Whole-brain dynamics of human sensorimotor adaptation

Dominic I. Standage*[1,3,4], Corson N. Areshenkoff[1], Daniel J. Gale[1], Joseph Y. Nashed[1,3], J. Randall Flanagan[1,2], Jason P. Gallivan[1,2,3]

[1]Centre for Neuroscience Studies, [2]Department of Psychology, and [3]Department of Biomedical and Molecular Sciences, Queen's University, Kingston, Ontario, Canada. [4]School of Psychology, University of Birmingham, UK

*Correspondence:

Dr. Dominic Standage
School of Psychology
University of Birmingham
d.standage@bham.ac.uk

Abstract

Individuals exhibit differences in learning ability, but the reasons for these differences are unclear. We used human functional magnetic resonance imaging (fMRI) to show that dynamic changes in whole-brain networks during the early stages of sensorimotor adaptation predict patterns of learning that emerge across two days of learning and relearning. A clustering of participant behavioural data revealed three distinct profiles of learners: individuals who learned quickly on both days, individuals who learned slowly on both days, and individuals who learned slowly on the first day, but quickly on the second day. These learning profiles were associated with the degree of modular reconfiguration during early learning on the first day, and with the selective recruitment of a cognitive network of regions, including areas in anterior temporal and prefrontal cortex. These findings demonstrate that across-day profiles of adaptation can be traced to whole-brain dynamics that manifest during initial learning.

Introduction

Successful behaviour relies on the brain's ability to flexibly adapt to environmental changes by learning new sensory-motor mappings. For example, a tennis player may favour a particular style of racket, but she will quickly adjust to a racket with a longer handle. For decades, it was widely believed that such sensorimotor adaptation was largely implicit or procedural in nature, driven by sensory prediction errors reflecting the difference between the expected and sensed consequences of movement (Wolpert 1995; 1998; 2000). Recent studies, however, have shown that explicit (declarative) cognitive strategies can also play an important role in guiding these adjustments, in parallel with implicit learning (see Krakauer & Mazzoni, 2011; McDougle et al., 2016; Wolpert et al., 2011). These two learning components have been well-characterized behaviourally and have been shown to operate on different timescales, with slow and fast adaptation linked to the implicit and explicit components respectively (McDougle et al., 2015; Smith et al., 2006; Taylor et al., 2014). There is emerging behavioural evidence that the implicit component is quite stable across individuals, with individual differences in the initial rate of learning being largely driven by the explicit component (de Brouwer et al., 2018; Fernandez-Ruiz et al., 2011). Moreover, it has been suggested that the explicit component is responsible for savings, or faster relearning upon reexposure to the task (Haith et al., 2015; Morehead et al., 2015). While a large body of work indicates that the cerebellum is essential for implicit learning (Izawa et al., 2012; Krakauer & Mazzoni, 2011; Martin et al., 1996; McDougle et al., 2016; Taylor et al., 2010; Taylor & Ivry, 2014; Tseng et al., 2007; Wolpert et al., 2011), the neural bases of explicit learning remains poorly understood.

Previous studies have used functional neuroimaging methods to document changes in the activity of individual brain regions during sensorimotor adaptation (Anguera et al., 2007; Dayan & Cohen, 2011; Doyon et al., 2009; S. Kim et al., 2015; Krakauer et al., 2004; Shadmehr, 1997). It is widely acknowledged, however, that this type of learning is supported by whole-brain distributed circuitry, spanning the cerebrum, cerebellum and striatum (Doyon et al., 2009). Understanding such distributed neural systems requires an investigation of the interactions between brain regions in large-scale networks. Thus, there is a critical need for studies to characterise functional networks supporting sensorimotor adaptation and to quantify changes in these networks during learning (Makino et al., 2016; Poldrack, 2006).

In recent years, the use of graph theory to construct and analyse formal networks derived from functional magnetic resonance imaging (fMRI) data has been a highly informative approach for investigating whole-brain networks (Danielle S. Bassett & Mattar, 2017; Bullmore & Sporns, 2009a; Hutchison et al., 2013; Sporns, 2014). Under this *network neuroscience* approach (Danielle S. Bassett & Sporns, 2017), coactive brain regions are hypothesised to form a functional network during a time window of interest. A compelling aspect of these networks is their modularity, or their decomposition (a.k.a. partitioning) into subnetworks characterised by dense connectivity between regions in the same module and sparse connectivity between regions in different modules. Modularity has long been recognised as a crucial property of complex biological systems (see D. S. Bassett et al., 2011) and is hypothesized to limit the degree to which new learning interferes with previous learning (Ellefsen et al., 2015). The structural and functional organization of the human brain is highly modular (Bullmore & Sporns, 2009b; Chen et al., 2008; Ferrarini et al., 2009) and recent fMRI studies have shown that the propensity of participants' brain regions to change modular affiliation (known as flexibility) correlates with cognitive task performance (Braun et

al., 2015) and motor skill acquisition (Danielle S. Bassett et al., 2015; D. S. Bassett et al., 2011).

Here, we used a similar network neuroscience approach to investigate the whole-brain bases of sensorimotor adaptation learning. We collected fMRI data while participants performed a classic sensorimotor adaptation task (visuomotor rotation learning) on two days, separated by twenty-four hours. We tested three principal hypotheses. Firstly, we hypothesized that individual variation in learning would cluster into distinct learning profiles across days. Secondly, we hypothesized that modular flexibility during early learning on the first day would be associated with a faster learning profile. We based this hypothesis on behavioural findings demonstrating links between participants' initial learning rates and the use of explicit, cognitive strategies (Miyamoto et al., 2020) and on evidence that cognitive processes are supported by modular flexibility (Braun et al., 2015). Thirdly, building on this second hypothesis, we predicted that a network of cognitive brain regions (e.g. areas in anterior temporal and prefrontal cortex) would be involved in learning, and that the recruitment of this cognitive network during early learning on the first day would be associated with a faster profile of learning.

Results

On two testing days, separated by 24 hours, we recorded whole-brain fMRI activity while participants (N=32) adapted, de-adapted, and then re-adapted (on the second day) their movements during a visuomotor rotation (VMR) task. This error-based learning task required participants to use their right (dominant) hand to perform center-out movements of a cursor to single targets, pseudo-randomly selected without replacement from 8 possible locations. To establish baseline performance, participants executed 120 standard, non-rotation trials (i.e., 15 bins of 8 trials). Following this baseline condition (henceforth Baseline), a 45° clockwise rotation was applied for the next 320 trials (Learning), where the viewed cursor was rotated about the hand's start location, requiring that participants learn to adapt their movements to the perturbation. Finally, to allow for an investigation of relearning the following day, we de-adapted participants' learning with 120 more baseline trials (Washout). One day later, participants returned to undergo the same testing regimen, thereby allowing us to measure savings [i.e. faster (re)learning on Day 2 than Day 1] upon re-exposure to the VMR.

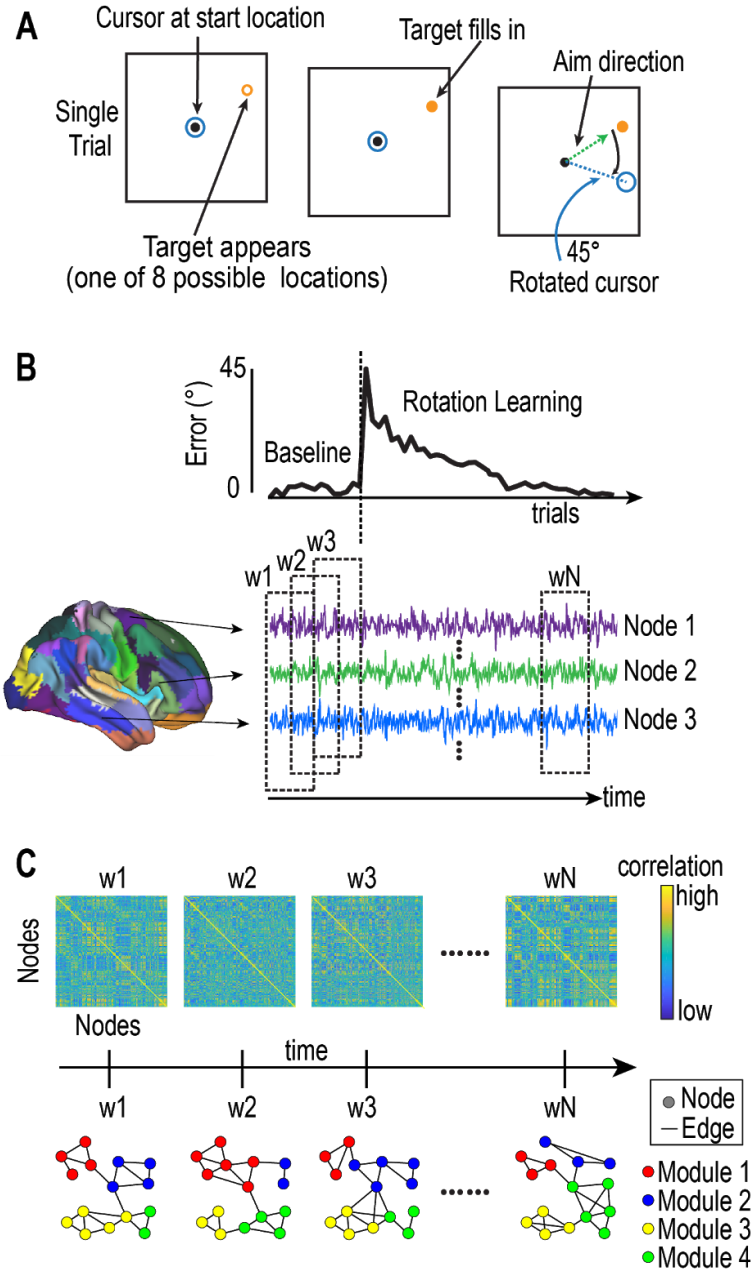

Figure 1: Overview of task and neural analyses. (A) Visuomotor rotation (VMR) task, where the viewed cursor, controlled by the hand, is rotated about the movement start location. (B) Each participant's cerebrum, striatum and cerebellum were parcellated into discrete brain regions and the average % Blood Oxygenation Level-Dependent (BOLD) time series was extracted from each region for the entire task on Day 1 (three example cerebral regions shown). The Pearson correlation coefficient (corrected for false discovery rate, see text) was calculated for each pair of regions in sliding, half-overlapping windows, constructing whole-brain functional connectivity matrices for each window (w1-wN; shown in 'C'). (C) Time-resolved clustering methods were applied to the networks constructed in 'B', linking module labels (for each node) across time slices. This procedure identified network modules that evolved during learning (4 modules shown in the schematic).

Behavioural analyses reveal three profiles of learning across days

The left panel in Figure 2A shows mean learning curves over all participants. Consistent with prior work (Krakauer, 2009; Morehead et al., 2015), these curves show that, on average, participants learned to reduce their errors during task performance and exhibited savings. However, these group-level findings obscure significant inter-participant variability in the patterns of learning across days. For example, participant P1 (Figure 2B) showed fast learning on both days (fast-fast, FF), participant P2 showed slow learning on both days (slow-slow, SS) and participant P3 showed a pattern of slow learning on Day 1 (magenta trace) but fast learning on Day 2 (cyan trace, slow-fast, SF). To test our first hypothesis that inter-participant variability in learning could be accounted for by a limited number of distinct learning profiles, we performed *k*-means clustering across all participants' early and late

binned errors on each day, along with their measure of savings (i.e. five behavioural measures per participant). These measures captured each participant's rate of learning (early binned error on each day), completeness of learning (late binned error on each day) and rate of relearning (savings) respectively. The clustering algorithm separated participants into three distinct subgroups or profiles of learners (Figure 2C) that showed very similar patterns of learning (Figure 2D) to the FF, SS and SF participants shown in Figure 2B (15 FF, 10 SS and 7 SF participants).

These subgroups revealed two important features of across-day learning (Figure 2D-E) that were obscured at the group level. Firstly, learning on Day 1 did not always predict learning on Day 2, given that a sizeable proportion of 'slow' learners on Day 1 (7 of 17 participants) became 'fast' learners on Day 2 (i.e. the SF subgroup). Secondly, savings was not ubiquitous across participants, as suggested by the group-level data in Figure 2A. Rather, savings differed significantly between the subgroups (Figure 2D inset), with the largest magnitude of savings (~28°) expressed by SF participants, followed by FF participants (~10°). In contrast, the SS subgroup did not show significant savings. Together, these results indicate that a large portion of the savings observed at the group level was driven by the SF subgroup, since participants in the FF subgroup were already fast learners on Day 1.

Having defined the three behavioural subgroups (FF, SS and SF), we next sought to identify a single scalar measure that reliably captures subgroup membership, in order to test our hypotheses on the relationships between learning profiles and whole-brain dynamics with standard correlation and regression techniques. To this end, we performed principal component analysis (PCA) on the same five behavioural measures used to derive the subgroups, finding that the first principal component (PC1) not only accounted for 61% of the variance in the behavioural data, but also correctly classified 31/32 participants' subgroup membership (Figure 2E). Guided by the PCA factor loadings (Supplementary Figure S1), we found that PC1 primarily captured participants' mean binned early error across days ($r=-0.98$, see Figure 2E inset), providing an intuitive interpretation of its meaning in relation to learning performance and subgroup membership. Note that we reversed the sign of PC1, so that higher magnitude corresponds to faster learning and smaller errors.

With a view to our analysis of participants' neural data, the above behavioural findings generate two competing predictions. If participants' Day 1 neural data tracks their Day 1 learning performance, then whole-brain dynamics of SS and SF participants will be similar, differing from FF participants. In contrast, if participants' Day 1 neural data anticipates their Day 2 learning performance, then whole-brain dynamics of FF and SF participants will be similar, differing from SS participants, e.g. if SF participants were trying (unsuccessfully) to implement a strategy on Day 1 (Miyamoto et al., 2020).

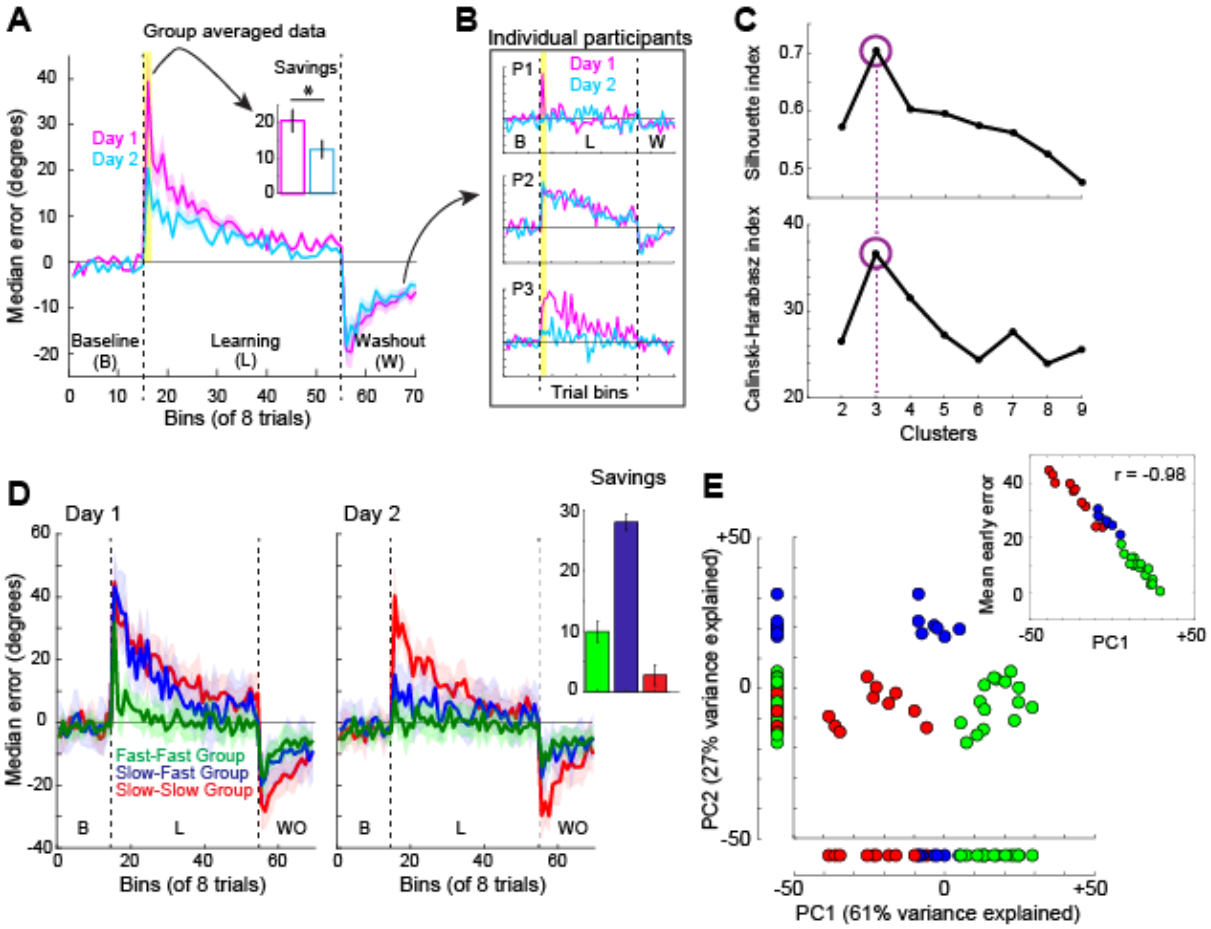

Figure 2. Clustering of participants' learning data revealed 3 behavioural subgroups. (A) Mean (across participants) of the binned median error (in degrees), during Baseline (non-rotation), Learning (45 degree rotation of the cursor) and Washout (non-rotation) on Day 1 (magenta) and Day 2 (blue). Shading shows standard error (±1 SE) and the dashed vertical lines demark the three task components. Note that savings (inset) is significant at the group level [paired t-test; t(31)=6.122, p=8.666e-7]. **(B)** The group-averaged approach in panel A obscures differences in learning across individuals (3 example participants, P1-P3) **(C)** *k*-means clustering of participants' early and late errors and savings across Days 1 and 2 (5 variables in total) identified three subgroups of participants. Plot shows the silhouette and Calinski-Harabasz (C-H) indices (see Methods), measuring the goodness of clustering. Both measures are maximized by three clusters (purple circles and dashed line indicate the best clustering solution). **(D)** The mean (across participants) binned median error for the three behavioural subgroups (identified in B,C) during Baseline, Learning and Washout on Day 1 (left) and Day 2 (right). The subgroups are readily distinguished by their early error on each day and thus we refer to their patterns of learning as fast-fast (FF, green, N=15), slow-slow (SS, red, N=10) and slow-fast (SF, blue, N=7). On Day 1, mean early error of FF was significantly lower than that of SS [two-sample t-test, t(23)=-9.963, p=8.203e-10] and SF [t(20)=-12.74, p=4.691e-11], but there was no statistical difference between SF and SS [t(15)=1.0931, p=0.292]. On Day 2, mean early error of FF was significantly lower than that of SS [t(23)=-7.413, p=8.476e-10] and SF [t(20)=-3.106, p=0.006], and mean early error of SF was significantly lower than that of SS [t(15)=-6.67, p=7.495e-6]. Inset shows savings by behavioural group. The FF and SF subgroups showed significant savings [t-test versus zero; FF: t(14)=5.817, p=4.473e-5; SF: t(6)=20.703, p=8.265e-7], but SS did not [t-test versus zero: t(9)=1.688, p=0.126]. **(E)** Top two components of the PCA for the 5 learning measures across participants. Single data points represent participants, with each participant color-coded by cluster-assigned subgroup. The x-axis of the plot shows that PC1 correctly classifies 31 of 32 participants (97% accuracy). Inset scatterplot shows that PC1 has a strong linear relationship with mean early error across days.

Modular flexibility is associated with learning profile

To test our second hypothesis that modular flexibility during early learning on Day 1 would be associated with a faster learning profile across days, we constructed multislice networks from each participant's fMRI data during Day 1 learning (see Methods) and partitioned these networks into spatio-temporal modules that maximized a quality function (Mucha et al., 2010). We then verified that participants' whole-brain networks showed significant modularity (Supplementary Figure S2) and quantified the flexibility of each brain region during each of early, mid and late Learning, defined as the number of times the region changed modules, relative to the number of possible changes (D. S. Bassett et al., 2011). We then computed the mean flexibility across all regions and modular partitions for each participant (see Methods) and calculated its relationship with PC1, our proxy for subgroup membership. Consistent with our hypothesis, we found a significant positive correlation between mean flexibility (henceforth flexibility) during early Learning and PC1 that was either weak or absent during mid or late Learning (Figure 3). To test whether the relationship between early flexibility and PC1 was learning-specific (i.e. only seen during Learning), we also computed the correlation between flexibility and PC1 during Baseline (i.e. during non-rotation trials prior to Learning). Notably, we observed a significant correlation between these measures (Figure 3 left), indicating that the positive relationship between modular flexibility and PC1 was not specific to learning. An analysis of visuomotor errors during Baseline revealed no differences between the behavioural subgroups (Supplementary Figure S3), indicating that flexibility did not support participants' current task performance. Rather, flexibility may reflect participants' state of attention (Shine et al., 2016) and/or mood (Betzel et al., 2017), which in turn relate to learning (Yerkes & Dodson, 1908). Thus, at first glance, the positive correlation between modular flexibility and PC1 supports our hypothesis that flexibility is linked to a faster learning profile, but this support does not imply a direct mechanistic relationship between the two measures.

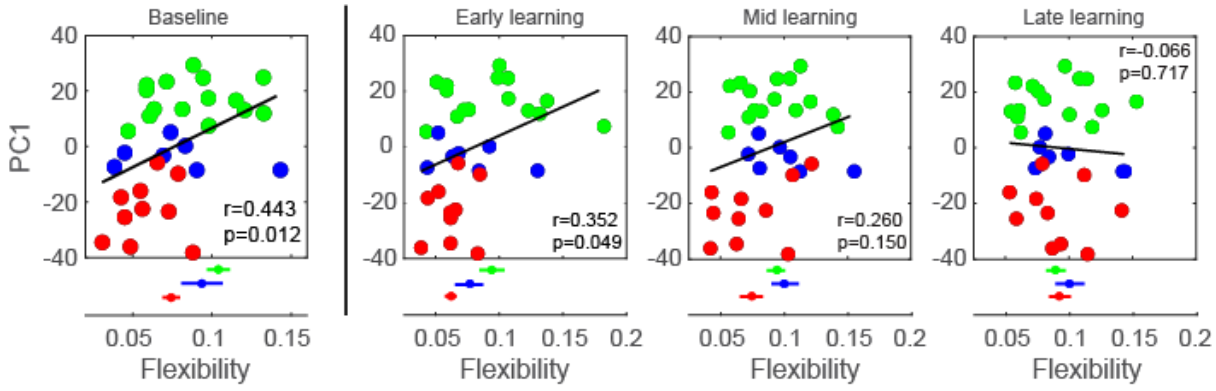

Figure 3. Modular flexibility during Baseline and early Day 1 learning is associated with participants' learning profiles across days. Scatter plots show PC1 as a function of modular flexibility during Baseline (left of vertical line) and during early, mid and late Learning on Day 1 (early: windows 1 - 13; mid: windows 14 - 26; late: windows 27 - 39). Green, red and blue circles correspond to the FF, SS and SF subgroups respectively. In each plot, fitted line shows the best linear fit (least squares method). R and p-values correspond to the coefficient and statistical significance of a Spearman correlation test. Subgroup means ±1 SE are shown below the scatter plots. Two-sample t-tests showed a significant difference between FF and SS during Baseline [$t(23)=2.966$, $p=0.007$] and early Learning [$t(23)=2.5$, $p=0.02$]. No other subgroup differences were significant during Baseline or any Learning epoch ($0.08 < p < 0.98$).

Dynamic network architecture is associated with learning profile

Having determined that participants' whole-brain functional networks showed significant modularity and that modular flexibility before and during learning was associated with participants' learning profiles, we sought to characterise the modules (brain subnetworks) undergoing this reconfiguration and to determine whether properties of their time-evolution predicted learning. In each Learning epoch (early, mid and late) we calculated the proportion of all modular partitions (over all participants, scans and time windows) in which each pair of regions was placed in the same module, referred to as a module allegiance matrix (Danielle S. Bassett et al., 2015). We then computed the average module allegiance matrix across epochs. To identify brain networks present across time and across participants, we clustered this matrix using symmetric nonnegative matrix factorization (Supplementary Figure S4). This clustering effectively identifies a consensus partition, or set of networks that summarises the structure of the temporally dynamic modules (Danielle S. Bassett et al., 2013). The summary network architecture and the module allegiance matrix from which it was derived are shown in Figure 4A-B. Consistent with our third hypothesis, we found that one of the derived networks (Network 1) was composed of primarily cognitive brain regions in anterior temporal and prefrontal cortex.

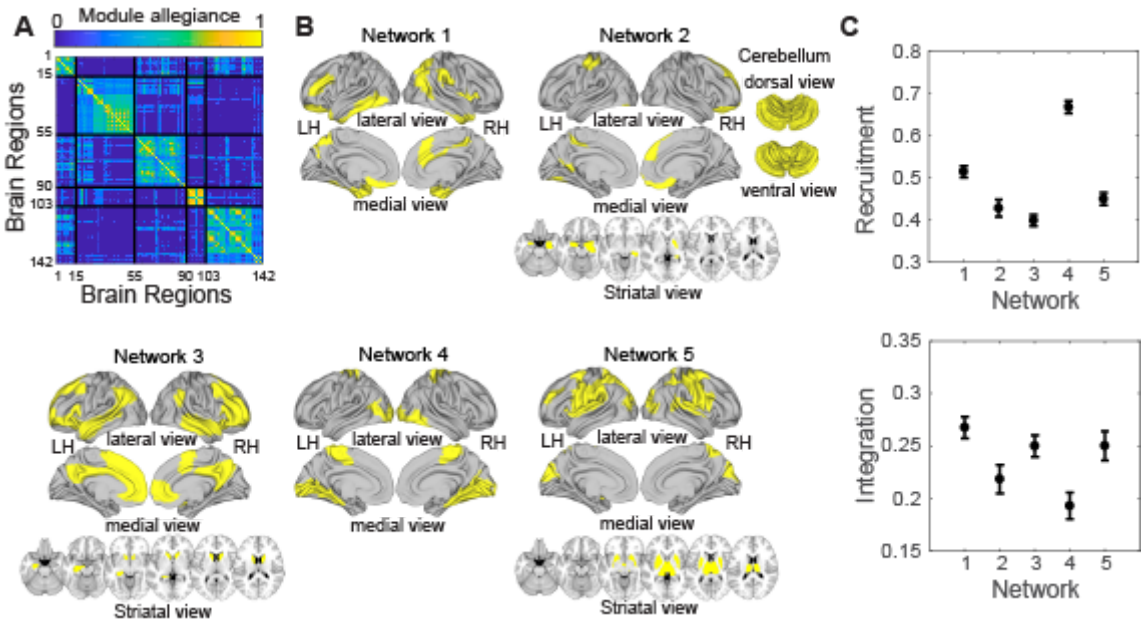

Figure 4. Summary of dynamic network architecture during adaptation learning. (A) Module-allegiance matrix showing the probability that each pair of brain regions on Day 1 was in the same temporal module, calculated over all participants, modular partitions and time windows of each Learning epoch (early, mid and late), before taking the mean over the epochs. The matrix is organized according to the consensus clustering solution (see text). **(B)** Network clusters from 'A', rendered onto the cortical and subcortical surfaces. Network 1 consisted of regions spanning ventrotemporal, anterior medial temporal, inferior parietal and dorsolateral prefrontal cortex. Network 2 consisted of regions spanning motor cortex, cerebellum, medial prefrontal cortex and several subcortical structures. Network 3 consisted of regions spanning angular gyrus, temporal gyrus, dorsal and medial prefrontal cortex, cingulate cortex, and several subcortical areas. Network 4 consisted of regions spanning visual and fusiform cortex, and medial somatomotor cortex. Lastly, Network 5 consisted of regions in visual cortex, medial and lateral parietal cortex, lateral somatomotor cortex and premotor cortex, along with bilateral thalamus and putamen. Note that we refrain from giving names to these networks, beyond their generic Network 1 - 5 descriptions. **(C)** Network recruitment (upper) and integration (lower) for each network (see text). The latter shows the mean integration of each network with the other four networks. Error bars show ±1 SE of the across-participant means.

To test our hypothesis that the recruitment of this cognitive brain network during early Learning on the first day would be correlated with a faster learning profile, we sought to determine the degree to which each of the identified summary networks interacted with one another over time. With each brain region assigned to one of the five networks, the interaction between any two networks can be measured by $I_{k1,k2} = (\sum i \in C_{k1}, j \in C_{k2} P_{i,j})/(|C_{k1}||C_{k2}|)$ (Danielle S. Bassett et al., 2015), where $C_{k \in 1,2}$ are modules, $|C_k|$ is the number of regions they contain, and $P_{i,j}$ is the proportion of the time regions i and j are in the same module (the module allegiance matrix, Figure 4A). We refer to the interaction of a network with itself as recruitment (Danielle S. Bassett et al., 2015), calculated by allowing $k1 = k2$. The integration between two networks $k1 \neq k2$ is the normalised interaction between them $I'_{k1,k2} = I_{k1,k2} / \sqrt{(I_{k1,k1} I_{k2,k2})}$, referred to as network integration (Danielle S. Bassett et al., 2015). In effect, recruitment captures the stability of a network's membership over time, whereas integration captures the interactions between separate networks over time. Figure 4C shows the recruitment and integration scores for each network. An analysis of variance showed a strong effect of network on each measure (recruitment: F(4)=35.01, p=8.134e-21; integration: F(4)=3.96, p=0.004).

To estimate the contribution of each network to participants' learning profiles, we fit separate multiple linear regression models $(PC1 \approx 1 + 1 \cdot x1 + 2 \cdot x2 + 3 \cdot x3 + 4 \cdot x4 + 5 \cdot x5)$ to the recruitment and integration scores for each network during Baseline and each Learning epoch (early, mid and late). Under this approach, the fitted -value for each network provides an estimate of the contribution of the network's recruitment/integration to PC1, the t-statistic corresponding to each estimate quantifies the statistical difference of from zero, and the F-statistic quantifies the overall goodness of fit of the model. Given the large number of comparisons required to assess these predictor and outcome relationships, we limited our focus to cases in which the p-values corresponding to both the F-statistic (for the fit) and the t-statistic (for a given -value) were statistically significant (p<0.05, highlighted by stars in Figure 5). During early Learning on Day 1, the recruitment of Networks 1 and 5 satisfied these criteria, but neither recruitment nor integration satisfied these criteria during Baseline, nor during mid or late Learning.

Recruitment of each of Networks 1 and 5 was associated with a positive contribution to PC1 during early Learning, that is, greater stability in the composition of these networks was associated with a faster learning profile (Figure 5A). Further inspection revealed a strong correlation between the recruitment of Network 1 and PC1 during early Learning, where recruitment was higher in the FF subgroup than the SS subgroup. There was also a significant correlation between the recruitment of Network 5 and PC1, where recruitment was higher in the FF subgroup than the SS and SF subgroups (Figure 5B, right lower left panel). Importantly, the correlation between the recruitment of Network 1 and PC1 retained its significance when we corrected for the same correlation during Baseline (r=0.369, p=0.041), whereas the significance of the correlation between Network 5 recruitment and PC1 did not (r=0.175, p=0.346). Thus, only the relationship between Network 1 recruitment and PC1 was learning-specific. This result is consistent with our hypothesis that the recruitment of a cognitive network during early learning on the first day would be associated with a faster learning profile. Of note, the differences in recruitment between behavioural subgroups (Figure 5B) suggest that the recruitment of Network 1 anticipates Day 2 learning, given the similarity between the FF and SF subgroups and their difference from SS. In

contrast, the recruitment of Network 5 tracks Day 1 learning, since the SS and SF subgroups differ from FF. We return to this interesting finding in the Discussion.

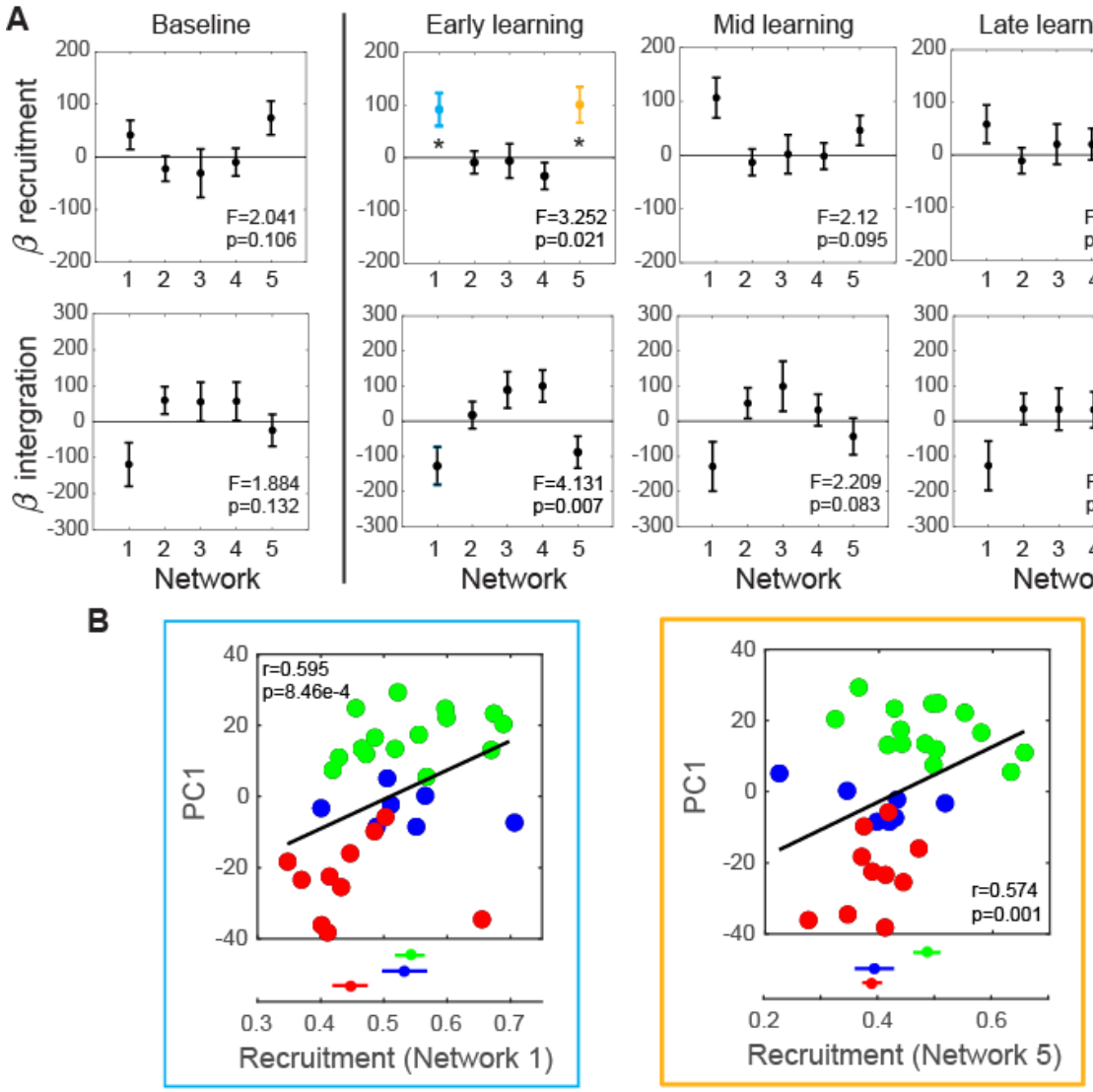

Figure 5. Contribution of network recruitment and integration to adaptation learning. (A) For each network, β-values from a linear regression model show the weighting of network recruitment (upper) and integration (lower) in predicting PC1 during Baseline (left of vertical line) and during early, mid and late Learning on Day 1 (right of vertical line). Error bars show ±1 SE of across-participant means. For significant fits (Bonferroni-corrected across regressions for recruitment and integration, *i.e.* $p < 0.05/2$), stars between β and zero denote statistical significance of the t-statistic for the estimate of each β-value (Bonferroni-corrected across networks, *i.e.* $p < 0.05/5$). Recruitment results are shown in blue and orange for Networks 1 and 5 respectively, where further analysis is provided in panel B. **(B)** Scatter plots show PC1 as a function of recruitment for significant results in 'A' (*i.e.* significant fits and β-values statistically different from zero). In each plot, r- and p-values correspond to Spearman correlation. Green, red and blue circles correspond to FF, SS and SF subgroups respectively. Line shows best linear fit (least squares method). Subgroup means ±1 SE are shown below the scatter plots. Blue and orange boxes correspond to the colour coded results in 'A'. During early Learning in Network 1 (blue), recruitment was significantly higher in FF than SS [two-sample t-test: $t(23)=2.612$, $p=0.016$] but did not differ significantly between FF and SF [$t(23)=0.204$, $p=0.841$] or between SS and SF [$t(15)=1.95$, $p=0.07$]. During early learning in Network 5 (orange), recruitment of Network 5 was significantly higher in FF than SS [$t(23)=2.977$, $p=0.007$] and SF [$t(20)=2.208$, $p=0.039$], but SS and SF did not differ significantly [$t(15)=0.111$, $p=0.913$]. For each of Networks 1 (blue) and 5 (orange), inset r- and p-values show results of a partial Spearman correlation between recruitment and PC1, controlling for recruitment of the other four networks.

Discussion

Humans show marked variability in their capacity for sensorimotor adaptation, the neural bases of which are unclear. We sought to characterise this variation and investigate its neural correlates by having participants perform a sensorimotor adaptation task on two testing days, separated by 24 hrs. A clustering analysis of participants' learning data across days revealed three distinct profiles of learners: participants who exhibited fast learning on both days (the FF subgroup), suggesting the use of explicit, cognitive strategies during learning (de Brouwer et al., 2018); participants who exhibited slow learning on both days (the SS subgroup), suggesting the use of primarily implicit-based learning (de Brouwer et al., 2018); and participants who exhibited slow learning on Day 1, but fast learning on Day 2 (the SF subgroup). Consistent with our hypothesis that modular flexibility would be associated with a faster learning profile, we found that the FF subgroup exhibited greater flexibility than the SS subgroup. However, this effect was not unique to early learning, as it was also observed during baseline trials, when performance across the subgroups was indistinguishable. This finding suggests that modular flexibility reflects a task-invariant network property that is associated with successful learning, but is not specific to learning. Finally, consistent with our hypothesis that a cognitive network would be involved in learning, we found that the recruitment of a network of anterior temporal and prefrontal areas (Network 1) was associated with a faster learning profile. Taken together, these findings provide a novel characterisation of whole-brain dynamics during human sensorimotor adaptation learning and suggest that the selective recruitment of cognitive systems during early learning underpins distinct learning profiles.

Prior neuroimaging studies of adaptation learning in humans have documented changes in activation during learning in a broad range of brain regions, including prefrontal, cingulate, premotor, motor and parietal cortices, as well as the striatum and cerebellum (Anguera et al., 2007; Dayan & Cohen, 2011; Doyon et al., 2009; Hardwick et al., 2013; S. Kim et al., 2015; Krakauer et al., 2004; Shadmehr, 1997). Identifying the specific roles of these regions in learning has proved challenging (Dayan & Cohen, 2011; Hardwick et al., 2013), but a key theme to emerge from this work is that the early and late phases of adaptation learning appear to engage different regions. For example, studies have shown increases in the activation of several higher-order brain regions during early learning, such as dorsolateral prefrontal cortex (DLPFC) and anterior cingulate cortex (Anguera et al., 2010, 2011), that appear to shift towards other regions during late learning, such as cerebellum (Flament et al., 1996; S. Kim et al., 2015; Seidler & Noll, 2008). Additionally, the storage limitations of visuospatial working memory (Christou et al., 2016) and the activation of DLPFC (Anguera et al., 2010) have been shown to correlate with faster adaptation learning across participants. These findings suggest an important role for cognitive brain systems during the early stages of learning in particular, and may explain why cognitive deficits associated with aging and explicit memory have been linked to slower adaptation learning (Anguera et al., 2011; Trewartha et al., 2014). Our results are consistent with this trend. By capitalising on recent methods from network neuroscience (Danielle S. Bassett & Sporns, 2017) and by quantifying learning-related changes in network structure over the course of Day 1 learning, we found that during early learning, the selective recruitment of a cognitive brain network (Network 1) involving regions in anterior temporal and prefrontal cortex was associated with faster learning across days. One possible explanation of this result is that this cognitive network underlies the explicit, cognitive component of adaptation learning (Taylor et al., 2014), the magnitude of which has been shown to positively correlate with faster rates of adaptation (de Brouwer et al., 2018).

Consistent with this explanation, we found that the recruitment of Network 2, composed of regions spanning the primary motor cortex and cerebellum, was *not* associated with learning performance. This null result is consistent with earlier results indicating that these structures are primarily involved in the implicit component of sensorimotor adaptation (Izawa et al., 2012; Norris et al., 2011; Schlerf et al., 2012; Taylor et al., 2010; Tseng et al., 2007), which is always present during learning and has been shown to be similar across participants (de Brouwer et al., 2018; Haith et al., 2015; Huberdeau et al., 2019). Thus, unlike the cognitive circuitry discussed above, such motor-related brain regions would not be expected to contribute to individual differences in learning performance.

Our behavioural findings also warrant further discussion. It is interesting that learning on Day 1 did not always predict learning on Day 2. This result was asymmetric, since fast learners on Day 1 were always fast learners on Day 2 (FF), whereas slow learners on Day 1 were either slow (SS) or fast (SF) learners on Day 2. One of the striking features of our behavioural clusters is the extent to which savings differed across subgroups. Savings in visuomotor rotation tasks are widely believed to be driven by the recall of a successful re-aiming strategy, which has been directly linked to the explicit component (Haith et al., 2015; Morehead et al., 2015). If so, faster learners should express greater savings (de Brouwer et al., 2018). Here, we found that fast learners on Day 1 did indeed show significant savings (the FF subgroup), but the largest savings were expressed by the SF subgroup, who were slow learners on Day 1 (~3 times larger savings than FF). This result conflicts with conventional interpretations of how savings relates to participants' learning rates, but is nevertheless intuitive in terms of the calculation of savings, i.e. by subtracting early error on Day 2 from early error on Day 1. In this regard, fast learners on Day 1 have limited scope for savings, since they are already fast learners. In contrast, slow learners on Day 1 have ample scope for improvement when they reencounter the rotation. If so, what explains the very different trajectories of Day 2 learning between the SF and SS subgroups in our study?

One possibility is that SF participants experienced a 'eureka' moment during the initial stages of relearning on Day 2, that is, the SF subgroup may have suddenly identified a successful re-aiming strategy. Another possibility is that some kind of sensorimotor consolidation process occurred during sleep between Day 1 and Day 2 (Albouy et al., 2013; Korman et al., 2007) allowing the SF subgroup to relearn the rotation more quickly on Day 2. While these possibilities cannot be ruled out, we believe it is more likely that SF participants identified the effective re-aiming strategy at some point on Day 1, which they reimplemented on Day 2 (see, for e.g., Sailer et al., 2005). There are two lines of evidence to support this conjecture, based on our behavioural and neural data. Firstly, the FF and SF subgroups exhibited similar de-adaptation during Washout (see Fig. 2D). This observation suggests that by the end of Day 1 learning, SF participants had identified the re-aiming strategy and could readily discard it when the rotation was removed (Benson et al., 2011; de Brouwer et al., 2018). Secondly, the FF and SF subgroups were similar in their recruitment of Network 1 during early learning (Fig. 5B). Thus, the recruitment of Network 1 by SF participants may reflect their exploration of explicit strategies, eventually allowing them to identify the correct re-aiming strategy. If so, what explains the similar learning trajectories of SS and SF participants on Day 1?

One possibility is that the explicit strategies supported by Network 1 were erratic among SF participants during early learning. Recent behavioural evidence suggests that erratic explicit strategies during adaptation learning are compensated by the implicit component (Miyamoto et al., 2020), potentially explaining why SF participants learned at approximately the same

rate as SS participants, i.e. this compensation may have mitigated against their erratic re-aiming. Another possibility is that SF and SS participants relied on the recruitment of Network 5 during early learning, since recruitment of this network was significantly lower in SS and SF participants than FF (Figure 5B). Network 5 was mainly composed of parietal and premotor areas that are often implicated in visuomotor control (Fig. 4), so perhaps an inability to fully engage this higher-order, motor-related network during early learning impacted SF participants' rate of learning. Future studies should investigate such possibilities.

Another important finding from our study is that modular flexibility during early learning was associated with better overall learning performance across days. On the one hand, this finding is consistent with earlier studies linking flexibility to motor sequence learning (D. S. Bassett et al., 2011) and reinforcement learning (Gerraty et al., 2018). On the other hand, the unexpected correlation between flexibility during Baseline and future learning casts its mechanistic role in an interesting light. During baseline trials on Day 1, participants had not encountered the visuomotor rotation and would not have had knowledge about strategies for enhancing performance on the task. Furthermore, behavioural performance during Baseline did not differ significantly between the FF, SF and SS subgroups. Thus, the positive correlation between flexibility during early learning on Day 1 and overall learning performance across days cannot be explained by participant variation in the processing of error-related visual feedback, nor by participants' implementation of a cognitive strategy. Rather, this finding suggests that flexibility reflects a general property of whole-brain dynamics that is conducive to cognitive performance (Braun et al., 2015), but may not directly support any specific cognitive process. In this way, modular flexibility may be likened to the psychological construct of cognitive flexibility (Telesford et al., 2017), explaining why it has been associated with performance across several cognitive domains, from learning (D. S. Bassett et al., 2011) and working memory (Braun et al., 2015), to language-related processing (Chai et al., 2016) and participants' level of arousal and fatigue (Betzel et al., 2017).

Methodological considerations

Our findings should be interpreted in light of several methodological considerations. Firstly, we derived our five summary networks from the propensity of brain regions to cluster together over time, rather than utilising network architectures revealed through resting-state functional connectivity. It is increasingly accepted that the structure of task-based functional connectivity can differ markedly from resting-state activity (Danielle S. Bassett et al., 2015) and our preference was to let the data speak for themselves in this regard. Secondly, our measurement of brain dynamics provides very different characterisations of neural processes than traditional studies of task-related activity. For example, our task did not use a block design or event-related structure, so the covariation of any two brain regions' activation was not a marker of their being activated or deactivated by task-related events, but rather, it revealed how their activation was functionally synchronized over time. We focused on the modularity of these functional networks, a well-characterised property of neural, social and technological systems (Danielle S. Bassett et al., 2010, 2011; Fortunato, 2010) that has been informative about the whole-brain bases of related tasks in recent years (D. S. Bassett et al., 2011; Braun et al., 2015). As with other analyses of functional connectivity (Fox & Raichle, 2007) our findings do not speak to causal or directional interactions between regions, which require methods for examining effective connectivity, such as dynamic causal modelling or structural equation modelling.

Thirdly, we leveraged the natural variation in participants' learning rates to infer the relative contribution of the explicit component of adaptation learning, rather than directly estimating its contribution, such as by having participants report their aiming directions (Taylor et al., 2014). We acknowledge that a more direct approach may have allowed us to make stronger claims about the links between cognitive strategies and brain networks, but our approach was based on several factors. Firstly, earlier behavioural studies by our group have established a tight correspondence between faster rates of adaptation learning and the use of explicit cognitive strategies (de Brouwer et al., 2018). Secondly, this earlier work established that 'reporting trials' magnify (and thereby distort) the use of such strategies. Given our present focus on individual differences, this issue was of particular concern, since reporting trials would likely have biased our sample in favour of faster learners (de Brouwer et al., 2018). Finally, and perhaps most importantly, we were concerned that the insertion of intermittent reporting trials would alter the nature of the continuous BOLD data collected, thereby impacting the construction and interpretation of temporal networks.

Conclusion

We have characterized different profiles of sensorimotor adaptation learning that emerge across days and have shown that these profiles are linked to the flexible reconfiguration of whole-brain networks early on Day 1 (i.e. during Baseline and early Learning). We have further identified a cognitive brain network involving regions in anterior temporal and prefrontal cortices, the recruitment of which was associated with future learning performance. These findings provide novel evidence for the roles of modular reconfiguration and higher-order cognitive systems in driving individual differences in learning.

Methods

Participants

40 right-handed participants between the ages of 18 and 35 (M = 22.5, SD = 4.51; 13 males) received financial compensation for their time. Data from 8 participants were excluded, due to either head motion in the MRI scanner (N=4; motion greater than 2 mm or 2° rotation within a single scan) or their inability to learn the rotation task (N=4), leaving 32 participants in the final analysis. Right-handedness was assessed with the Edinburgh handedness questionnaire (Oldfield, 1971). Participants' written, informed consent was obtained before commencement of the experimental protocol. The Queen's University Research Ethics Board approved the study and it was conducted in accordance with the principles outlined in the Canadian Tri-Council Policy Statement on Ethical Conduct for Research Involving Humans and the principles of the Declaration of Helsinki (1964).

Apparatus

In the scanner, participants performed hand movements that were directed towards a target by applying a directional force onto an MRI-compatible force sensor (ATI technologies) using their right index finger and thumb. The target and cursor stimuli were rear-projected with an LCD projector (NEC LT265 DLP projector, 1024 x 768 resolution, 60 Hz refresh rate) onto a screen mounted behind the participant. The stimuli on the screen were viewed through a mirror fixed to the MRI coil directly above participants' eyes, thus preventing participants from being able to see the hand. The force sensor and associated cursor movement were sampled at 500 Hz.

Procedure

This experiment used a well-established visuomotor rotation task (Krakauer, 2009) to probe sensorimotor learning (Figure 1*A*). To start each trial, the cursor (20-pixel radius) appeared in a central start position (25-pixel radius). A white target circle (30-pixel radius) appeared at one of eight locations (0, 45, 90, 135, 180, 225, 270, 315º) on an invisible ring around the central position (300-pixel radius) and filled in (white) following a 200 ms delay. Once filled in, participants applied a brief directional force impulse on the force sensor (threshold of 1.5 N), which launched the cursor towards the target. Targets could randomly appear in each location, without replacement, during sets of 8 trials. Once launched, the cursor would travel the 300-pixel distance to the ring over a 400 ms period (with a bell-shaped velocity profile) before becoming stationary at the ring to provide participants with end-point error feedback, for up to 1 s. If the cursor overlapped with the target to any extent, the target would become green, signifying a target 'hit'. Each trial was separated by 4 s and within this period, participants had 3.0 s from target presentation to complete the trial (including the 200 ms target delay, participants' reaction times, the 400 ms cursor movement and the 1000 ms visual end-point feedback). At 3.0 s the trial was ended, the data saved, and participants would briefly wait for the next trial to begin. Reaction times were not stressed in this experimental procedure, as participants had up to 1400 ms on each trial to apply the target-directed force.

During each testing session, 120 baseline trials (15 sets of 8 trials) were completed without a rotation of the cursor. Following these baseline trials, 320 learning trials (40 sets of 8 trials) were completed, wherein a 45º clockwise rotation of the cursor was applied. The baseline and learning trials were completed during one continuous fMRI scan. Following this scan, conditions were restored to baseline (i.e. no rotation of cursor) in a separate scan and participants performed 120 washout trials. These washout trials allowed us to probe the rate of task relearning 24 hours later. In addition to these VMR-related task components on each testing day, we also interspersed three 6-minute resting state fMRI scans prior to, between, and following VMR Learning and Washout [note that spontaneous brain activity during these Rest scans is not considered in this paper]. During these resting-state scans, participants were instructed to rest with their eyes open, while fixating a central cross location presented on the screen. The total testing time was 75 minutes on each testing day.

MRI Acquisition

Participants were scanned using a 3-Tesla Siemens TIM MAGNETOM Trio MRI scanner located at the Centre for Neuroscience Studies, Queen's University (Kingston, Ontario, Canada). Functional MRI volumes were acquired using a 32-channel head coil and a T2*-weighted single-shot gradient-echo echo-planar imaging (EPI) acquisition sequence (time to repetition (TR) = 2000 ms, slice thickness = 4 mm, in-plane resolution = 3 mm x 3 mm, time to echo (TE) = 30 ms, field of view = 240 mm x 240 mm, matrix size = 80 x 80, flip angle = 90°, and acceleration factor (integrated parallel acquisition technologies, iPAT) = 2 with generalized auto-calibrating partially parallel acquisitions (GRAPPA) reconstruction. Each volume comprised 34 contiguous (no gap) oblique slices acquired at a ~30° caudal tilt with respect to the plane of the anterior and posterior commissure (AC-PC), providing whole-brain coverage of the cerebrum and cerebellum. A T1-weighted ADNI MPRAGE anatomical was also collected (TR = 1760 ms, TE = 2.98 ms, field of view = 192 mm x 240 mm x 256 mm, matrix size = 192 x 240 x 256, flip angle = 9°, 1 mm isotropic voxels). For each resting-state scan, 180 imaging volumes were collected. For the baseline and rotation scan, a single, continuous scan was collected of 896 imaging volumes duration. For the

washout scan, one scan of 256 imaging volumes was collected. Each scan included an additional 8 imaging volumes at both the beginning and end of the scan.

On Day 1, a separate practice session was carried out before the actual fMRI experiment to familiarize participants with the task. This session involved performing 60 practice baseline (non-rotation) trials. The fMRI testing session for each participant lasted 2 hours and included set-up time (~20 min.), practice (~10 min.), one high-resolution anatomical scan (~8 min duration), two DTI scans (one in the AP direction and the other in the PA direction; ~10 min duration), a resting-state scan (6 min duration), a baseline and rotation scan (~30 min duration), a resting-state scan (6 min. duration), a washout scan (~9 min duration), and a final resting-state scan (6 min. duration). For the purposes of the present paper, we only consider results associated with the baseline-rotation scan collected on Day 1.

MRI Preprocessing

Preprocessing was performed using fMRIPrep 1.4.0 (Esteban, Ciric, et al., 2019; Esteban et al., 2018) RRID:SCR_016216), which is based on Nipype 1.2.0 (Esteban, Markiewicz, et al., 2019); RRID:SCR_002502).

Anatomical data preprocessing. T1w scans of each session were corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection (Tustison et al., 2010), distributed with ANTs 2.2.0 (Avants et al., 2008) RRID:SCR_004757). The T1w-reference was then skull-stripped with a Nipype implementation of the antsBrainExtraction.sh workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast (FSL 5.0.9, RRID:SCR_002823, (Zhang et al., 2001) A T1w-reference map was computed after registration of the individual T1w scans (after INU-correction) using mri_robust_template (FreeSurfer 6.0.1, (Reuter et al., 2010). Brain surfaces were reconstructed using recon-all (FreeSurfer 6.0.1, RRID:SCR_001847, (Dale et al., 1999), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle (RRID:SCR_002438, (Klein et al., 2017). Volume-based spatial normalization to FSL's MNI ICBM 152 non-linear 6th Generation Asymmetric Average Brain Stereotaxic Registration Model ((Evans et al., 2012), RRID:SCR_002823; TemplateFlow ID: MNI152NLin6Asym) was performed through nonlinear registration with antsRegistration (ANTs 2.2.0), using brain-extracted versions of both T1w reference and the T1w template.

Functional data preprocessing. For each of the 10 BOLD runs per participant (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. The BOLD reference was then co-registered to the T1w reference using bbregister (FreeSurfer) which implements boundary-based registration (Greve & Fischl, 2009). Co-registration was configured with nine degrees of freedom to account for distortions remaining in the BOLD reference. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using mcflirt (FSL 5.0.9, (Jenkinson et al., 2002). BOLD runs were slice-time corrected using 3dTshift from AFNI 20160207 ((Cox & Hyde, 1997), RRID:SCR_005927). The BOLD time-series were resampled into standard space (MNI152NLin6Asym), correspondingly generating spatially-normalized, preprocessed BOLD runs. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. Several confounding time-series were calculated based on the

preprocessed BOLD. The six head-motion estimates calculated in the correction step were included as confounding time series, along with the temporal derivatives and quadratic terms for each. A set of physiological regressors were extracted to allow for component-based noise correction (Behzadi et al., 2007). Here, principal components are estimated after high-pass filtering the preprocessed BOLD time-series (using a discrete cosine filter with 128s cut-off) based on anatomical masks (aCompCor). Components are calculated within the intersection of the aforementioned mask and the union of CSF and WM masks calculated in T1w space, after their projection to the native space of each functional run (using the inverse BOLD-to-T1w transformation). Components are also calculated separately within the WM and CSF masks. *k* components with the largest singular values are retained, and the remaining components are dropped from consideration. All resamplings were performed with a single interpolation step by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using antsApplyTransforms (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos, 1964). Non-gridded (surface) resamplings were performed using mri_vol2surf (FreeSurfer).

Many internal operations of fMRIPrep use Nilearn 0.5.2 (Abraham et al., 2014), RRID:SCR_001362), mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in fMRIPrep's documentation.

Region-of-Interest (ROI) signal extraction. ROI extraction and nuisance regression was performed using Nilearn 0.5.2 (Abraham et al., 2014). For each participant, the raw BOLD time series for each voxel inside a given ROI was isolated using the Schaeffer 100 atlas (Schaefer et al., 2018). The mean time course of each ROI was highpass-filtered (cutoff = .01Hz), temporally detrended, and z-scored across each run in order to ensure that all voxels had the same scale of activation after preprocessing. In order to account for motion confounds and physiological noise, the above-described nuisance regressors were included. That is, the participant's six motion parameters and their derivatives, squares, and the squares of the derivatives (24 total), and the top six aCompCor regressors. Note that nuisance regression was performed orthogonally to temporal filters so as to avoid reintroducing high-frequency noise from the regressors (see (Lindquist et al., 2019).

Data Analysis

Behavioural Analyses. To assess performance on the task, the angle of the cursor relative to the target at the moment the cursor reached the target distance (or nearest data point if this value could not be used) was calculated. This was done by subtracting the target angle from an angle derived from the horizontal and vertical position at a 300-pixel distance from the start point. The median endpoint cursor error for the set of 8 trials (1 trial of each target) was extracted and the mean end-point error for early trials (bins 1-3) and late trials (bins 38-40) of each day were computed to derive learning measures for the participants. In addition, savings for each participant was assessed by subtracting mean binned early error on Day 2 from the mean binned error on Day 1 (Morehead et al., 2015).

To identify distinct subgroups of learners in our sample, k-means clustering (Lloyd, 1982) was performed across participants using these five learning parameters (early and late error on both days, and savings). We used the L1 distance function (Manhatten distance) and applied the clustering algorithm 10,000 times, with a random initialization of centroid positions, to increase chances of escaping local minima. The optimal cluster number was

determined as a convergence of solutions through separately applying the Calinski-Harabasz (Calinski & Harabasz, 1974) and Silhouette index (Rousseeuw, 1987) cluster evaluation indices. We obtained identical results using the L2 (Euclidean) distance function.

Temporal modular structure in brain activity. The time series for each region was divided into windows of 64 seconds (32 imaging volumes; corresponding to two bins of 8 trials; see *Procedure* above) where contiguous windows overlapped by one bin, i.e. by 50% (Chai et al., 2016). We constructed functional networks in each time window by calculating the Pearson correlation coefficient between each pair of brain regions. To mitigate against the possibility of spurious correlations (at $p<0.05$), all correlations that did not pass false discovery rate correction (Benjamini & Yekutieli, 2001) were set to zero ($q=0.05$), as were all negative correlations, a requirement of the module detection approach used. We determined the modular structure of the resulting multislice networks with a generalized Louvain method for time-resolved clustering (Jeub et al, http://netwiki.amath.unc.edu/GenLouvain, 2011-2017). This algorithm was repeated 100 times, resulting in 100 clustering solutions (a.k.a. partitions). On each repetition, we initialized the algorithm at random (uniform distribution) and iterated until a stable partition was found, *i.e.* the output on iteration *n* served as the input on iteration *n+1* until the output matched the input (D. S. Bassett et al., 2011; Mucha et al., 2010)). We used the standard spatial and temporal resolution parameters $\gamma = 1$ and $\omega = 1$ respectively (see Supplementary Figure S5 for demonstrations of robustness).

The flexibility of each region refers to the number of times it changes module allegiance, relative to the opportunity to do so, i.e. flexibility = n_f / (T - 1), where n_f is the number of changes and T is the number of slices (time windows) in the multislice network (D. S. Bassett et al., 2011). We defined the flexibility for each partition as the mean flexibility of the regions, before calculating the mean network flexibility across all partitions.

Module allegiance matrix. We constructed a matrix *T*, where the elements $T_{i,j}$ refer to the number of times regions *i* and *j* were assigned to the same module over all time slices and participants of the epoch. We then constructed module allegiance matrix $P=(1/C)T$, where *C* is the total number of time slices in all of these partitions. The final module allegiance matrix was the mean of *P* over the three Learning epochs (early, mid and late), where values that did not exceed those of a null model were set to 0 (see Methods).

Clustering of the module allegiance matrix with non-negative matrix factorisation. We clustered the module allegiance matrix using symmetric nonnegative matrix factorization [SymNMF; (Kuang et al., 2012; Lee & Seung, 1999)], which decomposes a symmetric matrix Y as $Y = AA^T$, where A is a rectangular matrix containing a set of nonnegative factors which minimize a sum-of-squares error criterion. SymNMF has previously shown good performance for factoring similarity matrices (Kuang et al., 2012, 2015), and has the benefit of producing a natural clustering solution by assigning each observation to the factor on which it loads most strongly.

To choose the number of factors, we fit 250 random initializations of each factorization rank in the range of 2 to 15. For each rank, we computed the average root mean square reconstruction error, dispersion coefficient (H. Kim & Park, 2007), and cophenetic correlation (Brunet et al., 2004). The latter two measures quantify the consistency of the clustering solutions returned by different initializations.

We judged that a rank 5 solution was acceptable, as this was sufficient to explain a large majority of the variance in the observed data (over 80%), as this number was acceptable by both the dispersion and cophenetic correlation, and additional clusters provided little improvement in explained variance (Figure S3). We then selected the rank 5 solution with the lowest reconstruction error, and generated a clustering solution by assigning each brain region to the factor with the highest loading.

Statistical analysis

Experimenters were not blind to testing. Statistical significance was defined by alpha pre-set to 0.05. Error bars indicate +/- 1 standard errors of the mean (SEM). All the statistical tests are described in the figure legends and each test was selected based on data distributions using histograms. Two-tailed tests were used unless noted otherwise.

References

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., & Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, *8*, 14.

Albouy, G., King, B. R., Maquet, P., & Doyon, J. (2013). Hippocampus and striatum: dynamics and interaction during acquisition and sleep-related motor sequence memory consolidation. *Hippocampus*, *23*(11), 985–1004.

Anguera, J. A., Reuter-Lorenz, P. A., Willingham, D. T., & Seidler, R. D. (2010). Contributions of Spatial Working Memory to Visuomotor Learning. In *Journal of Cognitive Neuroscience* (Vol. 22, Issue 9, pp. 1917–1930). https://doi.org/10.1162/jocn.2009.21351

Anguera, J. A., Reuter-Lorenz, P. A., Willingham, D. T., & Seidler, R. D. (2011). Failure to engage spatial working memory contributes to age-related declines in visuomotor learning. *Journal of Cognitive Neuroscience*, *23*(1), 11–25.

Anguera, J. A., Russell, C. A., Noll, D. C., & Seidler, R. D. (2007). Neural correlates associated with intermanual transfer of sensorimotor adaptation. *Brain Research*, *1185*, 136–151.

Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, *12*(1), 26–41.

Bassett, D. S., Brown, J. A., Deshpande, V., Carlson, J. M., & Grafton, S. T. (2011). Conserved and variable architecture of human white matter connectivity. In *NeuroImage* (Vol. 54, Issue 2, pp. 1262–1279). https://doi.org/10.1016/j.neuroimage.2010.09.006

Bassett, D. S., Greenfield, D. L., Meyer-Lindenberg, A., Weinberger, D. R., Moore, S. W., & Bullmore, E. T. (2010). Efficient physical embedding of topologically complex information processing networks in brains and computer circuits. *PLoS Computational Biology*, *6*(4), e1000748.

Bassett, D. S., & Mattar, M. G. (2017). A Network Neuroscience of Human Learning: Potential to Inform Quantitative Theories of Brain and Behavior. *Trends in Cognitive Sciences*, *21*(4), 250–264.

Bassett, D. S., Porter, M. A., Wymbs, N. F., Grafton, S. T., Carlson, J. M., & Mucha, P. J. (2013). Robust detection of dynamic community structure in networks. *Chaos* , *23*(1), 013142.

Bassett, D. S., & Sporns, O. (2017). Network neuroscience. *Nature Neuroscience*, *20*(3), 353–364.

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*, *108*(18), 7641–7646.

Bassett, D. S., Yang, M., Wymbs, N. F., & Grafton, S. T. (2015). Learning-induced autonomy of sensorimotor systems. *Nature Neuroscience*, *18*(5), 744–751.

Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, *37*(1), 90–101.

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. In *The Annals of Statistics* (Vol. 29, Issue 4, pp. 1165–1188). https://doi.org/10.1214/aos/1013699998

Benson, B. L., Anguera, J. A., & Seidler, R. D. (2011). A spatial explicit strategy reduces error but interferes with sensorimotor adaptation. *Journal of Neurophysiology*, *105*(6), 2843–2851.

Betzel, R. F., Satterthwaite, T. D., Gold, J. I., & Bassett, D. S. (2017). Positive affect, surprise, and fatigue are correlates of network flexibility. *Scientific Reports*, *7*(1), 520.

Braun, U., Schäfer, 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. In *Proceedings of the National Academy of Sciences* (Vol. 112, Issue 37, pp. 11678–11683). https://doi.org/10.1073/pnas.1422487112

Brunet, J.-P., Tamayo, P., Golub, T. R., & Mesirov, J. P. (2004). Metagenes and molecular pattern discovery using matrix factorization. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(12), 4164–4169.

Bullmore, E., & Sporns, O. (2009a). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews. Neuroscience*, *10*(3), 186–198.

Bullmore, E., & Sporns, O. (2009b). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews. Neuroscience*, *10*(3), 186–198.

Calinski, T., & Harabasz, J. (1974). A Dendrite Method for Cluster Analysis. In *Communications in Statistics - Simulation and Computation* (Vol. 3, Issue 1, pp. 1–27). https://doi.org/10.1080/03610917408548446

Chai, L. R., Mattar, M. G., Blank, I. A., Fedorenko, E., & Bassett, D. S. (2016). Functional Network Dynamics of the Language System. *Cerebral Cortex* , *26*(11), 4148–4159.

Chen, Z. J., He, Y., Rosa-Neto, P., Germann, J., & Evans, A. C. (2008). Revealing modular architecture of human brain structural networks by using cortical thickness from MRI. *Cerebral Cortex* , *18*(10), 2374–2381.

Christou, A. I., Miall, R. C., McNab, F., & Galea, J. M. (2016). Individual differences in explicit and implicit visuomotor learning and working memory capacity. *Scientific Reports*, *6*, 36633.

Cox, R. W., & Hyde, J. S. (1997). Software tools for analysis and visualization of fMRI data. *NMR in Biomedicine*, *10*(4-5), 171–178.

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage*, *9*(2), 179–194.

Dayan, E., & Cohen, L. G. (2011). Neuroplasticity subserving motor skill learning. *Neuron*, *72*(3), 443–454.

de Brouwer, A. J., Albaghdadi, M., Flanagan, J. R., & Gallivan, J. P. (2018). Using gaze behavior to parcellate the explicit and implicit contributions to visuomotor learning. *Journal of Neurophysiology*, *120*(4), 1602–1615.

Doyon, J., Bellec, P., Amsel, R., Penhune, V., Monchi, O., Carrier, J., Lehéricy, S., & Benali, H. (2009). Contributions of the basal ganglia and functionally related brain structures to motor learning. *Behavioural Brain Research*, *199*(1), 61–75.

Ellefsen, K. O., Mouret, J.-B., & Clune, J. (2015). Neural modularity helps organisms evolve to learn new skills without forgetting old skills. *PLoS Computational Biology*, *11*(4), e1004128.

Esteban, O., Ciric, R., Finc, K., Blair, R., Markiewicz, C. J., Moodie, C. A., Kent, J. D., Goncalves, M., DuPre, E., Gomez, D. E. P., Ye, Z., Salo, T., Valabregue, R., Amlien, I. K., Liem, F., Jacoby, N., Stojić, H., Cieslak, M., Urchs, S., … Gorgolewski, K. J. (2019). Analysis of task-based functional MRI data preprocessed with fMRIPrep. In *bioRxiv* (p. 694364). https://doi.org/10.1101/694364

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Ilkay Isik, A., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2018). fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods*, 1.

Esteban, O., Markiewicz, C. J., Burns, C., Johnson, H., Ziegler, E., Manhães-Savio, A., Jarecka, D., Ellis, D. G., Yvernault, B., Hamalainen, C., Notter, M. P., Salo, T., Waskom, M., Goncalves, M., Jordan, K., Wong, J., Dewey, B. E., Madison, C., Clark, D., … Ghosh, S. (2019). *nipy/nipype: 1.4.0*. https://doi.org/10.5281/zenodo.3588470

Evans, A. C., Janke, A. L., Collins, D. L., & Baillet, S. (2012). Brain templates and atlases. *NeuroImage*, *62*(2), 911–922.

Fernandez-Ruiz, J., Wong, W., Armstrong, I. T., & Flanagan, J. R. (2011). Relation between reaction time and reach errors during visuomotor adaptation. *Behavioural Brain Research*, *219*(1), 8–14.

Ferrarini, L., Veer, I. M., Baerends, E., van Tol, M.-J., Renken, R. J., van der Wee, N. J. A., Veltman, D. J., Aleman, A., Zitman, F. G., Penninx, B. W. J. H., van Buchem, M. A., Reiber, J. H. C., Rombouts, S. A. R. B., & Milles, J. (2009). Hierarchical functional modularity in the resting-state human brain. *Human Brain Mapping*, *30*(7), 2220–2231.

Flament, D., Ellermann, J. M., Kim, S. G., Ugurbil, K., & Ebner, T. J. (1996). Functional magnetic resonance imaging of cerebellar activation during the learning of a visuomotor dissociation task. *Human Brain Mapping*, *4*(3), 210–226.

Fortunato, S. (2010). Community detection in graphs. In *Physics Reports* (Vol. 486, Issues 3-5, pp. 75–174). https://doi.org/10.1016/j.physrep.2009.11.002

Fox, M. D., & Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews. Neuroscience*, *8*(9), 700–711.

Gerraty, R. T., Davidow, J. Y., Foerde, K., Galvan, A., Bassett, D. S., & Shohamy, D. (2018). Dynamic Flexibility in Striatal-Cortical Circuits Supports Reinforcement Learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *38*(10), 2442–2453.

Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, *48*(1), 63–72.

Haith, A. M., Huberdeau, D. M., & Krakauer, J. W. (2015). The influence of movement preparation time on the expression of visuomotor learning and savings. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(13), 5109–5117.

Hardwick, R. M., Rottschy, C., Miall, R. C., & Eickhoff, S. B. (2013). A quantitative meta-analysis and review of motor learning in the human brain. *NeuroImage*, *67*, 283–297.

Huberdeau, D. M., Krakauer, J. W., & Haith, A. M. (2019). Practice induces a qualitative change in the memory representation for visuomotor learning. *Journal of Neurophysiology*, *122*(3), 1050–1059.

Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., Della Penna, S., Duyn, J. H., Glover, G. H., Gonzalez-Castillo, J., Handwerker, D. A., Keilholz, S., Kiviniemi, V., Leopold, D. A., de Pasquale, F., Sporns, O., Walter, M., &

Chang, C. (2013). Dynamic functional connectivity: promise, issues, and interpretations. *NeuroImage*, *80*, 360–378.

Izawa, J., Criscimagna-Hemminger, S. E., & Shadmehr, R. (2012). Cerebellar contributions to reach adaptation and learning sensory consequences of action. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(12), 4230–4239.

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

Kim, H., & Park, H. (2007). Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis. In *Bioinformatics* (Vol. 23, Issue 12, pp. 1495–1502). https://doi.org/10.1093/bioinformatics/btm134

Kim, S., Ogawa, K., Lv, J., Schweighofer, N., & Imamizu, H. (2015). Neural Substrates Related to Motor Memory with Multiple Timescales in Sensorimotor Adaptation. *PLoS Biology*, *13*(12), e1002312.

Klein, A., Ghosh, S. S., Bao, F. S., Giard, J., Häme, Y., Stavsky, E., Lee, N., Rossa, B., Reuter, M., Chaibub Neto, E., & Keshavan, A. (2017). Mindboggling morphometry of human brains. *PLoS Computational Biology*, *13*(2), e1005350.

Korman, M., Doyon, J., Doljansky, J., Carrier, J., Dagan, Y., & Karni, A. (2007). Daytime sleep condenses the time course of motor memory consolidation. *Nature Neuroscience*, *10*(9), 1206–1213.

Krakauer, J. W. (2009). Motor learning and consolidation: the case of visuomotor rotation. *Advances in Experimental Medicine and Biology*, *629*, 405–421.

Krakauer, J. W., Ghilardi, M.-F., Mentis, M., Barnes, A., Veytsman, M., Eidelberg, D., & Ghez, C. (2004). Differential cortical and subcortical activations in learning rotations and gains for reaching: a PET study. *Journal of Neurophysiology*, *91*(2), 924–933.

Krakauer, J. W., & Mazzoni, P. (2011). Human sensorimotor learning: adaptation, skill, and beyond. *Current Opinion in Neurobiology*, *21*(4), 636–644.

Kuang, D., Ding, C., & Park, H. (2012). Symmetric Nonnegative Matrix Factorization for Graph Clustering. In *Proceedings of the 2012 SIAM International Conference on Data Mining*. https://doi.org/10.1137/1.9781611972825.10

Kuang, D., Yun, S., & Park, H. (2015). SymNMF: nonnegative low-rank approximation of a similarity matrix for graph clustering. In *Journal of Global Optimization* (Vol. 62, Issue 3, pp. 545–574). https://doi.org/10.1007/s10898-014-0247-2

Lanczos, C. (1964). Evaluation of Noisy Data. *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis*, *1*(1), 76–85.

Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, *401*(6755), 788–791.

Lindquist, M. A., Geuter, S., Wager, T. D., & Caffo, B. S. (2019). Modular preprocessing pipelines can reintroduce artifacts into fMRI data. *Human Brain Mapping*, *40*(8), 2358–2376.

Lloyd, S. (1982). Least squares quantization in PCM. In *IEEE Transactions on Information Theory* (Vol. 28, Issue 2, pp. 129–137). https://doi.org/10.1109/tit.1982.1056489

Makino, H., Hwang, E. J., Hedrick, N. G., & Komiyama, T. (2016). Circuit Mechanisms of Sensorimotor Learning. *Neuron*, *92*(4), 705–721.

Martin, T. A., Keating, J. G., Goodkin, H. P., Bastian, A. J., & Thach, W. T. (1996). Throwing while looking through prisms: I. Focal olivocerebellar lesions impair adaptation. *Brain: A Journal of Neurology*, *119*(4), 1183–1198.

McDougle, S. D., Bond, K. M., & Taylor, J. A. (2015). Explicit and Implicit Processes Constitute the Fast and Slow Processes of Sensorimotor Learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(26), 9568–9579.

McDougle, S. D., Ivry, R. B., & Taylor, J. A. (2016). Taking Aim at the Cognitive Side of Learning in Sensorimotor Adaptation Tasks. *Trends in Cognitive Sciences*, *20*(7), 535–544.

Miyamoto, Y. R., Wang, S., & Smith, M. A. (2020). Implicit adaptation compensates for erratic explicit strategy in human motor learning. *Nature Neuroscience*, *23*(3), 443–455.

Morehead, J. R., Qasim, S. E., Crossley, M. J., & Ivry, R. (2015). Savings upon Re-Aiming in Visuomotor Adaptation. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(42), 14386–14396.

Mucha, P. J., Richardson, T., Macon, K., Porter, M. A., & Onnela, J.-P. (2010). Community structure in time-dependent, multiscale, and multiplex networks. *Science*, *328*(5980), 876–878.

Norris, S. A., Hathaway, E. N., Taylor, J. A., & Thach, W. T. (2011). Cerebellar inactivation impairs memory of learned prism gaze-reach calibrations. *Journal of Neurophysiology*, *105*(5), 2248–2259.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, *10*(2), 59–63.

Reuter, M., Rosas, H. D., & Fischl, B. (2010). Highly accurate inverse consistent registration: a robust approach. *NeuroImage*, *53*(4), 1181–1196.

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. In *Journal of Computational and Applied Mathematics* (Vol. 20, pp. 53–65). https://doi.org/10.1016/0377-0427(87)90125-7

Sailer, U., Flanagan, J. R., & Johansson, R. S. (2005). Eye-hand coordination during learning of a novel visuomotor task. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *25*(39), 8833–8842.

Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X.-N., Holmes, A. J., Eickhoff, S. B., & Yeo, B. T. T. (2018). Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cerebral Cortex* , *28*(9), 3095–3114.

Schlerf, J., Ivry, R. B., & Diedrichsen, J. (2012). Encoding of sensory prediction errors in the human cerebellum. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(14), 4913–4922.

Seidler, R. D., & Noll, D. C. (2008). Neuroanatomical correlates of motor acquisition and motor transfer. *Journal of Neurophysiology*, *99*(4), 1836–1845.

Shadmehr, R. (1997). Neural Correlates of Motor Memory Consolidation. *Science*, *277*(5327), 821–825.

Shine, J. M., Koyejo, O., & Poldrack, R. A. (2016). Temporal metastates are associated with differential patterns of time-resolved connectivity, network topology, and attention. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(35), 9888–9891.

Smith, M. A., Ghazizadeh, A., & Shadmehr, R. (2006). Interacting adaptive processes with different timescales underlie short-term motor learning. *PLoS Biology*, *4*(6), e179.

Sporns, O. (2014). Contributions and challenges for network models in cognitive neuroscience. *Nature Neuroscience*, *17*(5), 652–660.

Taylor, J. A., & Ivry, R. B. (2014). Cerebellar and prefrontal cortex contributions to adaptation, strategies, and reinforcement learning. *Progress in Brain Research*, *210*, 217–253.

Taylor, J. A., Klemfuss, N. M., & Ivry, R. B. (2010). An explicit strategy prevails when the cerebellum fails to compute movement errors. *Cerebellum* , *9*(4), 580–586.

Taylor, J. A., Krakauer, J. W., & Ivry, R. B. (2014). Explicit and implicit contributions to learning in a sensorimotor adaptation task. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(8), 3023–3032.

Telesford, Q. K., Ashourvan, A., Wymbs, N. F., Grafton, S. T., Vettel, J. M., & Bassett, D. S. (2017). Cohesive network reconfiguration accompanies extended training. *Human Brain Mapping*, *38*(9), 4744–4759.

Trewartha, K. M., Garcia, A., Wolpert, D. M., & Flanagan, J. R. (2014). Fast but fleeting: adaptive motor learning processes associated with aging and cognitive decline. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(40), 13411–13421.

Tseng, Y.-W., Diedrichsen, J., Krakauer, J. W., Shadmehr, R., & Bastian, A. J. (2007). Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *Journal of Neurophysiology*, *98*(1), 54–62.

Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: improved N3 bias correction. *IEEE Transactions on Medical Imaging*, *29*(6), 1310–1320.

Wolpert, D. M., Diedrichsen, J., & Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews. Neuroscience*, *12*(12), 739–751.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. In *Journal of Comparative Neurology and Psychology* (Vol. 18, Issue 5, pp. 459–482). https://doi.org/10.1002/cne.920180503

Zhang, Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, *20*(1), 45–57.