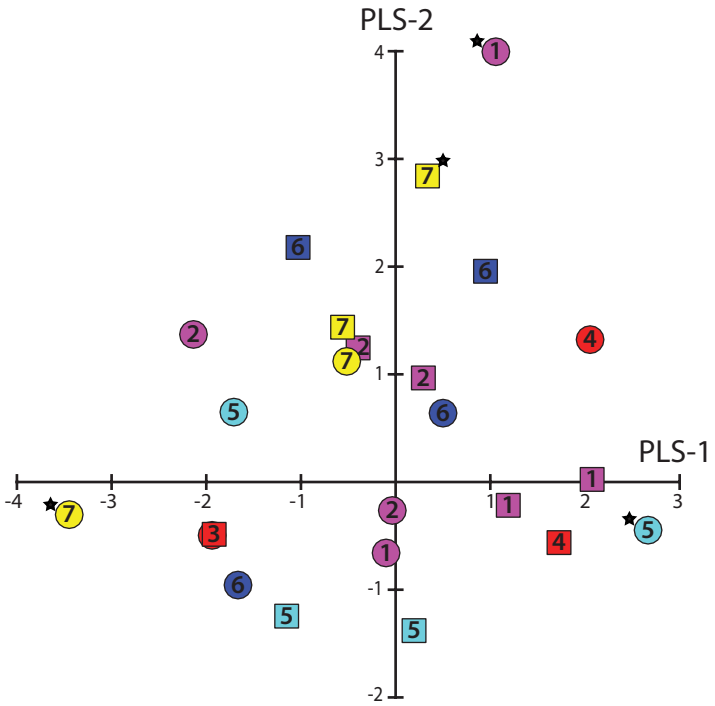


a. PLS weights for degree

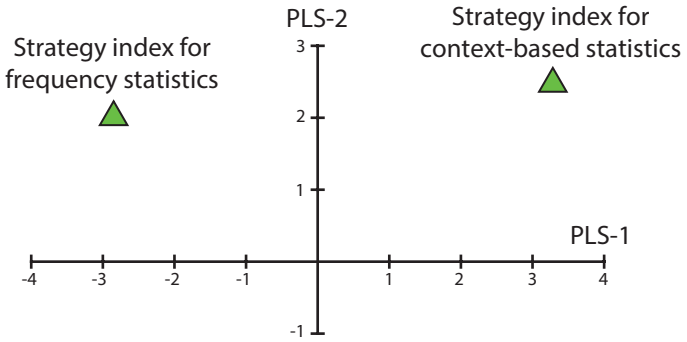


1	Caudate
2	Putamen
3	right Middle Frontal
4	left Inferior Frontal
5	Postcentral
6	Calcarine
7	Anterior Cingulate

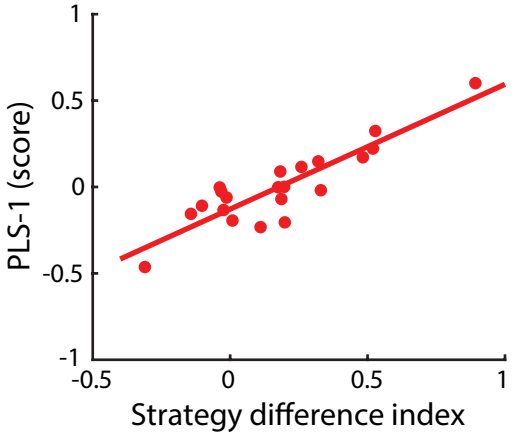
b. PLS weights for clustering coefficient

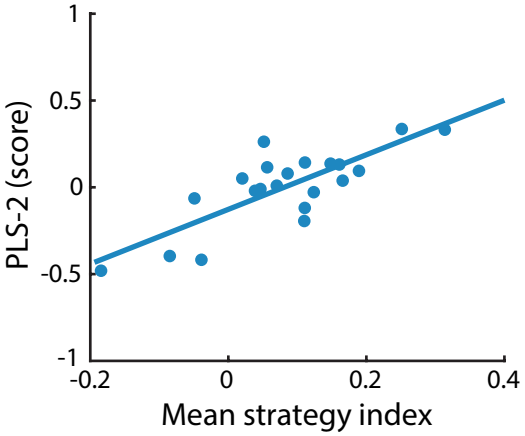


a. PLS weights for strategy

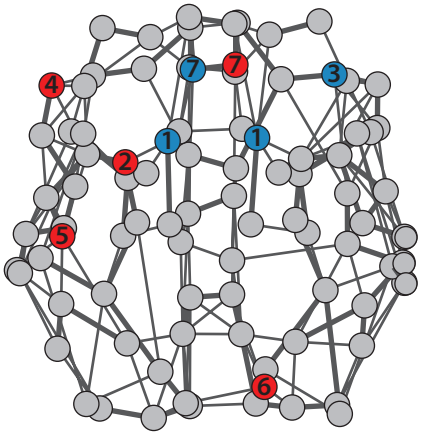


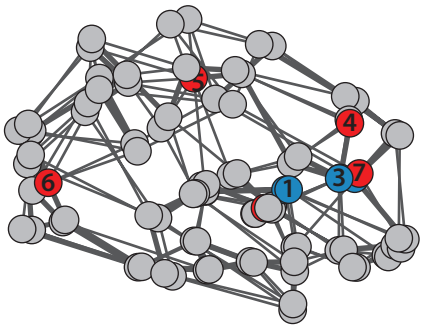
b. PLS correlations





c. PLS significant predictors





PLS-1 component
PLS-2 component

1	Caudate
2	Putamen
3	right Middle Frontal
4	left Inferior Frontal
5	Postcentral
6	Calcarine
7	Anterior Cingulate

Supplementary Methods

Stimuli: Stimuli comprised four symbols chosen from Ndjuká syllabary (**Figure 1a**) that were highly discriminable from each other and were unfamiliar to the participants. Each symbol subtended 8.5° of visual angle and was presented in black on a mid-grey background. Experiments were controlled using Matlab and the Psychophysics toolbox 3[1,2]. For the behavioral training sessions, stimuli were presented on a 21-inch CRT monitor (ViewSonic P225f 1280 x 1024 pixel, 85 Hz frame rate) at a distance of 45 cm. For the test sessions, stimuli were presented using a projector and a mirror set-up (1280 x 1024 pixel, 60 Hz frame rate) at a viewing distance of 67.5 cm. The physical size of the stimuli was adjusted so that the angular size was constant during training and test sessions.

Sequence design: We generated probabilistic sequences by using a temporal Markov model and varying the memory length (i.e. context length) of the sequence, following our previous work[3]. The model consists of a series of symbols, where the symbol at time *i* is determined probabilistically by the previous '*k*' symbols. We refer to the symbol presented at time *i*, *s(i)*, as the target and to the preceding *k*-tuple of symbols (*s(i-1), s(i-2), … , s(i-k)*) as the context. The value of '*k*' is the order or level of the sequence:

$$P(s(i) \mid s(i-1), s(i-2), \ldots, s(1)) = P(s(i) \mid s(i-1), s(i-2), \ldots, s(i-k)),\ k<i$$

In our study, we used three levels of memory length; for *k=0,1,2*. The simplest $k=0^{th}$ order model is a memory-less source. This generates, at each time step *i*, a symbol according to symbol probability *P(s)*, without taking into account the context (i.e. previously generated symbols). The order *k=1* Markov model generates symbol *s(i)* at each time *i* conditional on the previously generated symbol *s(i-1)*. This introduces a memory in the sequence; i.e. the probability of a particular symbol at time *i* strongly depends on the preceding symbol *s(i-1)*. Unconditional symbol probabilities *P(s(i))* for the case *k=0* are now replaced with conditional

ones, $P(s(i)|s(i-1))$. Similarly, an order $k=2$ Markov model generates a symbol $s(i)$ at each time i conditional on the two previously generated symbols $s(i-1)$, $s(i-2)$: $P(s(i)|s(i-1),s(i-2))$.

At each time the symbol that follows a given context is determined probabilistically, thus generating stochastic Markov sequences. The underlying Markov model can be represented through the associated context-conditional target probabilities (**Figure 1b**). We used 4 symbols that we refer to as items A, B, C and D. The correspondence between items and symbols was counterbalanced across participants. Note, that we designed the stochastic sources from which the sequences were generated so that the memory-conditional uncertainty remains the same across levels. In particular, for the zero-order source, only two symbols are likely to occur most of the time; the remaining two symbols have very low probability (0.05); this is introduced to ensure that there is no difference in the number of symbols across levels. Of the two dominant symbols, one is more probable (probability 0.72) than the other (probability 0.18). This structure is preserved in Markov chain of order 1 and 2, where conditional on the previous symbols, only two symbols are allowed to follow, one with higher probability (0.80) than the other (0.20). This ensures that the structure of the generated sequences across levels differs mainly in the memory length (i.e. context length) rather than the context-conditional probabilities.

In particular, for level-0 (zero-order), the Markov model was based on the probability of symbol occurrence: one symbol had a high probability of occurrence, one low probability, while the remaining two symbols appeared rarely (**Figure 1b**). For example, the probabilities of occurrence for the four symbols A, B, C and D were 0.18, 0.72, 0.05 and 0.05, respectively. Presentation of a given symbol was independent of the items that preceded it. For level-1 (first-order) and level-2 (second-order), the target depended on one or two immediately preceding items, respectively (**Figure 1b**). Given a context, only one of two targets could follow; one had a high probability of being presented and the other a low probability (e.g., 80% vs. 20%). For

example, when Symbol A was presented, only symbols B or C were allowed to follow, and B had a higher probability of occurrence than C.

Note, that we designed the stochastic sources from which the sequences were generated so that the memory-conditional uncertainty remains the same across levels. In particular, for the zero-order source (level-0), only two symbols are likely to occur most of the time; the remaining two symbols have very low probability (0.05); this is introduced to ensure that there is no difference in the number of symbols across levels. Of the two dominant symbols, one is more probable (probability 0.72) than the other (probability 0.18). This structure is preserved in Markov chain of order 1 (level-1) and 2 (level-2), where conditional on the previous symbols, only two symbols are allowed to follow, one with higher probability (0.80) than the other (0.20). This ensures that the structure of the generated sequences across levels differs mainly in the memory length (i.e. context length) rather than the context-conditional probabilities.

Procedure: Participants were initially familiarized with the task through a brief practice session (8 minutes) with random sequences (i.e. all four symbols were presented with equal probability 25% in a random order). Following this, participants took part in multiple behavioral training and test sessions that were conducted on different days. In addition, they participated in two brain imaging sessions, one before the first training session and one after the last training session. Participants were trained with structured sequences and tested with both structured and random sequences to ensure that training was specific to the trained sequences.

In the first test session (pre-training), participants were presented with level-0, level-1 and level-2 sequences and random sequences. Participants were then trained with level-0 sequences, and subsequently with level-1 and level-2 sequences. Training on level-0 sequences involves learning frequency statistics (i.e. participants are required to learn the occurrence

probability of each symbol), whereas training on level-1 and level-2 sequences involves learning context-based statistics (i.e. participants are required to learn the probability of a given symbol appearing depends on the preceding symbol(s)). For each level, participants completed a minimum of 3 and a maximum of 5 training sessions (840-1400 trials). Each training session comprised five blocks of structured sequences (56 trials per block) and lasted one hour. Training at each level ended when participants reached plateau performance (i.e. performance did not change significantly for two sessions). Participants were given feedback (i.e. score in the form of Performance Index) at the end of each block, rather than per-trial error feedback, which motivated them to continue with training. A post-training test session followed training per level (i.e. on the following day after completion of training) during which participants were presented with structured sequences determined by the statistics of the trained level and random sequences (90 trials each). In contrast to the training sessions, no feedback was given during test. The mean time interval (±standard deviation) between the pre-training and the post-training test sessions was 23.3 (±2.5) days.

For each trial, a sequence of 8-14 symbols appeared in the center of the screen, one at a time in a continuous stream (**Figure 1a**). This variable trial length ensured that participants maintained attention during the whole trial. The end of each trial was indicated by a red dot cue. Following this, all four symbols were shown in a 2x2 grid. The positions of test stimuli were randomized from trial to trial. Participants were asked to indicate which symbol they expected to appear following the preceding sequence by pressing a key corresponding to the location of the predicted symbol.

Psychophysical training: To ensure that sequences in each block were representative of the Markov model order per level, we generated 10,000 Markov sequences per level comprising

672 items per sequence. To quantify how close the generated sequence was to the ideal Markov model, we estimated the Kullback-Leibler divergence (KL divergence) as follows:

$$KL = \sum_{target} Q(target) \log\left(\frac{Q(target)}{P(target)}\right)$$

for the level-0 model, and

$$KL = \sum_{context} Q(context) \sum_{target} Q(target|context)\ log\left(\frac{Q(target|context)}{P(target|context)}\right)$$

for the level-1 and level-2 models, where P( ) refers to probabilities or conditional probabilities derived from the presented sequence and Q( ) refers to those specified by the ideal Markov model. KL divergence is a standard measure of distance between distributions and values close to 0 indicate small differences between the distributions. We selected fifty sequences with the lowest KL divergence (i.e. these sequences matched closely the Markov model per level). The sequences presented to the participants during the experiments were selected randomly from this sequence set.

For each trial, a sequence of 8-14 symbols appeared in the center of the screen, one at a time in a continuous stream, each for 300ms followed by a central white fixation dot (ISI) for 500ms (**Figure 1a**). This variable trial length ensured that participants maintained attention during the whole trial. Each block comprised equal number of trials with the same number of items. The end of each trial was indicated by a red dot cue that was presented for 500ms. Following this, all four symbols were shown in a 2x2 grid. The positions of test stimuli were randomized from trial to trial. Participants were asked to indicate which symbol they expected to appear following the preceding sequence by pressing a key corresponding to the location of the predicted symbol. Participants learned a stimulus-key mapping during the familiarization phase: key ‘8’, ‘9’, ‘5’ and ‘6’ in the number pad corresponded to the four positions of the test stimuli —upper left, upper right, lower left and lower right, respectively. After the participant’s response, a white circle appeared on the selected item for 300ms to indicate the participant’s

choice, followed by a fixation dot for 150ms (ITI) before the start of the next trial. If no response was made within 2s, a null response was recorded and the next trial started.

Test sessions: The pre-training test session (Pre) included nine runs (i.e. three runs per level), the order of which was randomized across participants. Test sessions after training per level included nine runs of structured sequences determined by the same statistics as the corresponding trained level and random sequences. Each run comprised five blocks of structured and five blocks of random sequences presented in a random counterbalanced order (2 trials per block; a total of 10 structured and 10 random trials per run), with an additional two 16s fixation blocks, one at the beginning and one at the end of each run. Each trial comprised a sequence of 10 stimuli which were presented for 250ms each, separated by a blank interval during which a white fixation dot was presented for 250ms. Following the sequence, a response cue (central red dot) appeared on the screen for 4s before the test display (comprising four test stimuli) appeared for 1.5s. Participants were asked to indicate which symbol they expected to appear following the preceding sequence by pressing a key corresponding to the location of the predicted symbol. A white fixation was then presented for 5.5s before the start of the next trial.

Performance index: We assessed participant responses in a probabilistic manner. We computed a performance index per context that quantifies the minimum overlap (min: minimum) between the distribution of participant responses and the distribution of presented targets estimated across 56 trials per block by:

$$PI(context) = \sum \min \left(P_{resp}(s_t|context_t), P_{pres}(s_t|context_t)\right)$$

where *t* is the trial index and the target *s* is from the symbol set A, B, C and D.

The overall performance index is then computed as the average of the performance indices across contexts, PI(context), weighted by the corresponding context probabilities:

$$PI = \sum PI(context) \cdot P(context).$$

To compare across different levels, we defined a normalized PI measure that quantifies relative participant performance above random guessing. We computed a random guess baseline; i.e. performance index PI_{rand} that reflects participant responses to targets with a) equal probability of 25% for each target per trial for level-0 ($PI_{rand} = 0.53$); b) equal probability for each target for a given context for level-1 ($PI_{rand} = 0.45$) and level-2 ($PI_{rand} = 0.44$). To correct for differences in random-guess baselines across levels, we subtracted the random guess baseline from the performance index ($PI_{normalized} = PI - PI_{rand}$).

Strategy choice and strategy index: To quantify each participant's strategy, we compared individual participant response distributions (response-based model) to two baseline models: (i) a probability matching model, where probabilistic distributions of possible outcomes are derived from the Markov models that generated the presented sequences (Model-matching), and (ii) a probability maximization model, where only the most likely outcome is allowed for each context (Model-maximization). We used KL divergence to quantify how close the response distribution is to matching and maximization distributions. KL divergence close to 0 indicates small difference between the distributions. KL is defined as follows:

$$KL = \sum_{target} M(target)\, log\left(\frac{M(target)}{R(target)}\right)$$

for the level-0 model, and

$$KL = \sum_{context} M(context) \sum_{target} M(target|context)\, log\left(\frac{M(target|context)}{R(target|context)}\right)$$

for the level-1 and level-2 models, where R( ) and M( ) denote the probability distribution or conditional probability distribution derived from the human responses and the models (i.e. probability matching or maximization) respectively, across all the conditions.

We quantified the difference between the KL divergence from the response-based model to Model-matching and the KL divergence from the response-based model to Model-maximization. We refer to this quantity as strategy choice indicated by ΔKL(Model-maximization, Model-matching) and it reflects the participant's preference towards matching or maximization. We then derived an individual strategy index by calculating the integral of each participant's strategy curve across trials and subtracting it from the integral of the exact matching curve across trials, as defined by Model-matching. We defined the integral curve difference (ICD) between individual strategy and exact matching as the individual strategy index. That is, strategy index close to zero indicates a strategy closer to matching, while higher positive values indicate deviation from matching towards maximization.

Supplementary Figure 1 illustrates how the response probability distributions may yield negative or positive strategy index values. For example, for level-1, Table A shows the context-target probability distribution that defines the matching model; a participant response distribution matching this model would indicate exact matching strategy. Table B represents the exact maximization model; that is, a participant whose response distribution follows this model chooses consistently the most probable outcome. Table C represents a random response model; that is, the participant chooses all context-target contingencies with equal probability. Participants may demonstrate this random distribution of responses at the beginning of learning before they have extracted the structure of the sequence or the exact context-target contingencies. Following training, participants may show response distributions closer to matching or deviating from matching towards maximization. Underestimating the probability of the most probable context-target contingency (e.g. Table D) will result in response distributions between the matching and the random model and yield a negative strategy index. In contrast, overestimating the probability of the most probable context-target contingency (e.g.

Table E) will result in response distributions between the matching and maximization models and yield a positive strategy index.

Further, response distributions during training (i.e. strategy choice per block: ΔKL(Model-maximization, Model-matching)) from three representative participants are shown in comparison to these models (matching, maximization, random) (**Supplementary Figure 1c**). Note that the strategy index is computed as the integral between the values of participant strategy choice and the matching model across blocks. As a result, calculating the strategy index for a participant that starts with a strategy closer to random and then deviates closer to the matching model may result in a negative (e.g. participant A) or a positive value (e.g. participant B). For example, data from a participant A that underestimates the probability of the most probable context-target contingency during most of the training blocks yield a negative strategy index. However, data from a participant B that overestimates the probability of the most probable context-target contingency in some of the training blocks yield a positive strategy index, as the integral becomes positive when the participant strategy crosses the matching model curve. In contrast, strategy choice data for a participant C that deviates from matching towards maximization yields a higher positive strategy index.

Further, we provide a mathematical description of strategy index variability. In particular, we generated synthetic response data from a virtual participant and present a two-parameter model characterizing the participant response distribution. Response distribution (denoted as P) is described as the mixture of two components, P_1 and P_2. To control the contribution of these two components, we define a parameter β as the weight of the two components ($0 \leq \beta \leq 1$): $P = \beta P_1 + (1-\beta) P_2$. The first component is the random model (i.e. equal probabilities for all context-target contingencies). Participants may follow this random model of responses at the beginning of training before they have learned the sequence structure and relative probabilities. The second component reflects the probability distribution of the items

in the sequence presented to the participant, e.g. P_2 = [0.2, 0.8, 0, 0]. This specification assumes that (1) only two items have non-zero probability; (2) the high probable target is four times more frequent than the less probable target. To capture how the participants learn these contingencies, we parameterized this distribution as follows: P_2 = [1-α, α, 0, 0], where $0 \leq \alpha \leq 1$. In particular, for (i) α = 1, the participant predicts always the most probable target (i.e. maximization); (ii) α = 0.8, the participant responses match the target distribution (i.e. matching); (iii) α = 0.5, the participant predicts equally the two possible (non-zero probability) targets; (iv) α < 0.5, the participant predicts the less probable target more frequently than the more probable target. In sum, we formulate our synthetic response model as follows: P = β [0.25, 0.25, 0.25, 0.25] + (1-β) [1-α, α, 0, 0].

To illustrate how the strategy index varies with parameters α and β, we computed the strategy index for all possible combinations of α and β values, where α and β vary between 0 and 1. This generated a strategy index surface as a function of α and β (**Supplementary Figure 2**). In particular, for β = 1 the strategy index is invariant to the parameter α and reflects equal responses for all targets (i.e. random model); yielding a strategy index value of -0.26. For β = 0, the model is reduced to P = [1-α, α, 0, 0] and is fully described by the P_2 component (see above). Therefore, (i) for α = 1 the model describes a maximization response (i.e. strategy index = 0.63), (ii) for α = 0.8 it describes a matching response (i.e. strategy index = 0), (iii) for α = 0.5 it describes a random response between the two possible targets (i.e. strategy index = -0.26) and (iv) for α < 0.5 it describes predictions of the less probable target more frequently than the more probable target (i.e. strategy index < -0.26). Further, for $0.5<\alpha<0.8$ the participant would underestimate the probability of the most probable target and yield a strategy index between -0.26 and 0; whereas for $0.8<\alpha<1$ the participant would overestimate the probability of the most probable target and yield a strategy index between 0 and 0.63. Note that the strategy index increases monotonically with α for a fixed β.

Supplementary Figure 2 presents data from three representative participants based on this two-parameter model. In particular, we present the evolution of their strategy index across training blocks as a walk on the model surface. That is, we fitted the two-parameter model on the participants' response data per block and estimated the parameters α and β per participant and block. We then computed the participant strategy index as the difference between the participant strategy choice and the matching model. In particular, we observed that all participants started close to the random model ($\beta \approx 1$) and then deviated towards higher α and lower β values. However, the trajectory and end point of the individual participants varied and therefore yielded different strategy index values. That is, participant A showed $0.5<\alpha<0.8$ throughout most of the training blocks (i.e. underestimated the highly probable targets) while $\alpha \approx 0.8$ (i.e. close to matching) at the end of the training, yielding a negative strategy index. In contrast, participant B showed $\alpha \approx 0.8$ consistently across blocks and therefore yielded a strategy index close to 0 (i.e. matching). Finally, participant C overestimated the highly probable targets (i.e. $0.8<\alpha<1$) and yielded a higher positive strategy index (i.e. closer to maximization).

MRI data acquisition: Scanning was conducted using a 3T Philips Achieva MRI scanner with a 32-channel head coil. T1-weighted anatomical data (175 slices; 1×1×1 mm^3 resolution) were collected during the first scanning session. Resting-state echo-planar imaging (EPI) data (gradient echo-pulse sequences) were acquired in both scanning sessions (whole brain coverage; 180 volumes; TR=2s; TE=35ms; 32 slices; 2.5x2.5x4 mm^3 resolution; SENSE). The benefit of non-isotropic resolution is acquisition speed; that is, faster acquisition of fewer slices at higher in-plane resolution (keeping voxel volume constant and signal-to-noise ratio similar). This is advantageous for resting-state fMRI (rs-fMRI) that requires relatively high temporal resolution. We employed standard pipelines (i.e. SPM) that have been extensively used to model fMRI data at non-isotropic resolution. We employed a well-established volumetric

analysis (i.e. Group Independent Component Analysis-GICA) to investigate functional connectivity at rest that has been developed and validated on non-isotropic data[4–8]. Finally, a recent study[9] has shown highly similar ICA results between isotropic and anisotropic datasets.

We collected rs-fMRI from three runs that each lasted for 6 minutes. Participants were instructed to keep their eyes open and maintain fixation to a white dot presented at the center of the screen. Diffusion Tensor Imaging (DTI) data were also collected in both scanning sessions and the acquisition consisted of 60 isotropically-distributed diffusion weighted directions (b=1500 smm^{-2}; TR=9.5s; TE=78ms; 75 slices; 2x2x2 mm^3 resolution; SENSE) plus a single volume without diffusion weighting (b=0 smm^{-2}, denoted as b0). The DTI sequence was repeated twice during each session, once following the Anterior-to-Posterior phase-encoding direction and once the Posterior-to-Anterior direction. This acquisition scheme was implemented to allow correction of susceptibility-induced geometric distortions[10].

DTI connectivity-based segmentation of striatum: Previous work across species[11,12] has shown that dissociable cortical projections from anatomically-defined striatal subdivisions mediate distinct brain functions. To investigate learning-dependent changes in these cortico-striatal connections, we defined the striatum (i.e. caudate and putamen) anatomically from the Automated Anatomical Labeling (AAL) atlas[13]. We then conducted a DTI connectivity-based segmentation to segment the striatum into finer subdivisions (i.e. segments) based on their whole-brain connectivity profile[14].

We pre-processed and analyzed the DTI data in FSL 5.0.8 (FMRIB Software Library, http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/). We first corrected the data for susceptibility distortions, eddy currents and motion artifacts (FSL topup and FSL eddy)[15] and rotated the gradient directions (bvecs) to correct for the estimated motion rotation[16,17]. We generated a distribution model in each voxel using *FSL BedpostX*[18] with default parameters.

To simulate tracts from a seed defined in MNI space, we computed the transformation matrix from MNI to native space per participant (FSL flirt). We followed a 4-step registration procedure: (a) aligned the non-weighted diffusion volume (b0) of each session to their midspace and create a midspace-template (rigid-body)[19,20], (b) aligned the midspace-template to the anatomical (T1) scan (affine), (c) aligned the T1 image to the MNI template (affine) and (d) inverted and combined all the transformation matrices of the previous steps to obtain the MNI-to-native registration. The results of each step were visually inspected to ensure that the alignment was successful.

We then simulated tracts (i.e. probabilistic streamlines) starting from the seed area (i.e. striatum) to the rest of the brain (i.e. target area) using the *ProbtrackX* algorithm[21]. Following a hypothesis-free classification method[22], we down-sampled the target area (AAL atlas excluding the seed: bilateral caudate and putamen) to 4x4x4 mm^3 resolution. As the seed areas were in MNI space, we provided the MNI-to-native transformation matrix and used the *omatrix2* option to create a seed-by-target connectivity matrix (the *ProbtrackX* algorithm transforms the seed from MNI to native space and performs the probabilistic tractography simulation in native space; the results are then transformed back into MNI space). We used a mid-sagittal exclusion mask to prevent tracts from crossing hemispheres[21] and length correction to account for the distance-from-the-seed bias towards shorter connections[22]. The parameters we used in *ProbtrackX* are: 5000 samples per voxel, 2000 steps per sample until conversion, 0.5mm step length, 0.2 curvature threshold, 0.01 volume fraction threshold and loopcheck enabled to prevent tracts from forming loops. We repeated this procedure for each hemisphere (**Supplementary Figure 3**).

This analysis generated a connectivity matrix from each voxel in the seed area to every voxel in the target area. Defining the seed in the MNI space guaranteed the same number of voxels in the seed across participants (after the data were transformed back from native to MNI

space), alleviating differences in individual brain size. Subsequently, we concatenated the connectivity matrices across participants and groups and correlated the connectivity values from and to each voxel in the seed; generating a seed-by-seed correlation matrix. We then performed k-means clustering on the correlation matrix for 2 to 8 classes (squared Euclidean distance). Lastly, we converted each class to a binary mask in MNI space to create the striatal segments and down-sampled them to the resting-state resolution (3x3x4 mm^3) for further analysis.

To find the optimal number of clusters, we computed the mean silhouette value per clustering by averaging the values across voxels[23]. The silhouette value shows how similar each voxel is to voxels of its class compared to voxels of other classes. Therefore, we selected the highest number of clusters that shows the maximum mean silhouette value averaged for the two hemispheres. This method resulted in 4 striatal segments per hemisphere (average silhouette value of 0.4) that corresponded to known anatomical subdivisions of the striatum (**Figure 3a, Supplementary Table 1**).

Resting-state data pre-processing: We pre-processed the resting-state data in SPM12.2 software package (http://www.fil.ion.ucl.ac.uk/spm/software/spm12/) following the optimized pipeline described in recent work[5]. We first processed the T1-weighted anatomical images by applying brain extraction and segmentation (SPM segment). From the segmented T1 we created a white matter (WM) mask and a cerebrospinal fluid (CSF) mask. For each resting-state run, we corrected the EPI data for slice scan timing (i.e. to remove time shifts in slice acquisition, SPM slice timing) and motion (least squares correction, SPM realign). We co-registered all EPI runs to the first run per participant (rigid body) and subsequently to the T1 image (rigid body, resliced to 1 x 1 x 1 mm^3) and calculated the mean CSF and WM signal per volume (SPM coregister & reslice). We then aligned the T1 image to the MNI space (affine)

and applied the same transformation to the EPI data (SPM normalise). We resliced the aligned EPI data to 3 x 3 x 4 mm^3 resolution and applied spatial smoothing with a 5mm isotropic FWHM Gaussian kernel (SPM smooth). Finally, we despiked any secondary motion artifacts using the Brain Wavelet Toolbox[24], regressed out the signal from CSF and the motion parameters (translation, rotation and their squares and derivatives[25]) and applied linear detrending[26]. Note that the pipeline we followed[5] does not include the global signal as a nuisance regressor, consistent with a recent review[27] suggesting that global signal regression may not be appropriate for comparisons between sessions and groups.

Independent Component Analysis (ICA): We used spatial GICA[6,28] to extract participant- and session-specific hemodynamic source locations using the Group ICA fMRI Toolbox (GIFT) (http://mialab.mrn.org/software/gift/). Pre-processed EPI data from both groups (i.e. training, no-training control) from both sessions (i.e. Pre, Post) were included in the GICA. Following pre-processing of each run, the mean value per voxel was removed and dimensionality reduction was performed. We used the Minimum Description Length criteria (MDL)[29] to estimate the dimensionality and determine the number of independent components. We used a two-level dimensionality reduction procedure using Principal Component Analysis (PCA); first at the participant level and then at the group level. The ICA estimation (Infomax algorithm) was run 20 times and the component stability was estimated using ICASSO[30].

This procedure resulted in 22 spatially independent components. We then generated participant-specific spatial maps for each component using GICA3 back reconstruction[4]. Lastly, participant and group spatial maps were scaled to z maps for further analysis[31]. We then used a quantitative method, as described in previous work[32], to remove components of non-neuronal origin. We first thresholded the group spatial maps at z=1.0 and calculated the spatial correlation of each component with CSF and grey matter (GM) probabilistic maps (as extracted

from the MNI template). We rejected any component with a spatial correlation of $R^2 > 0.025$ with CSF or of $R^2 < 0.025$ with GM. To supplement this method, we visually inspected all rejected components to verify that they were not of neuronal origin. This method resulted in 5 rejected components: 2 components had high spatial correlations with CSF and 3 components had low spatial correlations with GM.

We correlated the thresholded maps of the remaining components with known network templates and labeled each component based on its highest correlation value to these templates[7,33]. We selected 7 components (**Figure 3b, Supplementary Table 2**) that showed high correlation with templates of cortical regions involved in executive, motor, visual and motivational networks[11,12].

To extract the resting-state timecourse for each cortical ICA-based component and DTI-based striatal segment, we used an autoregressive AR(1) model (SPM first-level analysis) on the pre-processed data before ICA to treat for serial correlations[34]. Following the whole-brain modeling, we extracted the timecourse per voxel per component (SPM VOI extraction), as defined by participant-specific spatial maps thresholded at z=2.576 (p=0.01). We then applied a 5th order Butterworth band-pass filter, between 0.01 and 0.08 Hz to remove effects of scanner noise and physiological signals (respiration, heart beat)[35]. In addition, we extracted the first eigenvariate across all voxels in each component to derive a single timecourse per component for subsequent connectivity analysis.

Graph analysis: To construct a functional connectivity matrix for each participant, we followed the same processing steps as for the extrinsic connectivity analysis. We extracted the first eigenvariate across all voxels in each AAL region (90 areas; excluding Cerebellum and Vermis) and constructed a 90x90 correlation matrix by correlating the timecourse of each AAL region with every other AAL region. We then standardized the correlation coefficients using

Fisher z-transform and averaged the z-values across the three rs-fMRI runs to derive a single functional connectivity matrix for each participant and session.

To construct a structural connectivity matrix for each participant, we simulated tracts (i.e. probabilistic streamlines) from each AAL area (i.e. seed mask) to any other AAL area (i.e. termination masks; excluding Cerebellum and Vermis) in native space using the *Probabilistic Tracking* algorithm (FSL ProbtrackX)[21]. The parameters we used in *ProbtrackX* are: 5000 samples per voxel, 2000 steps per sample until conversion, 0.5mm step length, 0.2 curvature threshold, 0.01 volume fraction threshold and loopcheck enabled to prevent tracts from forming loops. To control for differences in volume across seeds and participants, we normalized the tract count (i.e. the number of streamlines reaching area *j* when seeded from areas *i*) by the total number of tracts started from the seed region[36]. Finally, we averaged the normalized tract count from area *i* to area *j* and from area *j* to area *i* to create a symmetric structural connectivity matrix for each participant and session.

We then constructed participant-specific binary graphs based on the connectivity matrices for each modality (i.e. rs-fMRI, DTI). We first generated the Minimum Spanning Tree[37] per matrix to create a connected graph for each participant and session. We then iteratively added the strongest edges irrespective of the sign (i.e. using the absolute functional connectivity value), until we reached a certain density level. Previous work in a similar-sized parcellation[38] has shown that density lower than 15% may result in sparse graphs and higher than 25% in graphs without small-world topology. Thus, we generated graphs at 20% density and then evaluated the stability of our findings in a range of density levels: from 10 to 30% in increments of 5. We used the Brain Connectivity Toolbox[39] to calculate graph metrics per participant and modality.

We note that the DTI and rs-fMRI metrics used in our graph analysis were derived by data pre-processed at native vs. standard space. In particular, DTI tractography is typically

performed in the native space to achieve best performance of the tracking algorithms[21], whereas rs-fMRI data are typically normalized to a standard space (e.g. MNI) before computing functional connectivity[5]. Following previous studies, we analyzed the DTI data in native space, while the rs-fMRI data in standard space (i.e. data were normalized to MNI), as these data needed to be in a common space for group analysis across participants. While some recent studies recommend performing the rs-fMRI analysis in native space to minimize the effect of interpolation and improve localization[40,41], others have found no difference with and without the inclusion of the normalization step[42]. Further, our analysis approach makes it unlikely that these differences in interpolation between data types (i.e. rs-fMRI, DTI) have a significant effect on our results. First, we selected brain regions for both the rs-fMRI and DTI graph analysis based on the AAL parcellation, resulting in larger size brain regions. This makes it unlikely that small differences in the interpolation step would significantly affect the connectivity values estimated across all voxels in each brain region. Second, for the rs-fMRI data we computed the first eigenvariate when we extracted the timecourse per brain region and computed functional connectivity from these values. This step extracts the most representative timecourse from all the voxels in each brain region based on their common variance; therefore, it minimizes the effects of noise and interpolation[43]. Third, for each imaging modality (i.e. rs-fMRI, DTI) we generated binary graphs and compared the connectivity values to select the strongest connections within-modality rather than comparing connectivity across modalities. That is, we created binary graphs at 20% density level by selecting the edges with the top 20% connectivity values, for each modality and session. We computed degree and clustering coefficient from these graphs per modality and used these metrics in the PLS regression to combine data from both modalities.

Partial Least Squares (PLS) modeling: control analyses: Results in the main text are presented for a network density of 20%. Here we show the robustness of these results in a range of densities (10%-30%) typically used in brain network analyses[38]. We calculated degree and clustering for 10% to 30% density in increments of 5% per session (Pre, Post). We computed the difference between the two curves (Post minus Pre) for each metric (degree, clustering coefficient)[44] and performed the same PLS regression analysis as before. We tested for model significance using permutation testing (10,000 permutations) and then correlated the estimated PLS components and bootstrapped weights (1,000 samples) with the components and weights estimated for 20% density as shown in the main text. We found that the first PLS component across densities was significant compared to the null ($p=0.05$) and showed a high correlation with the PLS-1 component for 20% density ($r(19)=0.94$, $p<0.001$, CI=[0.85, 0.98]). Further, the predictor weights across densities showed a high correlation with the weights for 20% density ($r(46)=0.84$, $p<0.001$, CI=[0.67, 0.93]). PLS-2 across densities was not significant in comparison to the null model; however, it showed a high correlation with the PLS-2 component and its weights for 20% density (component: $r(19)=0.89$, $p<0.001$, CI=[0.75, 0.95]; weights: $r(46)=0.89$, $p<0.001$, CI=[0.83, 0.94]). Similarly, PLS-3 across densities was not significant compared to the null and showed weaker correlations with the PLS-3 component for 20% density (component: $r(19)=0.77$, $p<0.001$, CI=[0.63, 0.88]; weights: $r(46)=0.48$, $p<0.001$, CI=[0.11, 0.71]). We therefore restricted the main analysis to the first two components. **Supplementary Figure 6** summarizes the weights (combinations of nodes and metrics) for PLS-1 and PLS-2 for the average metrics (10% to 30% density).

Further, to test whether our findings generalize to other parcellation schemes than the AAL atlas, we created graphs at 20% density using the Shen[45] and Brainnetome[46] atlases that provide a finer whole brain parcellation. We selected nodes that corresponded to the same anatomical areas as the selected AAL nodes and performed a similar PLS regression analysis.

We found that both atlases yielded significant results (Shen: first three components; Brainnetome: first four components). Moreover, we found that the first two components for these atlases were highly similar to our results when using the AAL atlas (Shen: PLS-1: $r(19)=0.75$, $p<0.001$, CI=[0.42, 0.92], PLS-2: $r(19)=0.83$, $p<0.001$, CI=[0.53, 0.93]; Brainnetome: PLS-1: $r(19)=0.73$, $p<0.001$, CI=[0.44, 0.89], PLS-2: $r(19)=0.87$, $p<0.001$, CI=[0.68, 0.94]). Note that the Brainnetome atlas provides a parcellation of the striatum (i.e. ventral caudate, dorsal caudate, dorsolateral putamen and ventromedial putamen) that is comparable to our DTI-based segmentation (**Figure 3a**). Further, the significant predictors for PLS-1 were: a) degree change in right ventral caudate (rs-fMRI), left dorsal caudate (rs-fMRI), left ACC (DTI) and left postcentral (rs-fMRI); b) clustering change in right ventral caudate (DTI) and left postcentral (rs-fMRI); whereas for PLS-2 were: a) degree change in right MFG (DTI) and left postcentral (DTI); b) clustering change in left ACC (DTI), right dorsolateral putamen (rs-fMRI) and right ACC (rs-fMRI). Taken together, these findings suggest that our graph analysis is robust across parcellation schemes that segment the striatum at different scales, making it unlikely that our results were confounded by the selected parcellation atlas.

Finally, we tested whether our findings generalize to other graph metrics that relate to global and local integration. In particular, we tested: a) the average shortest path length (i.e. average number of a node's transitions via graph edges to any other node in the network) and betweenness centrality (i.e. number of shortest paths that traverse through a certain node) as measures of global integration[47,48], b) the local efficiency (i.e. how efficiently a node's neighbors communicate if this node is removed) as measure of local integration[49]. These measures have been previously shown to relate to learning and brain plasticity[50–52].We conducted similar PLS regression analyses as for our main model (i.e. Model-1: degree and clustering coefficient) for the following models based on combinations between global and local integration metrics: a) Model-2: average shortest path length and clustering coefficient,

b) Model-3: average shortest path length and local efficiency, c) Model-4: degree and local efficiency, d) Model-5: betweenness centrality and clustering coefficient, e) Model-6: betweenness centrality and local efficiency. All models showed significant results when tested for 10,000 permutations (Model-2: first component, p=0.010; Model-3: first two components, p=0.044; Model-4: first three components, p=0.012; Model-5: first three components, p=0.026; Model-6: first component, p=0.022). Further, the first two components for these models were highly correlated to the components of the main model (Model-1) including degree and clustering coefficient (**Supplementary Table 5**). Thus, our findings showing that learning-dependent plasticity in cortico-striatal networks predicts individual behavior (i.e. decision strategy) are not limited only to selected measures of global or local integration.

Further, including all the above graph metrics in the same PLS model (Model-7: degree, average shortest path length, betweenness centrality, clustering coefficient and local efficiency), the model was significant for the first three PLS components compared to a null model (p=0.045, 10,000 permutations). In addition, the first two components for this model were highly correlated to the components of Model-1 (**Supplementary Table 5**), generalizing our results to a larger number of metrics that characterize whole-brain network connectivity.

No-training control experiment: Scanning for the no-training control experiment was conducted using a 3T MRI scanner with a 32-channel head coil. T1-weighted anatomical data (175 slices; 1×1×1 mm^3 resolution) were collected during the first scanning session. Resting-state EPI data (gradient echo-pulse sequences) were acquired in both scanning sessions with the same sequence as the one used in the training experiment (whole brain coverage; 180 volumes; TR=2s; TE=30ms; 36 slices; 2.5x2.5x4 mm^3 resolution; GRAPPA). We collected rs-fMRI from three runs that each lasted for 6 minutes. DTI data were also collected in both scanning sessions and the acquisition parameters were matched as closely as possible to the

training group: 60 isotropically-distributed diffusion weighted directions (b=1500 smm^{-2}; TR=8.9s; TE=91ms; 72 slices; 2x2x2 mm^3 resolution; GRAPPA) plus a single volume without diffusion weighting (b=0 smm^{-2}). The DTI sequence was repeated twice during each session, once following the Anterior-to-Posterior phase-encoding direction and once the Posterior-to-Anterior direction.

To ensure that the data quality was similar between the two groups (training vs. no-training control) that were tested using highly similar sequences and scanning parameters, we tested for differences related to a) head movement and b) spikes for the rs-fMRI data, and a) head movement and b) diffusion tensor model fit for the DTI data. For the rs-fMRI data, we calculated the maximum root mean square (rms) movement per run (based on x,y,z motion parameters estimated by SPM realign) and the maximum number of spikes per run (based on the Spike Percentage output of the Brain Wavelet toolbox[24]). For the DTI data, we calculated the root mean square (rms) movement per session (based on *eddy*'s *restricted_movement_rms* output) and the sum of squared errors (sse) from diffusion tensor model fit[18]. No significant differences were observed between groups for head movement (rs-fMRI: $F(1,40)=0.31$, $p=0.578$, $\eta_p^2=0.008$; DTI: $F(1,40)=1.84$, $p=0.182$, $\eta_p^2=0.044$), number of spikes ($F(1,40)=1.19$, $p=0.283$, $\eta_p^2=0.029$) or diffusion tensor model fit for the seed areas, the whole brain and the white-matter ($F(1,40)=0.77$, $p=0.386$, $\eta_p^2=0.019$). Thus, these analyses suggest that it is unlikely that differences in connectivity between groups could be due to differences in data quality.

Supplementary References

1. Brainard, D. H. The Psychophysics Toolbox. *Spat. Vis.* **10,** 433–436 (1997).
2. Pelli, D. G. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* **10,** 437–442 (1997).
3. Wang, R., Shen, Y., Tino, P., Welchman, A. E. & Kourtzi, Z. Learning predictive statistics from temporal sequences: Dynamics and strategies. *J. Vis.* **17,** 1 (2017).
4. Calhoun, V. D., Adali, T., Pearlson, G. D. & Pekar, J. J. Functional neuroanatomy of visuo-spatial working memory in turner syndrome. *Hum. Brain Mapp.* **14,** 96–107 (2001).
5. Vergara, V. M., Mayer, A., Damaraju, E., Hutchison, K. & Calhoun, V. The effect of preprocessing pipelines in subject classification and detection of abnormal resting state functional network connectivity using group ICA. *Neuroimage* (2016). doi:10.1016/j.neuroimage.2016.03.038
6. Calhoun, V. D., Liu, J. & Adali, T. A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. *Neuroimage* **45,** 163–172 (2009).
7. Shirer, W. R., Ryali, S., Rykhlevskaia, E., Menon, V. & Greicius, M. D. Decoding Subject-Driven Cognitive States with Whole-Brain Connectivity Patterns. *Cereb. Cortex* **22,** 158–165 (2011).
8. Allen, E. A. *et al.* A baseline for the multivariate comparison of resting-state networks. *Front. Syst. Neurosci.* **5,** 2 (2011).
9. Chen, Z. & Calhoun, V. Effect of Spatial Smoothing on Task fMRI ICA and Functional Connectivity. *Front. Neurosci.* **12,** (2018).
10. Andersson, J. L. R., Skare, S. & Ashburner, J. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *Neuroimage* **20,** 870–888 (2003).
11. Alexander, G. E., DeLong, M. R. & Strick, P. L. Parallel Organization of Functionally Segregated Circuits Linking Basal Ganglia and Cortex. *Annu. Rev. Neurosci.* **9,** 357–381 (1986).
12. Seger, C. A. The Involvement of Corticostriatal Loops in Learning Across Tasks, Species, and Methodologies. in *The basal ganglia IX* 25–39 (Springer, 2009). doi:10.1007/978-1-4419-0340-2_2
13. Tzourio-Mazoyer, N. *et al.* Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. *Neuroimage* **15,** 273–289 (2002).
14. Behrens, T. E. J. *et al.* Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nat. Neurosci.* **6,** 750–757 (2003).
15. Andersson, J. L. R. & Sotiropoulos, S. N. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* **125,** 1063–1078 (2016).
16. Jones, D. K. & Cercignani, M. Twenty-five pitfalls in the analysis of diffusion MRI data. *NMR Biomed.* **23,** 803–820 (2010).
17. Leemans, A. & Jones, D. K. The B -matrix must be rotated when correcting for subject motion in DTI data. *Magn. Reson. Med.* **61,** 1336–1349 (2009).
18. Behrens, T. E. J. *et al.* Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magn. Reson. Med.* **50,** 1077–1088 (2003).
19. Smith, S. M., De Stefano, N., Jenkinson, M. & Matthews, P. M. Normalized accurate measurement of longitudinal brain change. *J. Comput. Assist. Tomogr.* **25,** 466–75 (2001).
20. Thomas, C. & Baker, C. I. Teaching an adult brain new tricks: A critical review of

evidence for training-dependent structural plasticity in humans. *Neuroimage* **73,** 225–236 (2013).
21. Behrens, T. E. J., Johansen-Berg, H., Jbabdi, S., Rushworth, M. F. S. & Woolrich, M. W. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage* **34,** 144–155 (2007).
22. Tomassini, V. *et al.* Diffusion-Weighted Imaging Tractography-Based Parcellation of the Human Lateral Premotor Cortex Identifies Dorsal and Ventral Subregions with Anatomical and Functional Specializations. *J. Neurosci.* **27,** 10259–10269 (2007).
23. Rousseeuw, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20,** 53–65 (1987).
24. Patel, A. X. *et al.* A wavelet method for modeling and despiking motion artifacts from resting-state fMRI time series. *Neuroimage* **95,** 287–304 (2014).
25. Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S. J. & Turner, R. Movement-Related effects in fMRI time-series. *Magn. Reson. Med.* **35,** 346–355 (1996).
26. Van Dijk, K. R. a *et al.* Intrinsic Functional Connectivity As a Tool For Human Connectomics: Theory, Properties, and Optimization. *J. Neurophysiol.* **103,** 297–321 (2010).
27. Murphy, K. & Fox, M. D. Towards a Consensus Regarding Global Signal Regression for Resting State Functional Connectivity MRI. *Neuroimage* 0–1 (2016). doi:10.1016/j.neuroimage.2016.11.052
28. McKeown, M. J. *et al.* Analysis of fMRI data by blind separation into independent spatial components. *Hum. Brain Mapp.* **6,** 160–188 (1998).
29. Rissanen, J. Modeling by shortest data description. *Automatica* **14,** 465–471 (1978).
30. Himberg, J., Hyvärinen, A. & Esposito, F. Validating the independent components of neuroimaging time series via clustering and visualization. *Neuroimage* **22,** 1214–1222 (2004).
31. Ma, L., Narayana, S., Robin, D. A., Fox, P. T. & Xiong, J. Changes occur in resting state network of motor system during 4weeks of motor skill learning. *Neuroimage* **58,** 226–233 (2011).
32. Stevens, M. C., Kiehl, K. A., Pearlson, G. & Calhoun, V. D. Functional neural circuits for mental timekeeping. *Hum. Brain Mapp.* **28,** 394–408 (2007).
33. Laird, A. R. *et al.* Behavioral interpretations of intrinsic connectivity networks. *J. Cogn. Neurosci.* **23,** 4022–37 (2011).
34. Arbabshirani, M. R. *et al.* Impact of autocorrelation on functional connectivity. *Neuroimage* **102,** 294–308 (2014).
35. Murphy, K., Birn, R. M. & Bandettini, P. A. Resting-state fMRI confounds and cleanup. *Neuroimage* **80,** 349–359 (2013).
36. Johansen-Berg, H. & Rushworth, M. F. S. Using Diffusion Imaging to Study Human Connectional Anatomy. *Annu. Rev. Neurosci.* **32,** 75–94 (2009).
37. Kruskal, J. B. On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem. *Proc. Am. Math. Soc.* **7,** 48 (1956).
38. Bassett, D. S. *et al.* Hierarchical Organization of Human Cortical Networks in Health and Schizophrenia. *J. Neurosci.* **28,** 9239–9248 (2008).
39. Rubinov, M. & Sporns, O. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* **52,** 1059–1069 (2010).
40. Magalhães, R. *et al.* The impact of normalization and segmentation on resting state brain networks. *Brain Connect.* **5,** 1–28 (2015).
41. Razlighi, Q. R. *et al.* Unilateral disruptions in the default network with aging in native space. *Brain Behav.* **4,** 143–157 (2014).
42. van den Heuvel, M. P., Stam, C. J., Boersma, M. & Hulshoff Pol, H. E. Small-world

and scale-free organization of voxel-based resting-state functional connectivity in the human brain. *Neuroimage* **43,** 528–539 (2008).
43. Friston, K. J., Rotshtein, P., Geng, J. J., Sterzer, P. & Henson, R. N. A critique of functional localisers. *Neuroimage* **30,** 1077–1087 (2006).
44. Bassett, D. S., Nelson, B. G., Mueller, B. A., Camchong, J. & Lim, K. O. Altered resting state complexity in schizophrenia. *Neuroimage* **59,** 2196–2207 (2012).
45. Shen, X., Tokoglu, F., Papademetris, X. & Constable, R. T. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *Neuroimage* **82,** 403–415 (2013).
46. Fan, L. *et al.* The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture. *Cereb. Cortex* **26,** 3508–3526 (2016).
47. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature* **393,** 440–442 (1998).
48. Freeman, L. C. A Set of Measures of Centrality Based on Betweenness. *Sociometry* **40,** 35 (1977).
49. Latora, V. & Marchiori, M. Efficient Behavior of Small-World Networks. 3–6 (2001). doi:10.1103/PhysRevLett.87.198701
50. Román, F. J. *et al.* Enhanced structural connectivity within a brain sub-network supporting working memory and engagement processes after cognitive training. *Neurobiol. Learn. Mem.* **141,** 33–43 (2017).
51. Heitger, M. H. *et al.* Motor learning-induced changes in functional brain connectivity as revealed by means of graph-theoretical network analysis. *Neuroimage* **61,** 633–650 (2012).
52. Soto, F. A., Bassett, D. S. & Ashby, F. G. Dissociable changes in functional network topology underlie early category learning and development of automaticity. *Neuroimage* **141,** 220–241 (2016).

Supplementary Tables

Supplementary Table 1: Striatal segments. Four striatal segments for each hemisphere were estimated by a DTI connectivity-based and hypothesis-free classification method. The size of the segments and the MNI coordinates of their center of gravity are shown.

Hemisphere	Name	voxels	Center of gravity		
			x	y	z
Left	ventral striatum	102	-13	13	-9
	caudate head, anterior putamen	117	-16	14	-1
	caudate body/tail	120	-16	7	13
	posterior putamen	208	-27	-1	5
Right	ventral striatum	99	14	13	-8
	caudate head, anterior putamen	126	17	15	-1
	caudate body/tail	129	14	6	15
	posterior putamen	197	27	1	4

Supplementary Table 2: ICA components. Clusters within the 7 selected components are extracted from the group maps (z=1.96, p=0.05) and are organized into known functional groups[7,33]. The table shows the number of voxels within each cluster (clusters smaller than 20 voxels are not included), the MNI coordinates, the label of the corresponding AAL area and the t-statistic of the peak voxel.

Network	Component	Cluster	voxels	x	y	z	t-value
Executive	CP_9 (RCEN)	R MFG	718	39	23	50	3.87
		R IPL	477	48	-49	54	4.64
		L Cerebellum	39	-36	-70	-42	2.61
		R Cingulate	38	3	35	38	3.01
		R MTG	27	66	-25	-10	2.23
	CP_14 (LCEN)	L IFG triangular	510	-51	17	30	4.55
		L IPL	413	-33	-70	50	3.81
		L MFG	55	-27	17	58	2.8
		L MTG	47	-60	-49	-10	2.46
		L SFG medial	25	-3	29	42	2.71
Motor	CP_4 (Sensorimotor)	R SMA	853	0	-22	58	3.92
	CP_5 (Lateral Motor)	R Postcentral	368	51	-25	54	3.55
		L Postcentral	330	-51	-31	54	3.8
Visual	CP_2 (Secondary)	R MOG	726	33	-82	22	3.42
		L MOG	406	-24	-88	22	2.88
	CP_12 (Early)	R Calcarine	606	12	-97	-2	3.39
Motivational	CP_15 (ACC)	R ACC	620	0	44	-2	4.38

Supplementary Table 3: Intrinsic and extrinsic connectivity correlations with strategy index. Semipartial Pearson skipped correlations are reported for (a) intrinsic connectivity change (post minus pre-training) and (b) extrinsic connectivity change with strategy index for frequency and context-based statistics. Significant correlations are determined based on bootstrapped confidence intervals (CI) and denoted in bold. The r-value and 95% CI are shown for each statistical test (n=21).

a. Intrinsic connectivity analysis

Network	frequency statistics		context-based statistics	
	r	CI	r	CI
ACC	0.12	[-0.32, 0.51]	**0.35**	**[0.04, 0.63]**
RCEN	-0.17	[-0.61, 0.33]	-0.16	[-0.57, 0.33]
LCEN	-0.01	[-0.39, 0.41]	**0.42**	**[0.01, 0.68]**
Secondary Visual	-0.09	[-0.43, 0.29]	**-0.49**	**[-0.74, -0.10]**
Early Visual	-0.32	[-0.73, 0.16]	-0.03	[-0.44, 0.40]
Sensorimotor	0.20	[-0.13, 0.53]	0.23	[-0.22, 0.59]
Lateral Motor	**0.77**	**[0.60, 0.89]**	-0.07	[-0.50, 0.39]

b. Extrinsic connectivity analysis

Cortico-striatal pathways	frequency statistics		context-based statistics	
	r	CI	r	CI
ACC - right ventral striatum	-0.09	[-0.45, 0.28]	-0.15	[-0.43, 0.12]
ACC - left ventral striatum	-0.31	[-0.65, 0.12]	-0.14	[-0.53, 0.27]
RCEN - right caudate head, anterior putamen	-0.05	[-0.40, 0.36]	0.13	[-0.26, 0.42]
RCEN - left caudate head, anterior putamen	0.34	[-0.03, 0.66]	-0.14	[-0.41, 0.10]
LCEN - right caudate head, anterior putamen	0.17	[-0.31, 0.52]	0.22	[-0.19, 0.52]
LCEN - left caudate head, anterior putamen	0.03	[-0.34, 0.40]	0.01	[-0.35, 0.33]
Secondary Visual - right caudate body/tail	0.15	[-0.38, 0.57]	0.38	[-0.09, 0.72]
Secondary Visual - left caudate body/tail	0.19	[-0.25, 0.56]	0.21	[-0.28, 0.58]
Early Visual - right caudate body/tail	-0.04	[-0.50, 0.41]	0.05	[-0.41, 0.45]
Early Visual - left caudate body/tail	-0.19	[-0.60, 0.25]	**-0.46**	**[-0.83, -0.13]**
Sensorimotor - right posterior putamen	-0.14	[-0.49, 0.26]	0	[-0.35, 0.35]
Sensorimotor - left posterior putamen	0.01	[-0.55, 0.45]	0.03	[-0.37, 0.43]
Lateral Motor - right posterior putamen	**0.51**	**[0.20, 0.74]**	-0.19	[-0.59, 0.29]
Lateral Motor - left posterior putamen	0.13	[-0.41, 0.65]	0.03	[-0.50, 0.46]

Supplementary Table 4: PLS weights of the first two components: for (a) predictors and (b) response variables. Asterisks denote significant weights (|z|>2.576, p=0.01).

a. Weights for predictors

Node	Graph metric	PLS-1		PLS-2	
		rs-fMRI	DTI	rs-fMRI	DTI
L Caudate	Degree	1.79	-0.97	0.64	-2.84*
L Caudate	Clustering	1.18	1.05	-0.22	3.99*
R Caudate	Degree	2.30	-0.89	0.77	3.21*
R Caudate	Clustering	2.07	-0.10	0.03	-0.66
L Putamen	Degree	1.78	4.60*	1.38	-0.67
L Putamen	Clustering	0.29	-2.13	0.96	1.37
R Putamen	Degree	1.35	-2.06	0.31	0.34
R Putamen	Clustering	-0.40	-0.03	1.24	-0.27
R MFG	Degree	0.41	-0.22	0.39	2.67*
R MFG	Clustering	-1.92	-1.94	-0.49	-0.49
L IFG triangular	Degree	2.83*	1.50	0.11	1.24
L IFG triangular	Clustering	1.72	2.05	-0.57	1.32
L Postcentral	Degree	-1.86	-2.01	-1.69	-0.90
L Postcentral	Clustering	0.20	2.66*	-1.38	-0.44
R Postcentral	Degree	-0.74	0.15	-1.11	-0.69
R Postcentral	Clustering	-1.15	-1.71	-1.24	0.65
L Calcarine	Degree	-0.39	1.46	-0.23	-1.64
L Calcarine	Clustering	0.95	0.50	1.96	0.64
R Calcarine	Degree	0.40	3.58*	-0.67	0.02
R Calcarine	Clustering	-1.04	-1.67	2.18	-0.95
L ACC	Degree	0.39	-0.27	1.38	3.67*
L ACC	Clustering	0.34	-0.52	2.84*	1.12
R ACC	Degree	-0.18	2.16	2.55	1.21
R ACC	Clustering	-0.56	-3.45*	1.44	-0.30

b. Weights for response variables

Behavior	PLS-1	PLS-2
Strategy 0	-2.85*	2.01
Strategy 1&2	3.28*	2.47

Supplementary Table 5: **PLS results across graph metrics**. Pearson correlation of the first two PLS components between models (Model-1 is the reference model for the comparisons).

Model comparison	PLS-1	PLS-2
Model-2 vs. Model-1	r=0.94, CI=[0.81, 0.98]	r=0.89, CI=[0.75, 0.95]
Model-3 vs. Model-1	r=0.88, CI=[0.58, 0.97]	r=0.86, CI=[0.66, 0.96]
Model-4 vs. Model-1	r=0.99, CI=[0.96, 0.99]	r=0.98, CI=[0.94, 0.99]
Model-5 vs. Model-1	r=0.95, CI=[0.90, 0.98]	r=0.93, CI=[0.82, 0.97]
Model-6 vs. Model-1	r=0.92, CI=[0.80, 0.97]	r=0.89, CI=[0.73, 0.97]
Model-7 vs. Model-1	r=0.98, CI=[0.92, 0.99]	r=0.97, CI=[0.90, 0.99]

a

Table A		Target			
		A	B	C	D
Context	A	0	0.8	0.2	0
	B	0	0	0.8	0.2
	C	0.2	0	0	0.8
	D	0.8	0.2	0	0

Table B		Target			
		A	B	C	D
Context	A	0	1	0	0
	B	0	0	1	0
	C	0	0	0	1
	D	1	0	0	0

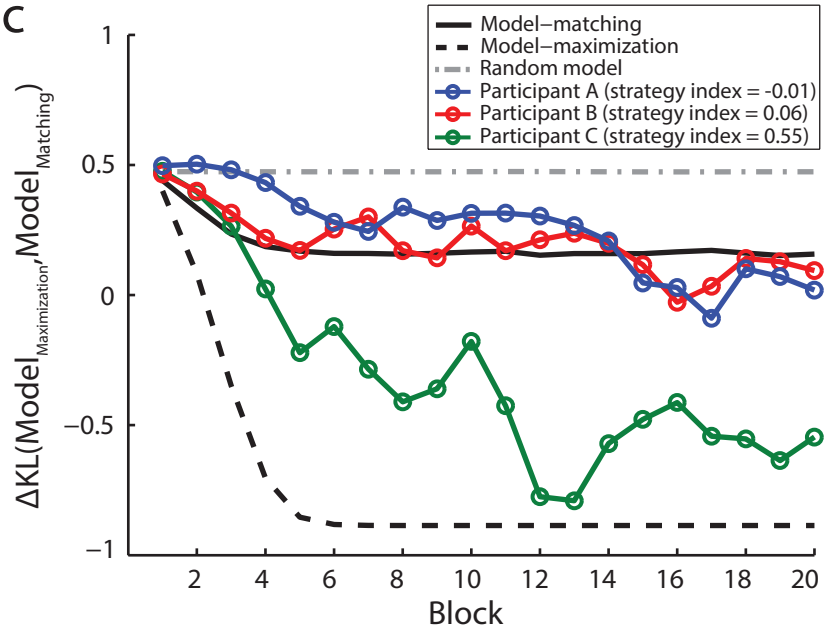
Table C		Target			
		A	B	C	D
Context	A	0.25	0.25	0.25	0.25
	B	0.25	0.25	0.25	0.25
	C	0.25	0.25	0.25	0.25
	D	0.25	0.25	0.25	0.25

b

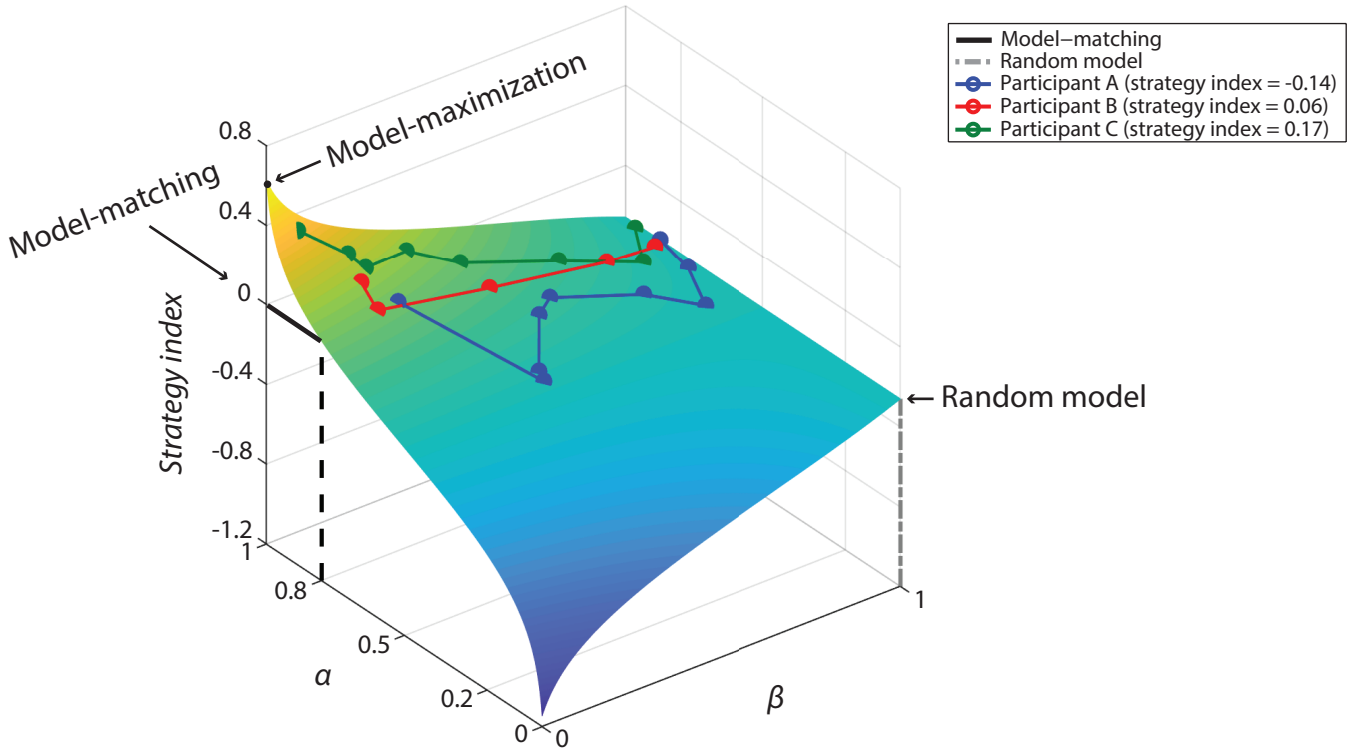
Table D		Target			
		A	B	C	D
Context	A	0	0.7	0.3	0
	B	0	0	0.7	0.3
	C	0.3	0	0	0.7
	D	0.7	0.3	0	0

Table E		Target			
		A	B	C	D
Context	A	0	0.9	0.1	0
	B	0	0	0.9	0.1
	C	0.1	0	0	0.9
	D	0.9	0.1	0	0

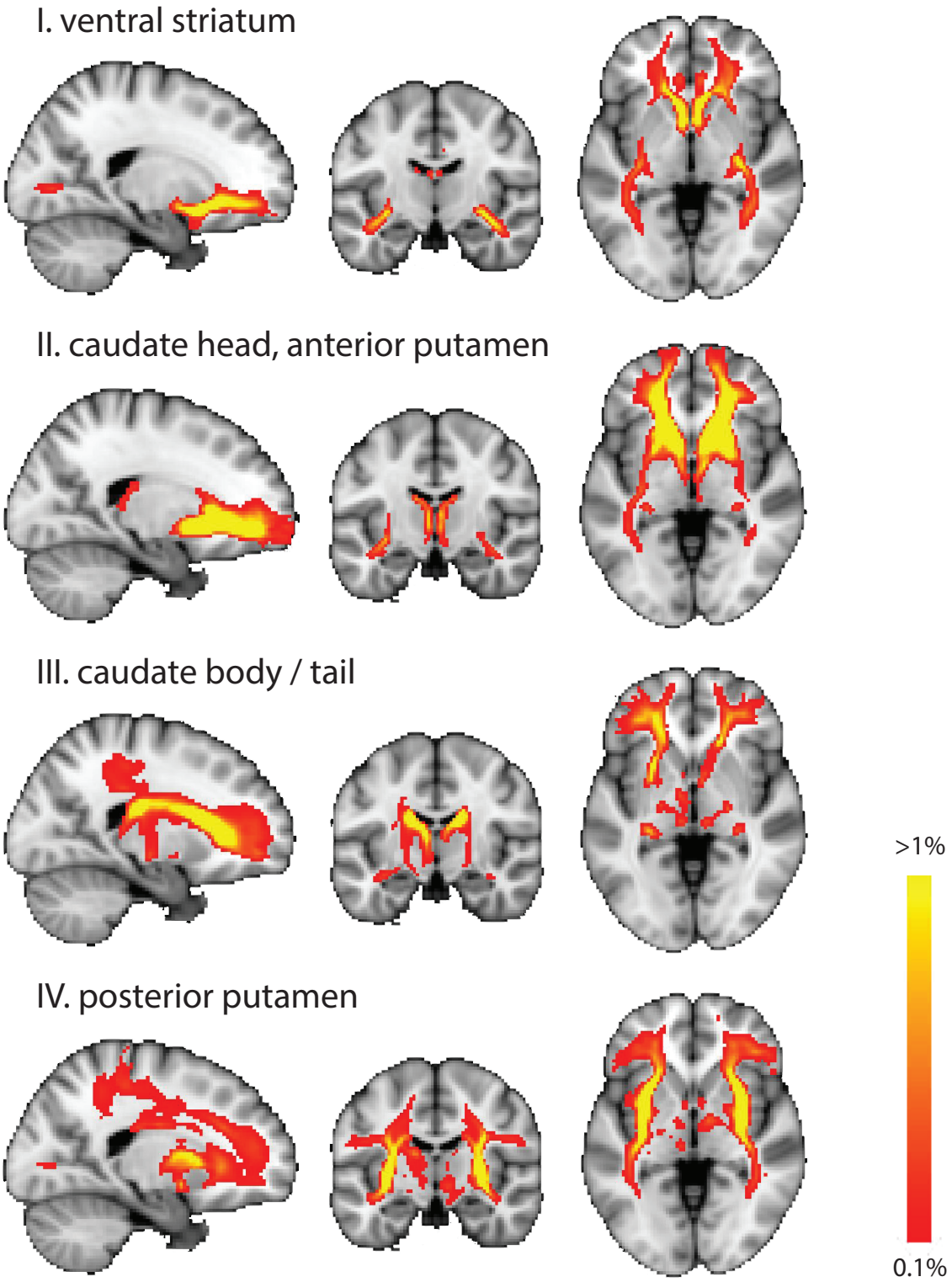
c



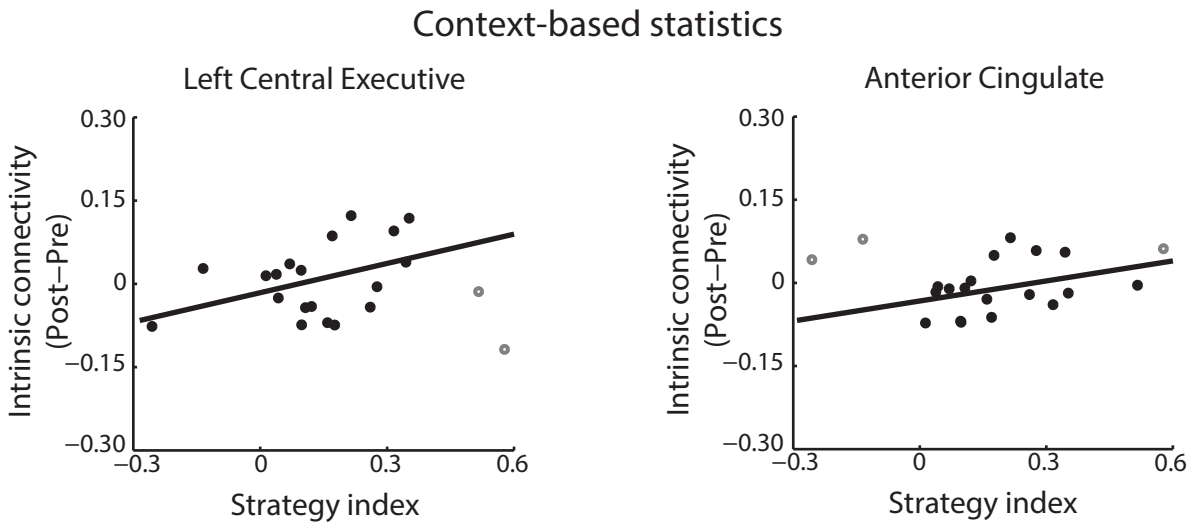
Supplementary Figure 1: Examples of participant responses for level-1 sequences. (a) Response tables for model-matching (Table A), model-maximization (Table B) and a random model (i.e. equal responses to all context-target contingencies; Table C). (b) Table D shows example responses for underestimating the probability of the most probable contingency (i.e. responses between random and model-matching). Table E shows example responses for overestimating the probability of the most probable contingency (i.e. responses between model-matching and model-maximization). (c) Participant strategy choice across training blocks for three representative participants (blue: participant A; red: participant B; green: participant C) against the three models (solid black line: model-matching; dashed black line: model-maximization; dashed gray line: random model). We computed the strategy index as the integral between the values of participant strategy choice and the model-matching across blocks.



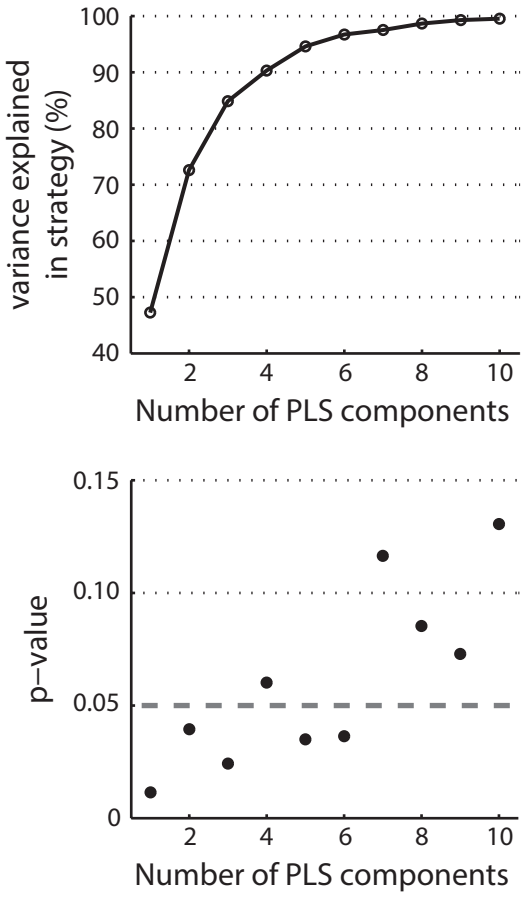
Supplementary Figure 2: Two-parameter model of participant response distribution. The surface of a two-parameter model depicted here describes the strategy index of a virtual participant as a function of α and β (P = β [0.25, 0.25, 0.25, 0.25] + (1-β) [1-α, α, 0, 0]). α describes participant preference for the more over the less probable target: (i) α=1 indicates maximization, (ii) α=0.8 indicates matching, (iii) α=0.5 indicates equal responses to the two possible targets, (iv) α<0.5 indicates participant preference of the less probable target. β describes participant preference for the random model: (i) β=1 indicates random model of responses (i.e. equal responses for all targets), (ii) β=0 indicates no random responses (i.e. the model is described by the probabilities of the two probable targets). Colder colors (e.g. blue) denote lower strategy index values, whereas warmer colors (e.g. yellow) denote higher strategy index values. Individual data of three representative participants are displayed as walks on the surface (blue: participant A; red: participant B; green: participant C). Individual data points start from the right (i.e. β≈1) and deviate towards the left of the surface (i.e. β≈0) showing three distinct behaviors: participant A underestimates the highly probable targets (i.e. negative strategy index close to matching), participant B matches the target distribution (i.e. zero strategy index close to matching) and participant C overestimates the highly probable targets (i.e. positive strategy index close to maximization).



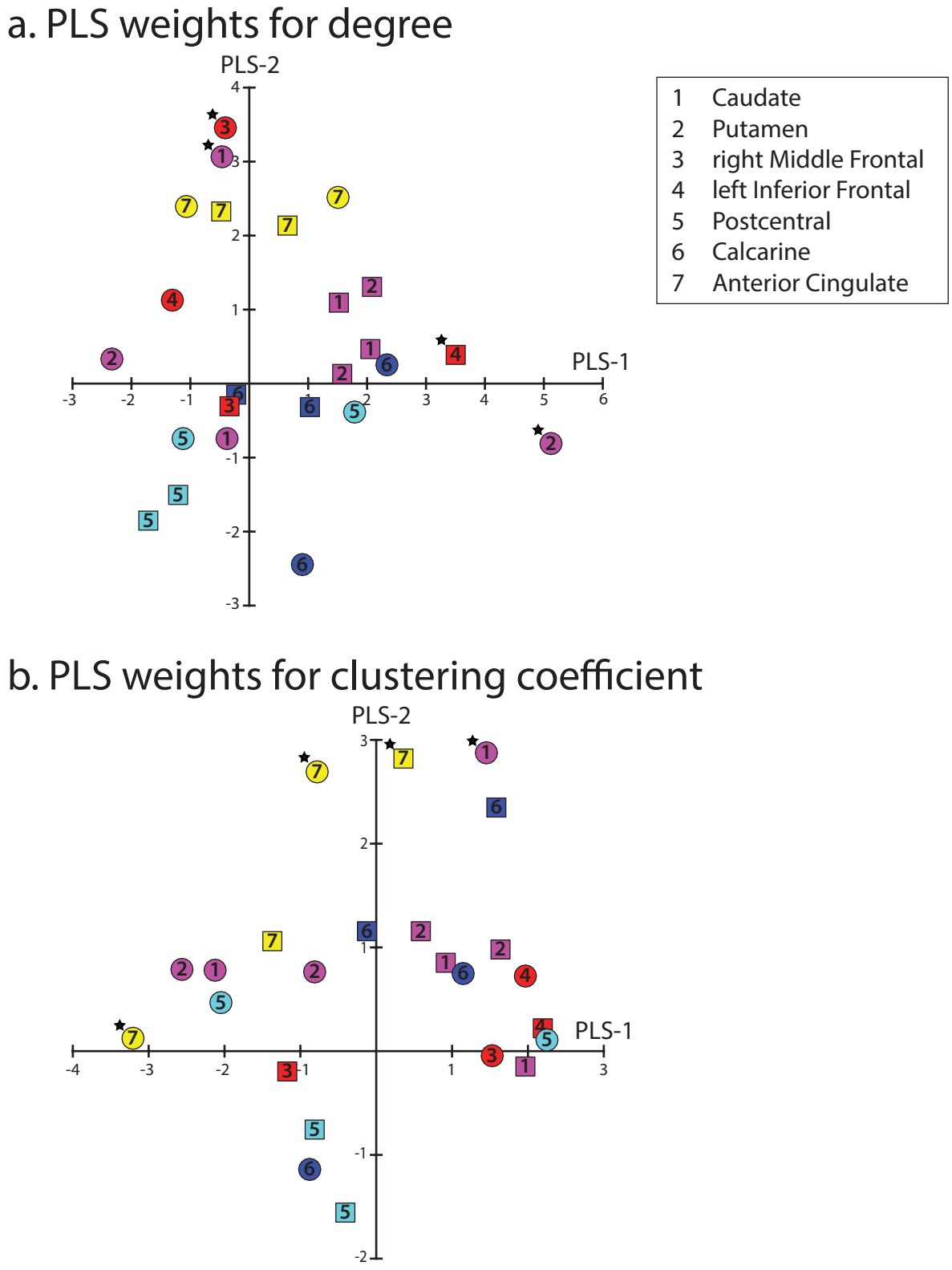
Supplementary Figure 3: DTI tractography for striatal segmentation. Striatal segments were estimated using a DTI connectivity-based and hypothesis-free classification method. Connection probability maps are displayed for each segment on the MNI template (neurological convention: left is left). Maps are thresholded at 0.1% of total tracts and averaged across groups and sessions. Whole brain tractography was computed separately for the left and right hemisphere and the maps were combined for visualization purposes (x=-20, y=-12, z=-2).



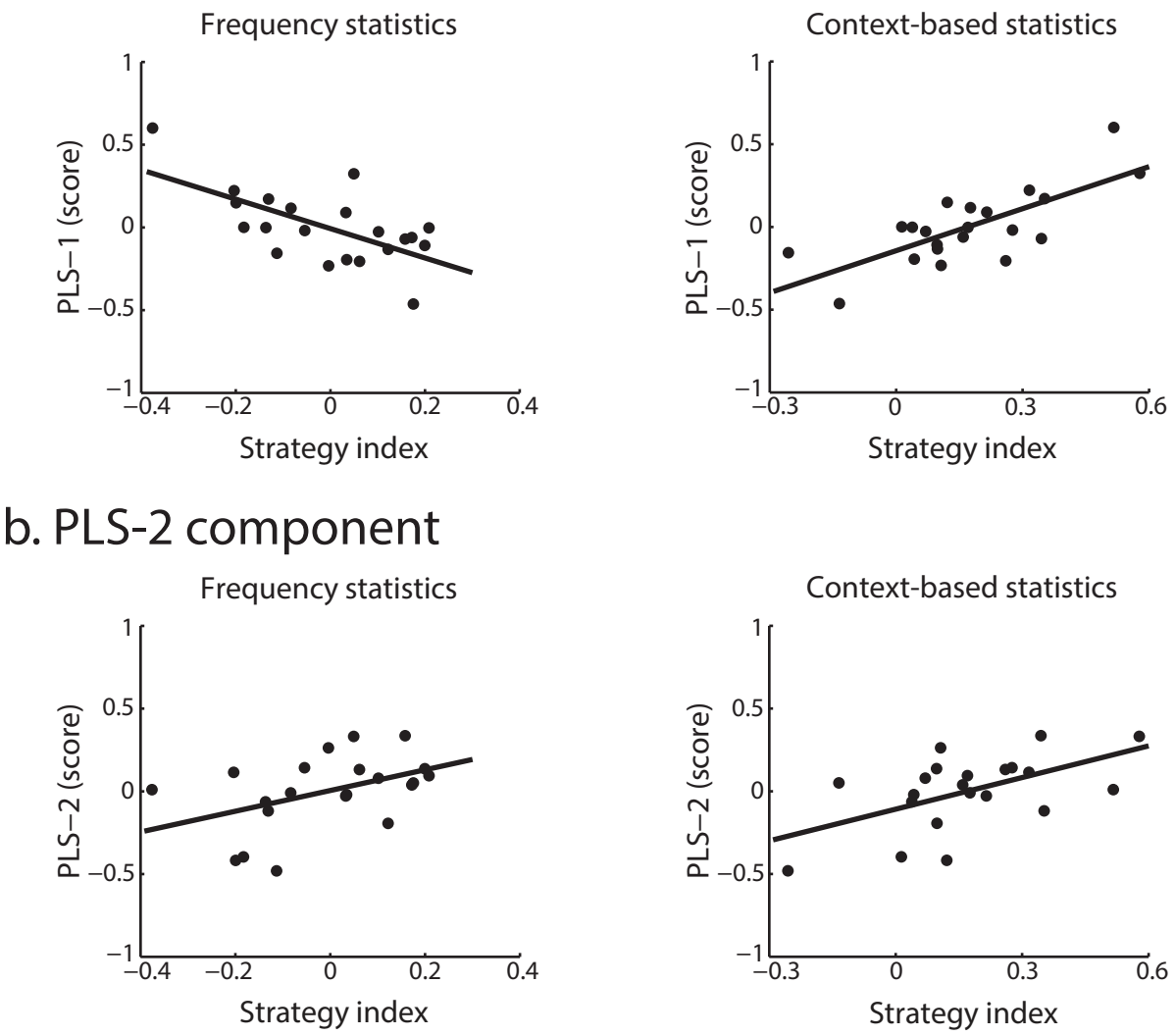
Supplementary Figure 4: Intrinsic connectivity analysis – supplementary results. Skipped Pearson correlations (two-sided, n=21) showed a magically significant relationship of intrinsic connectivity change (post- minus pre-training) in the Left Central Executive (LCEN) and Anterior Cingulate (ACC) networks with strategy index for frequency statistics (LCEN: r(19)=0.42, p=0.059, CI=[0.01, 0.68]; ACC: r(19)=0.35, p=0.121, CI=[0.04, 0.63]). Open circles in the correlation plots denote outliers as detected by the Robust Correlation Toolbox. Intrinsic connectivity was positive for all participants and sessions (pre-training, post-training); therefore, the sign of the change (Post minus Pre) indicates an increase (if positive) or a decrease (if negative) in the connectivity. In all but 5 cases (3 for posterior putamen - Lateral Motor connectivity; 2 for caudate body/tail - Early Visual connectivity) extrinsic connectivity change (Post minus Pre) had the same sign as the absolute connectivity change (|Post| minus |Pre|). Therefore, we interpret these correlations based on the change of the actual connectivity values (that is, Post>Pre is interpreted as increased connectivity). Performing the extrinsic connectivity analysis using the absolute connectivity change (|Post| minus |Pre|) showed similar results. That is, we found a) increased connectivity between the right posterior putamen and the Lateral Motor network correlated positively with strategy index for frequency statistics (r(16)=0.62, p=0.006, CI=[0.38, 0.79]), b) increased connectivity between the left body/tail of caudate and the Early Visual network correlated negatively with strategy index for context-based statistics (r(16)=-0.38, p=0.120, CI=[-0.74, -0.02]).



Supplementary Figure 5: Goodness of fit of PLS regression. Top panel shows variance explained in the response variables as a function of PLS components. Bottom panel shows the significance of the PLS model as a function of PLS components. Significance was determined by permutation testing (10,000 permutations); p-values below 0.05 indicate significant results.



Supplementary Figure 6: PLS results across a range of density levels (from 10% to 30%). Scatterplot of PLS-1 and PLS-2 weights for change (i.e. post- minus pre-training) in (a) degree and (b) clustering coefficient. PLS predictor weights for each selected node are indicated by symbols separately for DTI (circles) and rs-fMRI (squares) data. The color of the symbols corresponds to nodes in cortico-striatal circuits (**Figure 5**): caudate and putamen (magenta), right MFG and left IFG (red), postcentral gyrus (cyan), calcarine sulcus (blue), and ACC (yellow). PLS predictor weights with |z|>2.576 (p=0.01) are marked by an asterisk to denote significant predictors for the respective PLS component.



Supplementary Figure 7: PLS components related to strategy index. Illustration of the first two PLS components in relation to strategy index for frequency and context-based statistics (n=21). (a) Scatterplot of PLS-1 score with strategy index showing opposite patterns for frequency vs. context-based statistics. (b) Scatterplot of PLS-2 score with strategy index showing a similar pattern for frequency and context-based statistics. Note that the scatterplots between PLS components and strategy index are shown here for illustration purposes only. No further statistics were conducted to avoid circularity, as these two PLS components were shown to be significant predictors of the strategy index (**Figure 7a, Supplementary Table 4b**).