Received: 1 June 2022 | Revised: 8 November 2022 | Accepted: 11 November 2022

DOI: 10.1002/hbm.26158

RESEARCH ARTICLE

WILEY

Changes in global functional network properties predict individual differences in habit formation

Xiaoyu Wang | Katharina Zwosta | Uta Wolfensteller | Hannes Ruge

Fakultät Psychologie, Technische Universität Dresden, Dresden, Germany

Correspondence
Xiaoyu Wang, Fakultät Psychologie, Technische Universität Dresden, 01062 Dresden, Germany.
Email: xiaoyu.wang3@tu-dresden.de

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Number: SFB 940 project

Abstract

Prior evidence suggests that sensorimotor regions play a crucial role in habit formation. Yet, whether and how their global functional network properties might contribute to a more comprehensive characterization of habit formation still remains unclear. Capitalizing on advances in Elastic Net regression and predictive modeling, we examined whether learning-related functional connectivity alterations distributed across the whole brain could predict individual habit strength. Using the leave-one-subject-out cross-validation strategy, we found that the habit strength score of the novel unseen subjects could be successfully predicted. We further characterized the contribution of both, individual large-scale networks and individual brain regions by calculating their predictive weights. This highlighted the pivotal role of functional connectivity changes involving the sensorimotor network and the cingulo–opercular network in subject-specific habit strength prediction. These results contribute to the understanding the neural basis of human habit formation by demonstrating the importance of global functional network properties especially also for predicting the observable behavioral expression of habits.

KEYWORDS
fMRI, functional connectivity, goal-directed behavior, habit, multivariate linear regression, sensorimotor network

1 | INTRODUCTION

Everyday life is full of situations that require behavioral adaptation in order to achieve certain novel goals and further to perform fluently and effortlessly. Initial human learning is often driven by performance feedback signaling whether or not a pursued goal was met. Thereby, we can incrementally learn, which actions lead to desired outcomes or which actions serve to avoid adverse outcomes under given situational context conditions. Hence, feedback-driven learning is an effective way to establish novel goal-directed behaviors. But only more extensive training ultimately enables us to handle everyday challenges effortlessly. According to many theoretical accounts (Daw et al., 2005; Dickinson, 1985; Dolan & Dayan, 2013), human behavior is initially goal-directed involving the anticipation of future outcomes such that a response to a certain stimulus is selected to either achieve a rewarding outcome or to avoid a punishing outcome. With increasing practice, however, behavior is assumed to become less and less governed by outcome anticipation and instead to become more and more purely stimulus-bound or habitual. Such habitual behavior is assumed to be controlled directly by stimulus–response (S-R) associations, rendering our behavior much faster and less resource demanding but also highly inflexible, which is adaptive in a stable environment, but maladaptive in a changing environment (Dickinson, 1985; Seger & Spiering, 2011). Habitual inflexibility is

indicated by reduced sensitivity to outcome de- or re-valuation (de Wit et al., 2012; Delorme et al., 2016; Dickinson, 1985; Ersche et al., 2016; Niv et al., 2006). That is, if a stimulus that habitually triggers a response X now requires a different response Y in order to achieve a certain goal, then, initially, there is a perseverative tendency to stick with response X. In contrast, if the original response X was still under goal-directed control, it could be de-activated more flexibly by re-coding the response-associated outcome representation from "desirable" or "correct" to "non-desirable" or "incorrect."

Importantly, existing research is still rather unclear about how quickly goal-habit transition evolves (de Wit et al., 2018; Pool et al., 2022; Seger & Spiering, 2011), and there are certainly several modulating factors involved. Here, we made the key assumption that interindividual differences would be a relevant factor especially with an intermediate training duration that is long enough to ensure some overtraining at asymptotic behavioral accuracy levels in all subjects, but also still short enough to prevent ceiling habit levels in all subjects.

To better understand how quickly the goal-directed behavior transitions into habitual behavior for different individuals, and how this might be related to interindividual differences in neural processes, our group has previously developed a paradigm, which could not only characterize the dynamic alterations in brain activity during training but also help us to examine whether individual differences in training-related activity changes could predict interindividual differences in subject-specific habit strength after devaluation of the habitual response outcomes (Zwosta et al., 2018). To this end, after habit training, subjects were explicitly instructed that a trained response (e.g. R1) performed upon a given stimulus (e.g., S1) would no longer produce the previously valid outcome (here: monetary gain or no monetary loss). Instead, in the subsequent test phase, a given stimulus (e.g., S1) from now on required one out of two responses (R1 or R2) depending on which of two different outcomes (here: different colorings) needed to be generated. Hence, this goal-directed test phase response could be compatible or incompatible with the previously trained habitual response and the difference between incompatible and compatible trials enabled us to assess how strongly the previously trained habits interfered with the selection of goal-directed actions.

Hence, our expectation was that at the end of the training phase (after approximately 100 trials per S-R link) some subjects would have made considerable progress in goal-habit transition while others would still rely more heavily on goal-directed control. Indeed, our previously published fMRI results demonstrated that the activation of angular gyrus (AG), a sub-region of inferior parietal cortex, and a main hub region of several brain networks (Buckner et al., 2008; Hagmann et al., 2008; Igelstrom & Graziano, 2017; Vincent et al., 2008), decreased significantly across training, and most importantly, stronger activity decrease predicted stronger individual habit strength after instructed outcome devaluation (Zwosta et al., 2018). Together with other study results that also implicated the AG in goal-directed action control (Desmurget et al., 2009; Liljeholm et al., 2011, 2015; Melcher et al., 2013; Zwosta et al., 2015), this finding suggests that a less pronounced activity decrease in the AG reflects a less pronounced goal-habit transition. In other words, some subjects might still have relied on goal-directed control (reflected by continued AG engagement) even after around 100 training trials. In turn, consistent with this assumption, these subjects were better able to follow the devaluation instruction and hence showed a less pronounced goal-habit compatibility effect as a proxy for habit strength. At the same time, however, we could not identify (components of) a putative habit-related brain network to be associated with the behavioral goal-habit compatibility effect. While we found increased training-related functional connectivity between posterior putamen and premotor cortex (PMC) as a potential indication of ongoing habit formation, which has been demonstrated in both non-human (Ashby et al., 2010; Yin et al., 2004) and human subjects (de Wit et al., 2012; Tricomi et al., 2009; Yin & Knowlton, 2006), neither this increased connectivity nor any other training-related activity changes in putative habit-related brain regions were predictive of individual habit strength as measured by the goal-habit compatibility effect (Zwosta et al., 2018).

In light of these previous findings, the primary aim of the present examination was to advance our understanding of the goal-habit transition process by taking a more holistic approach. Specifically, we reasoned that the neural basis of goal-habit transition might be more adequately characterized in terms of a widespread re-organization and optimization of information flow across the whole brain. If this were the case, isolated connectivity changes between individual pairs of brain regions, for instance, between PMC and putamen might just be a small part of a global neural re-organization process and might therefore fail to predict individual habit strength if considered in isolation. In contrast, using a multiple regression approach that considers the multitude of brain-wide connectivity changes as a whole might be a more adequate basis for the successful prediction of individual habit strength than assessing each pairwise functional connectivity measure independently in a mass-univariate manner. In fact, with the growing popularization of large-scale network approaches in the neuroscience field, emerging human learning research has already demonstrated that alterations in local neural activity are accompanied by the whole brain functional connectivity reorganization (Bassett et al., 2011, 2013, 2015; Cole et al., 2013; Mohr et al., 2016), which exhibits a significant predictive role in subject-specific learning ability (Bassett et al., 2011, 2013, 2015; Braun et al., 2015; Mohr et al., 2016).

Our analysis proceeded in the following steps: First, we extracted the mean time series from the 333 cortical regions defined in the Gordon atlas (Gordon et al., 2016) as well as 14 subcortical regions using the Harvard-Oxford subcortical atlas and then computed the learning-related functional connectivity changes among all these sub-regions. Second, we filtered the training-induced functional connectivity changes and then fed the selected features into the Elastic Net model. Using a leave-one-subject-out cross-validation strategy, this machine learning multiple regression approach was employed to predict the behavioral marker of individual habit strength. Finally, to better characterize the contribution of individual network components, we calculated the predictive weights of each functional brain network as well as the associated single brain regions in habit strength prediction.

2 | MATERIALS AND METHODS

2.1 | Subjects

Data of the goal-habit learning task have been published before with a different focus (Zwosta et al., 2018). After excluding three subjects due to excessive head movement (see Section 2.4 for more details) and two subjects due to abnormal behavior performance (see Section 2.2 *phase 3* for more details) from further analyses, fMRI data of 48 subjects (27 females, 21 males; mean age: 23 years, range 19–31 years) were re-used in the current study. All subjects were right-handed, neurologically healthy and had normal or corrected vision including normal color vision. The experimental protocol was approved by the Ethics Committee of the Technische Universität Dresden. All subjects gave written informed consent prior to taking part in the experiment and they were compensated with 8€ per hour in addition to the money they gained during the experiment.

2.2 | Experimental procedure

The experimental paradigm consisted of three consecutive phases (Figure 1). During the first phase, which was performed outside of the MRI scanner, goal-directed behavior based on hierarchical S-R-O associations was established. During the second phase, which was performed inside the scanner, subjects were required to learn novel responses to gain monetary reward or to avoid monetary loss for a subset of the stimuli already used in Phase 1. Training was continued beyond asymptotic performance levels in order to further strengthen habitual S-R associations by overtraining. Finally, during the third phase, which was also performed inside the scanner, monetary

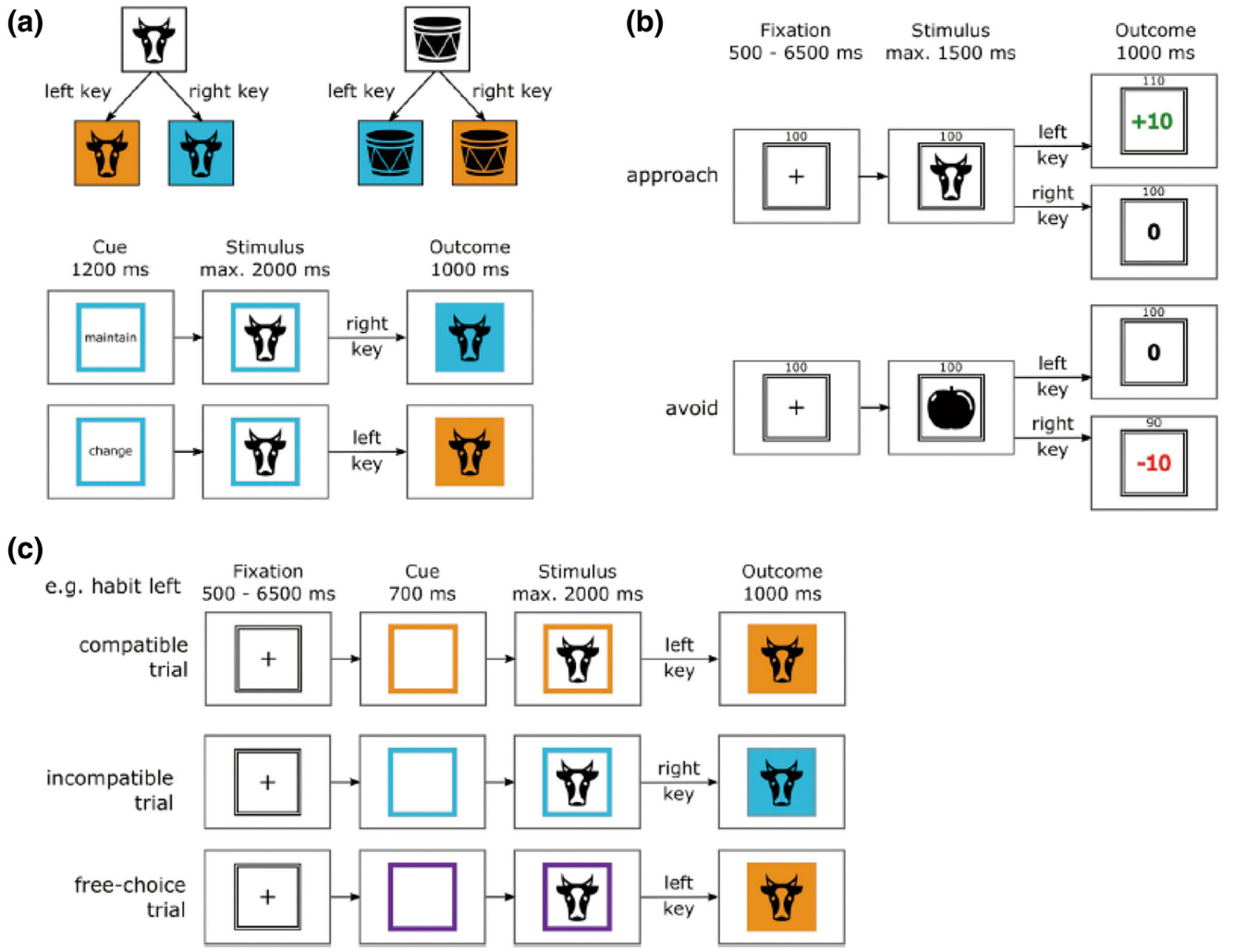

FIGURE 1 Experimental paradigm (for a detailed description, see main text). The experiment consisted of three consecutive phases. In Phase 1, goal-directed behavior was established, followed by Phase 2, which aimed at establishing habitual behavior through extensive training. In Phase 3, goal-directed responses established in Phase 1 were put into competition with responses trained in Phase 2 in order to test the individual habit strength. (a) Examples of the instructed hierarchical stimulus–response-outcome associations used in Phase 1 and two exemplary trials from Phase 1. (b) Exemplary trials from Phase 2. (c) Examples for compatible, incompatible, and free-choice trials in Phase 3

outcomes associated with habitual responses putatively established in Phase 2 were de-valuated by instruction, and habitual responding and goal-directed actions established in Phase 1 were put into competition to measure the habit strength developed in Phase 2. The whole experiment was controlled by E-Prime 2.0. In the current study, we utilized and analyzed fMRI data from Phase 2 and behavioral data from Phase 3.

Phase 1: The purpose of Phase 1 was to establish hierarchical S-(R-O) associations where the correct response upon a given stimulus category depended on the outcome to achieve. Hence, no stable associations between stimuli and responses (i.e., no S-R habits) could be learned. Ten stimuli were grouped into five "artificial" and five "natural" stimuli (natural stimuli: tree, snowflake, cow, mushroom, lungs; artificial stimuli: scissors, computer mouse, car, cupboard, ball). Responding to a stimulus from one group of stimuli (e.g., artificial) with the right key led to a blue outcome color and responding with the left key led to an orange outcome. This R-O association was inverted for the other group of stimuli (e.g., natural), such that pressing the right key led to an orange outcome and pressing the left key led to a blue outcome. Phase 1 comprised 240 trials and each trial started with the presentation of a cue containing either the German words for "change" or "maintain." This word was framed by a colored square displaying the present outcome color that was produced in the previous trial. The subjects' task was to press the key that would either change or maintain the current outcome color for the associated category the currently displayed stimulus belonged to. Our previously published study results demonstrated that the error rates and response times decreased across blocks, which indicated that subjects successfully acquired the R-O associations (Zwosta et al., 2018).

Phase 2: The purpose of phase 2 was to enable the formation of S-R habits for a subset of the stimuli already used in Phase 1. Eight of the 10 stimuli (four artificial and four natural stimuli) were re-used from Phase 1. At the beginning, subjects were instructed that the categories of the stimuli were now irrelevant and that they had to find out the correct key for each of the eight stimuli individually by trial and error. For both categories, each of the four stimuli belonging to one category was associated with one of the four combinations of correct responses (left or right) and outcome types (approach or avoidance). Subjects were also explicitly told that for four of the stimuli the correct response would allow them to gain points while for the other four stimuli the correct response would allow them to avoid losing points. Rewards were +10 points printed in green color, punishments were −10 points printed in red and outcomes of 0 points were printed in black. If subjects failed to execute any response during the response window, they also received the unfavorable outcome, that is, they lost 10 points ("−10") in avoidance trials and gained zero points ("0") in approach trials. Trials in Phase 2 were clustered into seven task blocks with 112 trials each (14 per stimulus). Hence, the whole Phase 2 consisted of 784 trials (98 per stimulus). Our previously published study results have shown that error rates and response times significantly decreased across blocks, $F(6,312) = 127.84$, $p < .001$, $\eta 2 = 0.71$ and $F(6,312) = 122.35$, $p < .001$, $\eta 2 = 0.70$, respectively (Zwosta et al., 2018).

Phase 3: At the beginning of Phase 3, subjects were instructed that they could no longer gain or lose any points and thereby removed the contingency between stimulus, response and monetary outcome (instructed outcome devaluation). Hence, any tendency to continue to perform the trained response established in Phase 2 should not be motivated by aiming to gain reward or avoid loss but should be based on habitual responding instead. Phase 3 had 384 trials and each trial started with a fixation cross, followed by a colored frame (cue) which was either one of the two outcome colors (blue and orange) previously introduced in phase 1, or a new third color (purple) indicating free-choice trials. If the frame was blue or orange then subjects were required to press the response that would lead to this particular outcome color for the displayed stimulus according to the R-O contingencies introduced in Phase 1 (goal-directed trials). If, however, the frame was purple then subjects could freely choose one of the two responses (free-choice trials). We were interested in two different trial categories: (1) trials for which the trained response toward the stimulus was identical to the required goal-directed response either because it had previously been rewarded or not punished (compatible condition, 96 trials in total). (2) Trials for which the trained response did not match the required goal-directed response (incompatible condition, 96 trials in total). The compatibility effect was considered as an indicator of habit strength, reflecting the impact of the trained habits on goal-directed behavior, and computed as the reaction time difference between correct incompatible and compatible trials. Since there was no difference between approach and avoidance conditions ($t(47) = 1.0172$, $p = .314$), we pooled compatibility effects across two conditions in further analysis to increase the statistical power. Compatibility effects of two subjects were identified as abnormal values (3 standard deviations above the mean) and hence discarded from the predictive analysis (see Section 2.8). Data of the remaining 48 subjects were normally distributed as assessed with d'Agostino-test ($p = .722$). Paired t-test was employed to assess the statistical significance of the compatibility effect, revealing that the reaction time in incompatible trials (mean RT = 814.515 ± 173.3 ms) was significantly higher than in compatible trials (mean RT = 792.141 ± 178.241 ms) with $t(47) = 4.3907$ ($p < .001$).

2.3 | fMRI scanning

MRI data were acquired on a 3 T Siemens whole body Trio System (Erlangen, Germany) equipped with a 32-channel head coil. Ear plugs dampened scanner noise. Structural images were acquired using a T1-weighted sequence (TR = 1900 ms, TE = 2.26 ms, T1 = 900 ms, flip = 9°) with a resolution of 1 mm × 1 mm × 1 mm. Functional images were acquired using a gradient echo planar sequence (TR = 2000 ms, TE = 30 ms, flip angle = 80°). Each volume contained 32 slices that were measured in ascending order. The voxel size was 4 × mm × 4 mm 4 × mm (gap: 20%). Only the fMRI data of experimental Phase 2 (goal-habit transition) were analyzed. Data of Phase 3 (goal-habit competition) were only used to compute the behavioral index of habit strength.

2.4 | fMRI preprocessing

Data preprocessing was performed with SPM12 running in Matlab 9.5. The initial three volumes were discarded to allow for the steady-state of the signal.

In order to assess the goal-habit transition during learning, the time series were cut into three parts (block1 and block2; block3, block4, and block5; block6 and block7) before preprocessing. The "early" learning phase (block1 and block2) and the "late" learning phase (block6 and block7) were utilized for further analysis. Preprocessing steps described below were applied to the early and late learning phase data independently (data of early and late phases were registered to their own first time point, respectively). Previous research has shown that even small amounts of head movements can substantially influence estimates of functional connectivity (Power et al., 2012, 2014; Van Dijk et al., 2012; Yan et al., 2013). In this study, frame-wise displacement over 0.2 mm was computed to identify spike events. Three subjects were excluded from further analyses since more than 20% of fMRI data samples in either the early or late learning phase were diagnosed as spike events.

Slice timing correction and motion correction were conducted. T1-weighted images were co-registered to the mean functional images and segmented into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). All functional images were then spatially normalized to the standard MNI space via the deformation fields derived from the tissue segmentation of the structural image (resampling to 3 mm resolution). SPM's a priori tissue probability maps (empirical thresholds: 90% for WM mask and 70% for CSF mask) were employed to create the average signals of WM and CSF, respectively (Yan et al., 2016).

After normalization, a general linear model (GLM) was run in order to regress out several nuisance variables. Several regressors were included in order to reduce the effects of motion and other noise confounds according to previous recommendations especially in the context of functional connectivity analyses (Geerligs et al., 2016). This included the original six motion parameters, average signals in WM, CSF masks, and their expansions. The expansions included the first-order temporal derivative, as well as their squares and squared derivatives. In addition, recent studies demonstrated that the global signal may represent an important confound in the search for individualized task-specific changes in functional connectivity analysis (Greene et al., 2018; Jangraw et al., 2018), which could substantially decrease the predictive model performance (Jangraw et al., 2018). Therefore, we also included the whole brain signal as a nuisance regressor in the GLM as suggested by previous research aiming to predict behavioral measures based on functional connectivity information (Cui et al., 2020; Cui & Gong, 2018; Dubois, Galdi, Han, et al., 2018; Dubois, Galdi, Paul, et al. 2018; Rosenberg et al., 2018; Rosenberg, Finn, et al., 2016; Rosenberg, Zhang, et al., 2016). In total, signal associated with 33 nuisance variables were regressed out through the GLM. Finally, the residual time series were spatially smoothed (6 mm FWHM).

2.5 | Task-activation regression for task-based functional connectivity

Previous research has shown that regressing out average task-related activity can reduce the spurious correlations between different brain regions (Cole et al., 2019), which can also improve the test–retest reliability in task-based functional connectivity analyses (Cao et al., 2014). We performed the single-subject GLM analysis to obtain the residual time series, which were then used for further functional connectivity computation. Learning trials were assigned to one of three categories: correct approach or avoidance and error trials. To appropriately capture BOLD activation, we used Fourier basis set regressors including 14 different sine-wave regressors spanning 30 s, which were time-locked to the onset of the learning trials (Cole et al., 2019). After that, breaks between task blocks were also included as regressors with an additional GLM, the break-related regressors were based on the standard hemodynamic response function of SPM12 and convolved with the duration of breaks which varied considerably. With each subject-specific GLM, the high-pass filter was set to a cutoff of 128 s in SPM12 and estimated with ordinary least squares (i.e., AR (1) off).

2.6 | Brain parcellation

The preprocessed time series were extracted from 347 predefined regions of interest (ROIs), including 333 cortical parcels (161 and 162 regions from the left and right hemispheres, respectively) associated with different functional networks as defined in the Gordon atlas (see Figure 2, cf. Van Essen et al., 2017). In addition, in order to ensure whole-brain coverage, as in previous study (Shine et al., 2016), we also included 14 additional subcortical regions from the Harvard-Oxford subcortical atlas including bilateral thalamus, caudate, putamen, globus pallidus, hippocampus, amygdala, and ventral striatum.

2.7 | Functional connectivity

Early and late task-related functional connectivity (FC) were estimated for each pairwise combination of the 347 ROIs (60,031 pairs in total) after averaging the signals across all voxels within each ROI. We then further calculated the changes in FC between the late and early learning phase (ΔFCs) among each pair of ROIs.

2.8 | Predictive modeling based on ΔFCs

We investigated the functional brain networks as well as the brain regions, whose connectivity changes contributed significantly to predicting the later behavioral habit strength. In the present study, habit strength was probed by a behavioral index that measures how strongly a newly acquired habit impacts the execution of goal-directed actions (habit-incompatible vs. compatible) after instructed

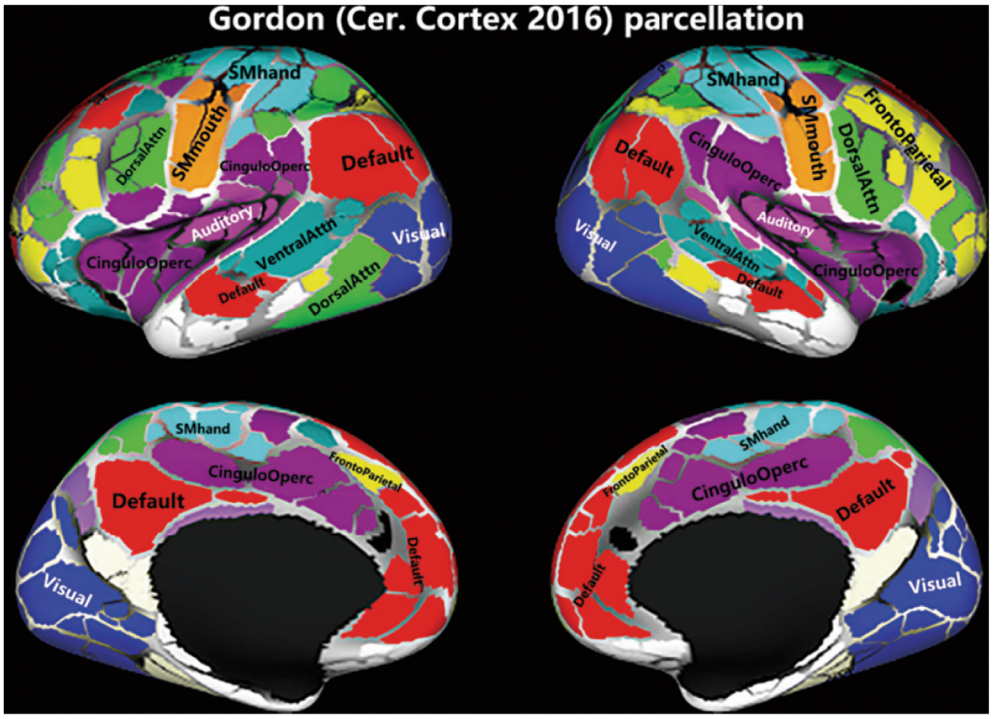

FIGURE 2 A diagram of the Gordon parcellation and how it is related to different large-scale functional networks (Gordon et al., 2016). *Source*: The figure was obtained from the website of the Brain Analysis Library of Spatial maps and Atlases (Van Essen et al., 2017)

devaluation of the habitual response outcomes in Phase 3. The connectivity changes, 60,031 features here, were used as the input feature-space, which leads to a large P—small N setting. We first used a univariate feature filtering approach to reduce the number of features and then entered the selected ΔFC-features into the Elastic Net model (Zou & Hastie, 2005) for individual habit strength prediction. Finally, we summed the absolute predictive weights across all features from each functional brain network as well as each single brain region, and investigated which network or brain region contributed most during the habit strength prediction.

2.8.1 | Feature selection

Due to our relatively small sample size ($N = 48$) and to obtain the highest possible estimates of regression accuracy, feature selection was performed using leave-one-subject-out cross-validation (LOOCV) (Lin et al., 2018; Plitt et al., 2015; Rosenberg et al., 2018; Rosenberg, Finn, et al., 2016). We set aside ΔFCs and behavioral data from one subject (novel left-out test subject) and performed the univariate Pearson correlation between each single ΔFC-feature and habit strength across the remaining 47 subjects. We discarded ΔFC-features for which the p value of the correlation with the behavioral score was greater than either .01 or .05 (Dubois, Galdi, Han, et al., 2018; Dubois, Galdi, Paul, et al., 2018; Greene et al., 2018; Kwak et al., 2021; Rosenberg et al., 2018). We also discarded ΔFC-features that correlated with head motion (see Section 2.9 for more details) (Rosenberg et al., 2018). This process was repeated 48 times each time with another subject left out of the training set, and the following predictive model training and testing were performed based on each LOOCV fold, as described next.

2.8.2 | Predictive model training and testing

Elastic Net regression modeling was applied here to predict novel individual habit strength based on ΔFCs across training. Although ridge regression normally shrinks coefficient values, it still keeps all the features and cannot produce a more parsimonious model. In contrast, lasso regression can create a sparse design matrix by setting coefficient values to exactly zero (Zou & Hastie, 2005). Here, we applied Elastic Net, which could select the optimal hyperparameter α (representing the weight of lasso vs. ridge optimization, with intermediate values representing Elastic Net optimization). The sparsity of the model can potentially be optimized to further maximize the accuracy of the predictive model, thereby avoiding any a-priory assumptions about the sparsity of the discriminative ground truth. For N observation pairs (x_i, y_i), Elastic Net solves the penalized residual sum of squares based on the user-defined α parameter, which is a compromise between ridge and lasso that ranges from 0 to 1. Note that it becomes the lasso when $\alpha = 1$ and the ridge regression when $\alpha = 0$. The loss function of the Elastic Net is:

$$\min_{(\beta_0,\,\beta)\,\in\,\mathbb{R}^{p+1}} \left[\frac{1}{2N}\sum_{i=1}^{N}\left(y_i-\beta_0-x_i^T\beta\right)^2+\lambda P_a(\beta)\right] \tag{1}$$

$$P_a(\beta)=(1-a)\frac{1}{2}\|\beta\|_{\ell_2}^2+\alpha\|\beta\|_{\ell_1} \tag{2}$$

$$= \sum_{j=1}^{p} \left[\frac{1}{2}(1-a)\beta_j^2 + \alpha|\beta_j| \right] \quad (3)$$

where λ is the tuning regularization parameter and P_a is the elastic-net penalty.

To investigate the optimal sparsity of the predictive model, we specified a set of α parameters which varied over 101 grid values ([0, 0.01, 0.02, ..., 1]). The grid range of the regularization parameter λ with maximally 100 values was generated based on the coordinate descent algorithm implemented in the glmnet package (http://hastie.su.domains/glmnet_matlab/). To achieve unbiased estimates, nested three-fold cross-validation (Kwak et al., 2021; Mumford et al., 2012; Plitt et al., 2015) was applied to select the optimal λ parameter. That is, for each LOOCV-fold comprising data of 47 subjects the dataset was randomly split into three similarly sized groups of which two were used for training while the remaining split of data was held back and used for assessing the performance of the prediction, thus optimizing the hyperparameters λ for each single individual α parameter. This whole procedure was repeated 100 times since the folds were selected at random, and the optimal combination of alpha and λ with the smallest average value of mean absolute error (MAE) across the 100 cross-validation runs was then obtained by the left-out testing subject. We repeated this procedure so that each subject was left out of the training set once and measured the models' predictive power by correlating the predicted and observed habit strength, controlling for motion (see Section 2.9). The significance of the correlation coefficient was determined using permutation tests within which each subject's actual habit strength was randomly permuted, and the same cross-validation described above was performed after each permutation. We performed 10,000 permutations to establish the empirical distribution of chance. The Elastic Net predictive analysis described above was performed using glmnet package in Matlab and analysis code is available here: https://github.com/xiaoyu-TUD/Elastic-Net.

2.8.3 | Predictive contribution of functional networks and single brain regions

In a final analysis step, we aimed to characterize the contribution of connectivity changes regarding specific brain networks and single brain regions for habit strength prediction. Since both, initial features selection and predictive model training and testing were defined using LOOCV, there were as many different models (48 unique models here) as there were rounds of cross-validation, each coming with a different set of selected features. To create a consistent set of features, we retained the features that were present in every round of cross-validation (Rosenberg et al., 2018; Rosenberg, Finn, et al., 2016). These common features were then used for characterizing the predictive contribution of brain networks and regions.

We then assigned the common predictive features to the associated brain network pairs ($13 \times 12/2 = 78$ pairs among 13 functional brain networks) and summed all the absolute predictive weights of each network pair for each LOOCV fold separately. The averaged summed value of each network pair across LOOCV folds was finally obtained. Thereafter, to investigate the predictive role of each single brain region, we summed the absolute predictive weights across all features from each single brain region for each LOOCV fold separately. One sample t-test based on the LOOCV-wise predictive brain maps was calculated and corrected for multiple comparison using FWE. The contribution of each single region for habit strength prediction was thereby expressed as the *t*-value.

2.9 | Head motion control

Since even small head motion can confound functional connectivity analyses, after regressing out head motion parameters, global signal, and excluding subjects showing large head motion (see Section 2.4 for details), we applied three different additional strategies to ensure that motion does not account for our current findings (Rosenberg et al., 2018; Rosenberg, Finn, et al., 2016).

First, we investigated whether the mean frame-wise displacement during learning was correlated with both, the observed and the predicted habit strength score. Second, we correlated each single ΔFC-feature with mean frame-wise displacement across subjects using Spearman's correlation before features selection, and excluded any features that were significantly (as in previous studies, p value $= .05$) correlated with motion across subjects (Rosenberg et al., 2018). Finally, we evaluated model prediction performance with partial correlations between observed and predicted habit strength, controlling for mean frame-wise displacement.

3 | RESULTS

3.1 | Habit strength prediction

After univariate feature selection and head motion-related feature reduction, predictive models trained on ΔFCs revealed that correlations between the predicted and observed behavioral habit strength r were significantly different from chance when using partial correlations, controlling for head motion, for both $p < .01$ feature selection threshold ($r = .539$, $p_{perm} < .001$) and $p < .05$ feature selection threshold ($r = .547$, $p_{perm} < .001$) (Figure 3). The additional head motion check demonstrated that the mean frame-wise displacement parameters did not correlate with both the observed ($r = .092$, $p = .532$) and predicted habit strength for both $p < .01$ ($r = .102$, $p = .492$) and $p < .05$ threshold ($r = .113$, $p = .447$). Since there was no significant improvement when considering more ΔFC-features for novel individualized habit strength prediction; in the following analysis, we therefore decided to apply the stricter threshold ($p < .01$) to exploit the contribution of connectivity changes from the functional brain networks and each single region in habit strength prediction.

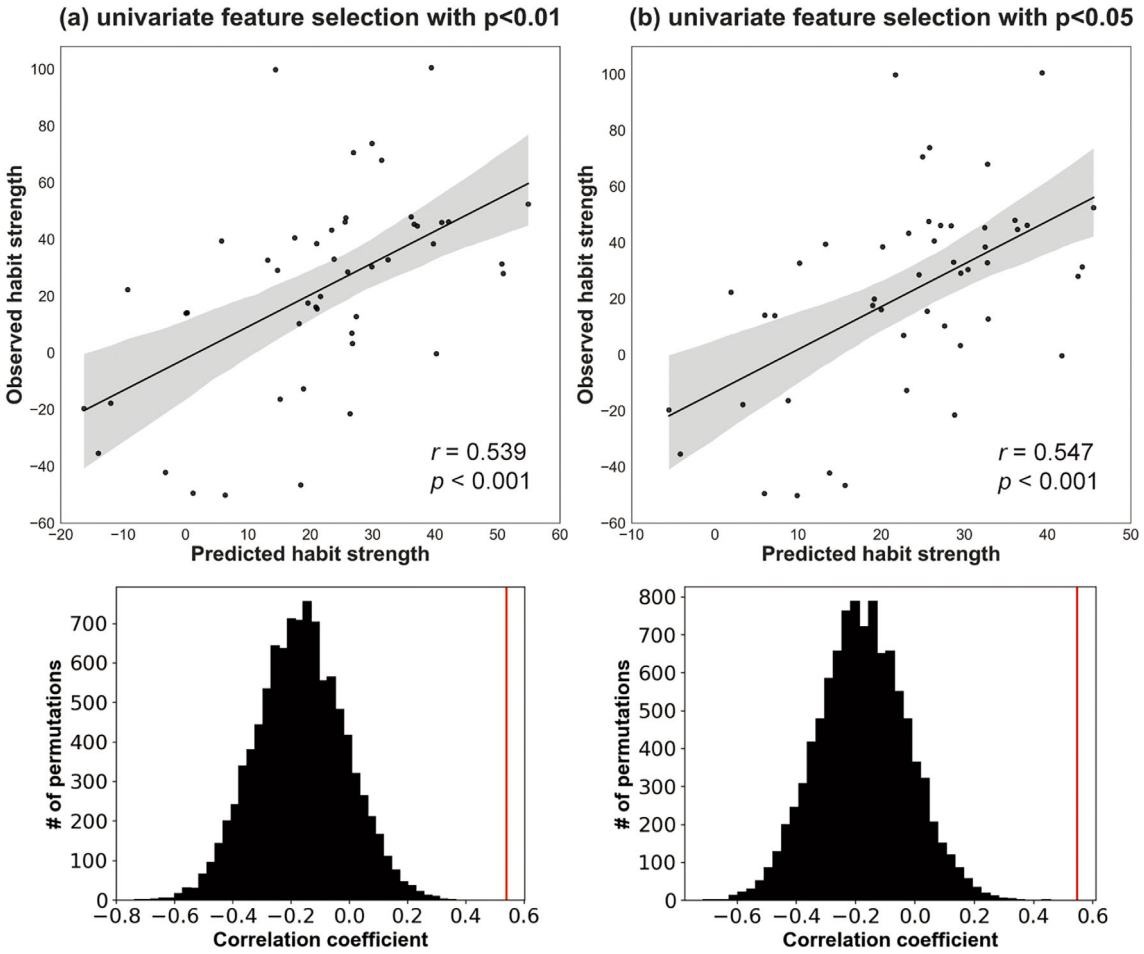

FIGURE 3 Partial correlations between the observed habit strength and habit strength predicted by learning-related ΔFCs, controlling for head motion. ΔFC-features were selected based on univariate Pearson correlation between habit strength and ΔFCs with $p < .01$ (a) or .05 (b)

3.2 | Predictive brain functional networks

The optimal α parameter of Elastic Net equaled 0.99 and 115 common ΔFC-features from 48 unique models were finally obtained, which implied a model primarily determined by LASSO regression. Those ΔFC-features were distributed into 134 different brain regions. We grouped and summed the absolute predictive weights of those features into their belonging functional brain networks based on Gordon et al. (2016) and further characterized the ΔFC-features contribution. As shown in Figure 4, ΔFCs from the sensorimotor network account for the largest proportion of predictive weights (14.64% of overall summed predictive weights) among the 13 networks, with the majority of ΔFC-links between the sensorimotor network and the cingulo-opercular network (5.14%), followed by the ΔFC-links between the sensorimotor network and the dorsal attention network (5.10%), and ΔFC-links within the sensorimotor network (3.25%). In addition, while ΔFCs from the visual and subcortical networks each account for 8.96% and 9.81% predictive weights, respectively, the ΔFC-links between the visual and subcortical networks alone account for a quite large proportion of 7.28% predictive weights of the overall 115 ΔFC-features suggesting a quite dominant contribution of this particular network connection.

3.3 | Predictive brain mapping

In order to characterize the contribution of individual sub-regions associated with each predictive functional network in greater detail, the absolute predictive weights were summed across all features from each single region and then mapped into each LOOCV fold. One sample t-tests revealed the brain regions[1] surviving FWE multiple comparison correction (Figure 5). The most strongly contributing brain region was located within the right fusiform gyrus, which belongs to the visual network. In addition, a region within the left middle occipital lobule bordering the inferior parietal lobule (belonging to the dorsal attention network) also demonstrated a significant role for predicting habit strength. Strikingly large *t*-values were also found for regions belonging to the cingulo–opercular network, such as the left precentral area and the anterior cingulate cortex (among other significant

[1]Individual brain regions were named based on the automatic anatomical labeling atlas (AAL).

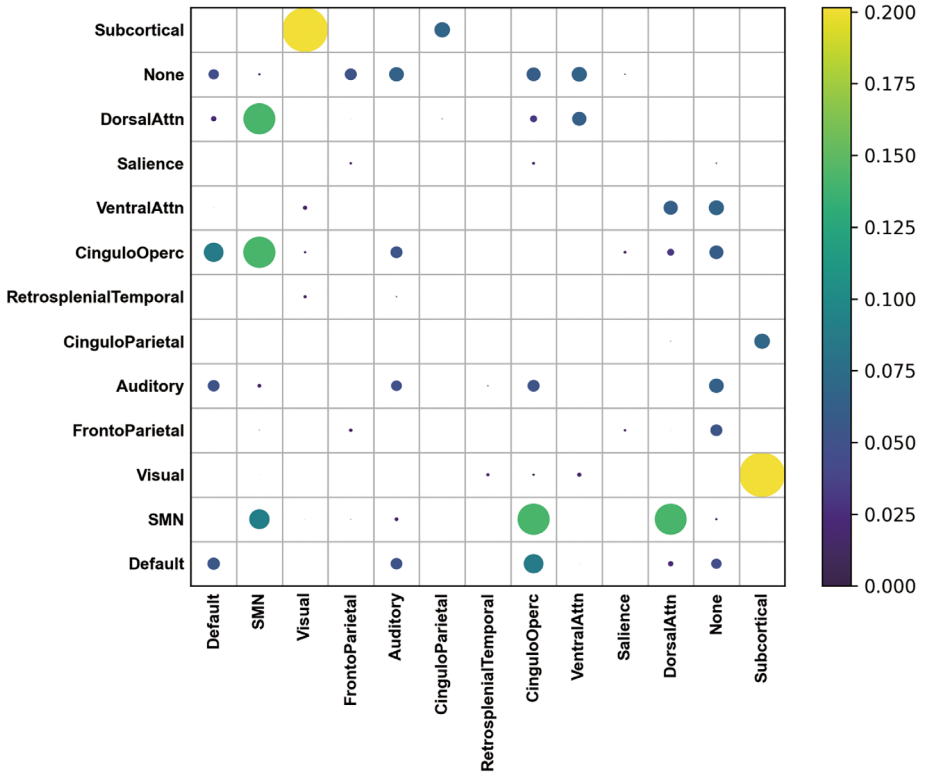

FIGURE 4 Predictive weights associated with 115 common ΔFC-features selected by Elastic Net were summed and grouped into 13 different networks. Color bar and circle size denoted the summed predictive weights of ΔFC-features either within or between different brain networks

regions within the opercular part of the cingulo–opercular network), as well as the right somatosensory area belonging to the sensorimotor network. The precuneus that is associated with the default mode network also played a significant predictive role. Finally, for the subcortical areas, only left caudate contributed significantly to habit strength prediction.

We finally performed an additional post hoc examination to identify specific ΔFC-links that contributed most strongly to the habit strength prediction. Here, the individual regions with a *t*-value higher than 20 (accounting for 42.22% predictive weights among all 115 features) were considered important and used for the post hoc examination. Note that excluding links with $t < 20$ was meant to highlight only the most reliable effects. As shown in Figure 6, the ΔFC between right fusiform and left caudate contributed most (mean = 0.184 ± 0.051), followed by the left middle occipital lobule and right postcentral area (mean = 0.138 ± 0.036). The ΔFC between left precentral area and right postcentral area (mean = 0.129 ± 0.036), and between left anterior cingulate cortex and right precuneus (mean = 0.0867 ± 0.025) also demonstrated an especially significant role in habit strength prediction.

4 | DISCUSSION

Here, we examined whether individual habit strength could be predicted by whole-brain FC changes during the transition from initially goal-directed to more habitual action. We first computed the pairwise FC changes (ΔFC) among 347 anatomically defined brain regions during learning. By entering selected ΔFC features into Elastic Net regression embedded within a LOOCV scheme, we could successfully predict the habit strength score of individual subjects. This basic result crucially demonstrates that global learning-related changes in large-scale functional connectivity are indeed associated with interindividual differences in acquired habit strength.

We further quantified the obtained predictive regression weights to identify the most relevant functional brain networks and their associated brain regions. Overall, we found widespread functional network changes contributing to the successful prediction of individual habit strength. Conceptually of outstanding importance, however, is the dominant role of regions from sensorimotor and cingulo–opercular networks (see especially Figure 4), as discussed next.

The significant predictive role of FC changes from the sensorimotor network (SMN) in habit strength prediction further highlights the often-assumed pivotal role of primary sensorimotor processes in habit formation as suggested by both human and nonhumans studies (Ashby et al., 2010; Bassett et al., 2015; Graybiel, 2008; Jahanshahi et al., 2015; Seger, 2018; Sidarta et al., 2016; Thorn et al., 2010; Vahdat et al., 2014; Wolfensteller & Ruge, 2012).

The predictive contribution of the SMN also comprised a strong SMN-CON component and there were additional widespread CON-related predictive connectivity changes. Together, this is consistent with previous reports of CON involvement in downstream (sensorimotor) output control (Dosenbach et al., 2006; Mohr et al., 2016; Newbold et al., 2021; Wallis et al., 2015). As can be seen in Figure 5, a number of opercular sub-regions of the CON were also involved but associated with overall smaller predictive weights and, to our

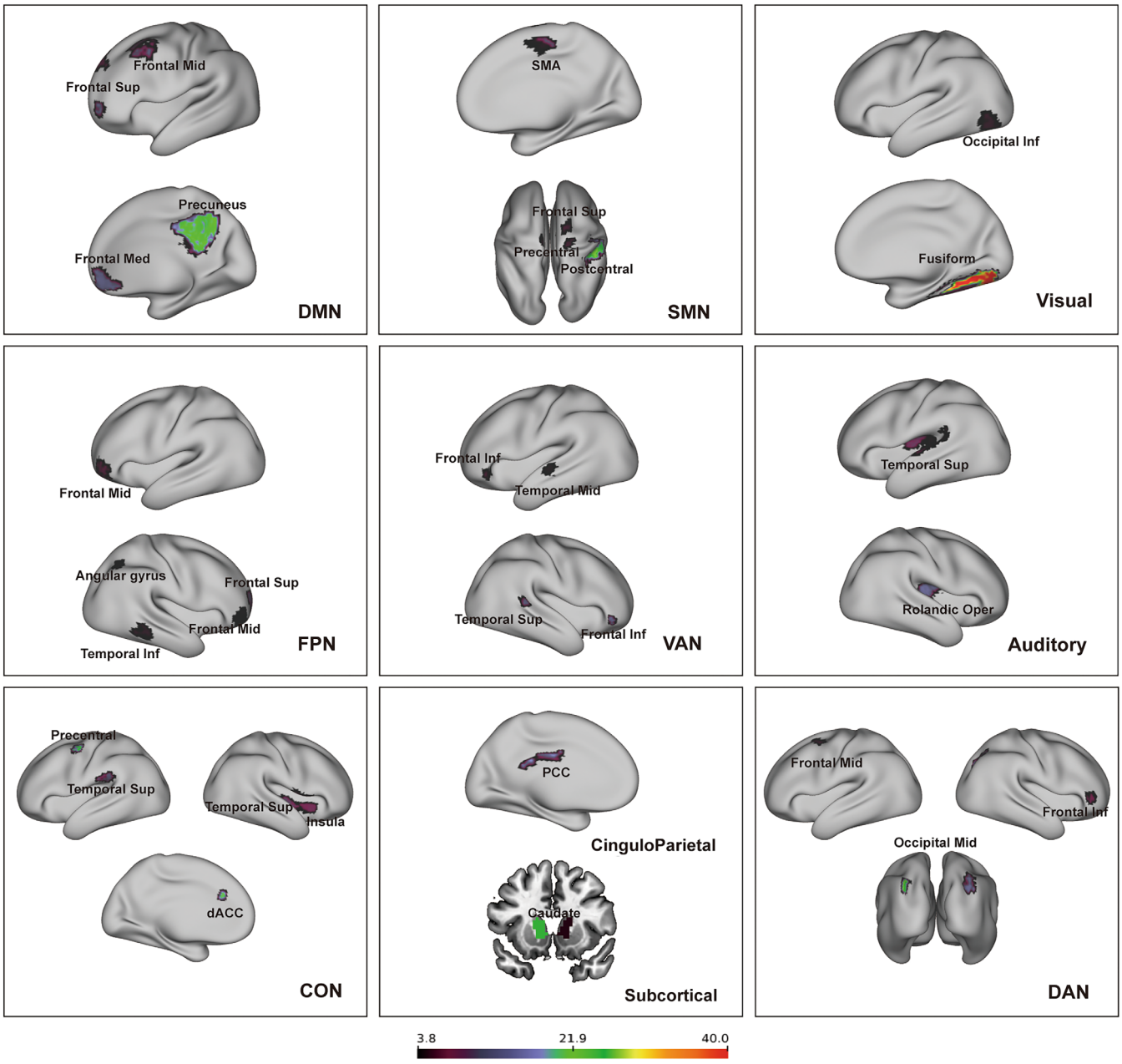

FIGURE 5 One sample *t*-test of the summed absolute predictive weights across LOOCV folds after FWE multiple comparison correction

knowledge, a less clear-cut functional role in sensorimotor control (but see, Ruge & Wolfensteller, 2010). Interestingly, the strongest CON-related contributions were from the precentral gyrus (more specifically from premotor cortex; PMC), and the dorsal anterior cingulate cortex (dACC). The functional role of precentral gyrus/PMC regarding habit-predictive connectivity changes is consistent with previously reported changes in local activation levels, both in the present paradigm (Zwosta et al., 2018) but also in other studies involving stimulus–response learning on various timescales (Ruge & Wolfensteller, 2010; Tschentscher et al., 2012; Wallis & Miller, 2003). Together, learning-related changes in local brain activation and functional connectivity, demonstrate the crucial role of the PMC for habit formation while transitioning from goal-directed to habitual action control. Also, the present involvement of the dACC as part of the cingulo–opercular network, is consistent with previously reported changes in local activation levels. Specifically, previous studies have shown that local dACC activation decreased significantly across S-R learning (Hampshire et al., 2016; Milham et al., 2003; Ruge & Wolfensteller, 2013; Sliwinska et al., 2017; Zwosta et al., 2018), which might suggest a particularly dominant role early in performance monitoring gradually fading out with progressing automatization (Noonan et al., 2011; Ruge & Wolfensteller, 2013; Walton et al., 2004).

Remarkably, connectivity changes involving the subcortical putamen were not relevant for habit strength prediction. This suggests that habit-strength, at least as assessed in the present study, is solely

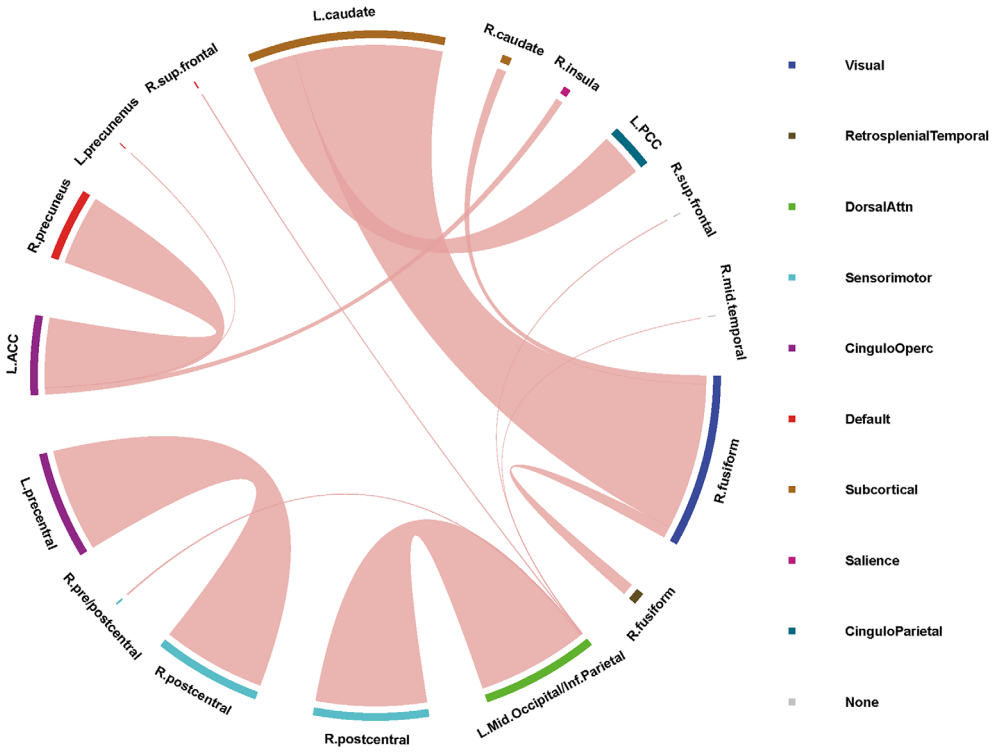

FIGURE 6 Visualization of specific ΔFCs that contributed most strongly to habit strength prediction (note that we excluded other significant, yet less strongly contributing ΔFCs in order to avoid an overly complex figure). Widths of ribbons are proportional to the predictive weights of the ΔFCs. Individual brain regions are named according to the AAL atlas

mediated by cortico-cortical plasticity already at an intermediate stage of automatization (compare Ashby et al., 2010).

As elaborated already in the Introduction, the previous analysis of training-related local activity changes strongly suggested a prominent role of the angular gyrus (AG) of the inferior parietal lobe for predicting habit strength (Zwosta et al., 2018). Given a putative role of AG in goal-directed action rather than habitual action, this previous finding highlighted the relevance of a decreasing engagement of goal-directed action control during learning for the subsequent expression of habitual tendencies after devaluation. However, in terms of connectivity changes, the present analysis highlights a prominent predictive role of the sensorimotor together with motor-related CON regions. Together, this suggests that for predicting later habit strength, it is relevant to consider training-related neural changes both in the goal-directed system (AG activity) and in the sensorimotor system together with the CON.

There are several potential limitations of the current study that should be noted. First, the current study lacks an additional external sample for establishing out-of-sample validity. Previous studies that predominantly focused on the unbiased prediction of behavior, additionally tested the performance of the internal predictive model using an independent external sample collected by different centers/scanning sites (Kwak et al., 2021; Rosenberg, Finn, et al., 2016; Spisak et al., 2020). However, compared with those studies, the current study emphasizes the predictive weights of connectivity changes from different functional brain networks and also the specific region associated with each network during the habit strength predictions. For this reason, we believe that this limitation might not necessarily affect our conclusions about which regions' connectivity changes contribute most during habit strength predictive analysis, though it will be desirable to obtain the additional external validation for the unbiased predictive performance in future research.

Second, while the current study focused on predicting individual differences in habit strength, it is mostly unknown whether this measure is trait-like—that is, stable and reliable over time within individuals. This would be important to establish in future studies as the theoretical maximum of predictive power is the reliability of the to-be-predicted behavioral measure.

5 | CONCLUSION AND OUTLOOK

The current study demonstrates the crucial role of functional connectivity changes associated with the sensorimotor network and the cingulo–opercular network in habit strength prediction. We therefore argue that these networks can be considered playing an important role during goal-habit transition. While we failed to demonstrate a predictive role of connectivity changes associated with the angular gyrus (as a putative part of the goal-directed network), previously reported study results showed that habit strength could be predicted based on learning-related local activity changes in the angular gyrus. Since some of the most relevant regions within the sensorimotor and the cingulo–opercular networks and also the angular gyrus are located superficially in the human brain, they could easily be reached by non-invasive brain stimulation techniques such as transcranial magnetic stimulation. We consider this a promising direction for future research to examine the potential causal role of those regions in human goal-habit transition.

ACKNOWLEDGMENTS

This work was supported by the German Research Foundation (DFG, SFB 940 project A2, Z2). Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

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

ORCID

Xiaoyu Wang https://orcid.org/0000-0002-5628-9531

REFERENCES

Ashby, F. G., Turner, B. O., & Horvitz, J. C. (2010). Cortical and basal ganglia contributions to habit learning and automaticity. *Trends in Cognitive Sciences*, *14*, 208–215.

Bassett, D. S., Wymbs, N. F., Porter, M. A., Mucha, P. J., Carlson, J. M., & Grafton, S. T. (2011). Dynamic reconfiguration of human brain networks during learning. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 7641–7646.

Bassett, D. S., Wymbs, N. F., Rombach, M. P., Porter, M. A., Mucha, P. J., & Grafton, S. T. (2013). Task-based core-periphery organization of human brain dynamics. *PLoS Computational Biology*, *9*, e1003171.

Bassett, D. S., Yang, M., Wymbs, N. F., & Grafton, S. T. (2015). Learning-induced autonomy of sensorimotor systems. *Nature Neuroscience*, *18*, 744–751.

Braun, U., Schafer, A., Walter, H., Erk, S., Romanczuk-Seiferth, N., Haddad, L., Schweiger, J. I., Grimm, O., Heinz, A., Tost, H., Meyer-Lindenberg, A., & Bassett, D. S. (2015). Dynamic reconfiguration of frontal brain networks during executive cognition in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 11678–11683.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, *1124*, 1–38.

Cao, H., Plichta, M. M., Schafer, A., Haddad, L., Grimm, O., Schneider, M., Esslinger, C., Kirsch, P., Meyer-Lindenberg, A., & Tost, H. (2014). Test-retest reliability of fMRI-based graph theoretical properties during working memory, emotion processing, and resting state. *NeuroImage*, *84*, 888–900.

Cole, M. W., Ito, T., Schultz, D., Mill, R., Chen, R., & Cocuzza, C. (2019). Task activations produce spurious but systematic inflation of task functional connectivity estimates. *NeuroImage*, *189*, 1–18.

Cole, M. W., Reynolds, J. R., Power, J. D., Repovs, G., Anticevic, A., & Braver, T. S. (2013). Multi-task connectivity reveals flexible hubs for adaptive task control. *Nature Neuroscience*, *16*, 1348–1355.

Cui, Z., & Gong, G. (2018). The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features. *NeuroImage*, *178*, 622–637.

Cui, Z., Li, H., Xia, C. H., Larsen, B., Adebimpe, A., Baum, G. L., Cieslak, M., Gur, R. E., Gur, R. C., Moore, T. M., Oathes, D. J., Alexander-Bloch, A. F., Raznahan, A., Roalf, D. R., Shinohara, R. T., Wolf, D. H., Davatzikos, C., Bassett, D. S., Fair, D. A., … Satterthwaite, T. D. (2020). Individual variation in functional topography of association networks in youth. *Neuron*, *106*, 340–353.e8.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*, 1704–1711.

de Wit, S., Kindt, M., Knot, S. L., Verhoeven, A. A. C., Robbins, T. W., Gasull-Camos, J., Evans, M., Mirza, H., & Gillan, C. M. (2018). Shifting the balance between goals and habits: Five failures in experimental habit induction. *Journal of Experimental Psychology. General*, *147*, 1043–1065.

de Wit, S., Watson, P., Harsay, H. A., Cohen, M. X., van de Vijver, I., & Ridderinkhof, K. R. (2012). Corticostriatal connectivity underlies individual differences in the balance between habitual and goal-directed action control. *The Journal of Neuroscience*, *32*, 12066–12075.

Delorme, C., Salvador, A., Valabregue, R., Roze, E., Palminteri, S., Vidailhet, M., de Wit, S., Robbins, T., Hartmann, A., & Worbe, Y. (2016). Enhanced habit formation in Gilles de la Tourette syndrome. *Brain*, *139*, 605–615.

Desmurget, M., Reilly, K. T., Richard, N., Szathmari, A., Mottolese, C., & Sirigu, A. (2009). Movement intention after parietal cortex stimulation in humans. *Science*, *324*, 811–813.

Dickinson, A. (1985). Actions and habits: The development of behavioural autonomy. *Philosophical transactions of the Royal Society of London B, Biological Sciences*, *308*, 67–78.

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron*, *80*, 312–325.

Dosenbach, N. U., Visscher, K. M., Palmer, E. D., Miezin, F. M., Wenger, K. K., Kang, H. C., Burgund, E. D., Grimes, A. L., Schlaggar, B. L., & Petersen, S. E. (2006). A core system for the implementation of task sets. *Neuron*, *50*, 799–812.

Dubois, J., Galdi, P., Han, Y., Paul, L. K., & Adolphs, R. (2018). Resting-state functional brain connectivity best predicts the personality dimension of openness to experience. *Personality Neuroscience*, *1*, e6.

Dubois, J., Galdi, P., Paul, L. K., & Adolphs, R. (2018). A distributed brain network predicts general intelligence from resting-state human neuroimaging data. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *373*, 1756.

Ersche, K. D., Gillan, C. M., Jones, P. S., Williams, G. B., Ward, L. H. E., Luijten, M., Wit, S. D., Sahakian, B. J., Bullmore, E. T., & Robbins, T. W. (2016). Carrots and sticks fail to change behavior in cocaine addiction. *Science*, *352*, 1468–1471.

Geerligs, L., Cam, C., & Henson, R. N. (2016). Functional connectivity and structural covariance between regions of interest can be measured more accurately using multivariate distance correlation. *NeuroImage*, *135*, 16–31.

Gordon, E. M., Laumann, T. O., Adeyemo, B., Huckins, J. F., Kelley, W. M., & Petersen, S. E. (2016). Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cerebral Cortex*, *26*, 288–303.

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience*, *31*, 359–387.

Greene, A. S., Gao, S., Scheinost, D., & Constable, R. T. (2018). Task-induced brain state manipulation improves prediction of individual traits. *Nature Communications*, *9*, 2807.

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., & Sporns, O. (2008). Mapping the structural core of human cerebral cortex. *PLoS Biology*, *6*, e159.

Hampshire, A., Hellyer, P. J., Parkin, B., Hiebert, N., MacDonald, P., Owen, A. M., Leech, R., & Rowe, J. (2016). Network mechanisms of intentional learning. *NeuroImage*, *127*, 123–134.

Igelstrom, K. M., & Graziano, M. S. A. (2017). The inferior parietal lobule and temporoparietal junction: A network perspective. *Neuropsychologia*, *105*, 70–83.

Jahanshahi, M., Obeso, I., Rothwell, J. C., & Obeso, J. A. (2015). A fronto-striato-subthalamic-pallidal network for goal-directed and habitual inhibition. *Nature Reviews. Neuroscience*, *16*, 719–732.

Jangraw, D. C., Gonzalez-Castillo, J., Handwerker, D. A., Ghane, M., Rosenberg, M. D., Panwar, P., & Bandettini, P. A. (2018). A functional connectivity-based neuromarker of sustained attention generalizes to predict recall in a reading task. *NeuroImage*, *166*, 99–109.

Kwak, S., Kim, H., Kim, H., Youm, Y., & Chey, J. (2021). Distributed functional connectivity predicts neuropsychological test performance among older adults. *Human Brain Mapping*, *42*, 3305–3325.

Liljeholm, M., Dunne, S., & O'Doherty, J. P. (2015). Differentiating neural systems mediating the acquisition vs. expression of goal-directed and habitual behavioral control. *The European Journal of Neuroscience*, *41*, 1358–1371.

Liljeholm, M., Tricomi, E., O'Doherty, J. P., & Balleine, B. W. (2011). Neural correlates of instrumental contingency learning: Differential effects of action-reward conjunction and disjunction. *The Journal of Neuroscience*, *31*, 2474–2480.

Lin, Q., Rosenberg, M. D., Yoo, K., Hsu, T. W., O'Connell, T. P., & Chun, M. M. (2018). Resting-state functional connectivity predicts cognitive impairment related to Alzheimer's disease. *Frontiers in Aging Neuroscience*, *10*, 94.

Melcher, T., Winter, D., Hommel, B., Pfister, R., Dechent, P., & Gruber, O. (2013). The neural substrate of the ideomotor principle revisited: Evidence for asymmetries in action-effect learning. *Neuroscience*, *231*, 13–27.

Milham, M. P., Banich, M. T., Claus, E. D., & Cohen, N. J. (2003). Practice-related effects demonstrate complementary roles of anterior cingulate and prefrontal cortices in attentional control. *NeuroImage*, *18*, 483–493.

Mohr, H., Wolfensteller, U., Betzel, R. F., Misic, B., Sporns, O., Richiardi, J., & Ruge, H. (2016). Integration and segregation of large-scale brain networks during short-term task automatization. *Nature Communications*, *7*, 13217.

Mumford, J. A., Turner, B. O., Ashby, F. G., & Poldrack, R. A. (2012). Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *Neuroimage*, *59*(3), 2636–2643.

Newbold, D. J., Gordon, E. M., Laumann, T. O., Seider, N. A., Montez, D. F., Gross, S. J., Zheng, A., Nielsen, A. N., Hoyt, C. R., Hampton, J. M., Ortega, M., Adeyemo, B., Miller, D. B., Van, A. N., Marek, S., Schlaggar, B. L., Carter, A. R., Kay, B. P., Greene, D. J., … Dosenbach, N. U. F. (2021). Cingulo-opercular control network and disused motor circuits joined in standby mode. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(13), e2019128118.

Niv, Y., Joel, D., & Dayan, P. (2006). A normative perspective on motivation. *Trends in Cognitive Sciences*, *10*, 375–381.

Noonan, M. P., Mars, R. B., & Rushworth, M. F. (2011). Distinct roles of three frontal cortical areas in reward-guided behavior. *The Journal of Neuroscience*, *31*, 14399–14412.

Plitt, M., Barnes, K. A., Wallace, G. L., Kenworthy, L., & Martin, A. (2015). Resting-state functional connectivity predicts longitudinal change in autistic traits and adaptive functioning in autism. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, E6699–E6706.

Pool, E. R., Gera, R., Fransen, A., Perez, O. D., Cremer, A., Aleksic, M., Tanwisuth, S., Quail, S., Ceceli, A. O., Manfredi, D. A., Nave, G., Tricomi, E., Balleine, B., Schonberg, T., Schwabe, L., & O'Doherty, J. P. (2022). Determining the effects of training duration on the behavioral expression of habitual control in humans: A multilaboratory investigation. *Learning & Memory*, *29*, 16–28.

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, *59*, 2142–2154.

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*, *84*, 320–341.

Rosenberg, M. D., Finn, E. S., Scheinost, D., Papademetris, X., Shen, X., Constable, R. T., & Chun, M. M. (2016). A neuromarker of sustained attention from whole-brain functional connectivity. *Nature Neuroscience*, *19*, 165–171.

Rosenberg, M. D., Hsu, W. T., Scheinost, D., Todd Constable, R., & Chun, M. M. (2018). Connectome-based models predict separable components of attention in novel individuals. *Journal of Cognitive Neuroscience*, *30*, 160–173.

Rosenberg, M. D., Zhang, S., Hsu, W. T., Scheinost, D., Finn, E. S., Shen, X., Constable, R. T., Li, C. S., & Chun, M. M. (2016). Methylphenidate modulates functional network connectivity to enhance attention. *The Journal of Neuroscience*, *36*, 9547–9557.

Ruge, H., & Wolfensteller, U. (2010). Rapid formation of pragmatic rule representations in the human brain during instruction-based learning. *Cerebral Cortex*, *20*, 1656–1667.

Ruge, H., & Wolfensteller, U. (2013). Functional integration processes underlying the instruction-based learning of novel goal-directed behaviors. *NeuroImage*, *68*, 162–172.

Seger, C. A. (2018). Corticostriatal foundations of habits. *Current Opinion in Behavioral Sciences*, *20*, 153–160.

Seger, C. A., & Spiering, B. J. (2011). A critical review of habit learning and the basal ganglia. *Frontiers in Systems Neuroscience*, *5*, 66.

Shine, J. M., Bissett, P. G., Bell, P. T., Koyejo, O., Balsters, J. H., Gorgolewski, K. J., Moodie, C. A., & Poldrack, R. A. (2016). The dynamics of functional brain networks: Integrated network states during cognitive task performance. *Neuron*, *92*, 544–554.

Sidarta, A., Vahdat, S., Bernardi, N. F., & Ostry, D. J. (2016). Somatic and reinforcement-based plasticity in the initial stages of human motor learning. *The Journal of Neuroscience*, *36*, 11682–11692.

Sliwinska, M. W., Violante, I. R., Wise, R. J. S., Leech, R., Devlin, J. T., Geranmayeh, F., & Hampshire, A. (2017). Stimulating multiple-demand cortex enhances vocabulary learning. *The Journal of Neuroscience*, *37*, 7606–7618.

Spisak, T., Kincses, B., Schlitt, F., Zunhammer, M., Schmidt-Wilcke, T., Kincses, Z. T., & Bingel, U. (2020). Pain-free resting-state functional brain connectivity predicts individual pain sensitivity. *Nature Communications*, *11*, 187.

Thorn, C. A., Atallah, H., Howe, M., & Graybiel, A. M. (2010). Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron*, *66*, 781–795.

Tricomi, E., Balleine, B. W., & O'Doherty, J. P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. *The European Journal of Neuroscience*, *29*, 2225–2232.

Tschentscher, N., Hauk, O., Fischer, M. H., & Pulvermuller, F. (2012). You can count on the motor cortex: Finger counting habits modulate motor cortex activation evoked by numbers. *NeuroImage*, *59*, 3139–3148.

Vahdat, S., Darainy, M., & Ostry, D. J. (2014). Structure of plasticity in human sensory and motor networks due to perceptual learning. *The Journal of Neuroscience*, *34*, 2451–2463.

Van Dijk, K. R., Sabuncu, M. R., & Buckner, R. L. (2012). The influence of head motion on intrinsic functional connectivity MRI. *NeuroImage*, *59*, 431–438.

Van Essen, D. C., Smith, J., Glasser, M. F., Elam, J., Donahue, C. J., Dierker, D. L., Reid, E. K., Coalson, T., & Harwell, J. (2017). The brain analysis library of spatial maps and atlases (BALSA) database. *NeuroImage*, *144*, 270–274.

Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E., & Buckner, R. L. (2008). Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *Journal of Neurophysiology*, *100*, 3328–3342.

Wallis, G., Stokes, M., Cousijn, H., Woolrich, M., & Nobre, A. C. (2015). Frontoparietal and cingulo-opercular networks play dissociable roles in control of working memory. *Journal of Cognitive Neuroscience*, *27*, 2019–2034.

Wallis, J., & Miller, E. (2003). From rule to response: Neuronal processes in the premotor and prefrontal cortex. *Journal of Neurophysiology*, *90*, 1790–1806.

Walton, M. E., Devlin, J. T., & Rushworth, M. F. (2004). Interactions between decision making and performance monitoring within prefrontal cortex. *Nature Neuroscience*, *7*, 1259–1265.

Wolfensteller, U., & Ruge, H. (2012). Frontostriatal mechanisms in instruction-based learning as a hallmark of flexible goal-directed behavior. *Frontiers in Psychology*, *3*, 192.

Yan, C. G., Cheung, B., Kelly, C., Colcombe, S., Craddock, R. C., Di Martino, A., Li, Q., Zuo, X. N., Castellanos, F. X., & Milham, M. P. (2013). A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics. *NeuroImage*, *76*, 183–201.

Yan, C. G., Wang, X. D., Zuo, X. N., & Zang, Y. F. (2016). DPABI: Data Processing & Analysis for (resting-state) brain imaging. *Neuroinformatics*, *14*, 339–351.

Yin, H., Knowlton, B., & Balleine, B. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *The European Journal of Neuroscience*, *19*, 181–189.

Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews. Neuroscience*, *7*, 464–476.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society*, *67*, 301–320.

Zwosta, K., Ruge, H., Goschke, T., & Wolfensteller, U. (2018). Habit strength is predicted by activity dynamics in goal-directed brain systems during training. *NeuroImage*, *165*, 125–137.

Zwosta, K., Ruge, H., & Wolfensteller, U. (2015). Neural mechanisms of goal-directed behavior: Outcome-based response selection is associated with increased functional coupling of the angular gyrus. *Frontiers in Human Neuroscience*, *9*, 180.

How to cite this article: Wang, X., Zwosta, K., Wolfensteller, U., & Ruge, H. (2023). Changes in global functional network properties predict individual differences in habit formation. *Human Brain Mapping*, *44*(4), 1565–1578. https://doi.org/10.1002/hbm.26158