Behavioral/Cognitive

Pre-supplementary Motor Cortex Mediates Learning Transfer from Perceptual to Motor Timing

Itzamná Sánchez-Moncada,[1] Luis Concha,[1,2] and Hugo Merchant[1]

[1]Instituto de Neurobiología, Querétaro 76230, México and [2]International Laboratory for Brain, Music and Sound (BRAMS), Montreal, Québec H2V 2S9, Canada

When we intensively train a timing skill, such as learning to play the piano, we not only produce brain changes associated with task-specific learning but also improve our performance in other temporal behaviors that depend on these tuned neural resources. Since the neural basis of time learning and generalization is still unknown, we measured the changes in neural activity associated with the transfer of learning from perceptual to motor timing in a large sample of subjects ($n = 65$; 39 women). We found that intense training in an interval discrimination task increased the acuity of time perception in a group of subjects that also exhibited learning transfer, expressed as a reduction in inter-tap interval variability during an internally driven periodic motor task. In addition, we found subjects with no learning and/or generalization effects. Notably, functional imaging showed an increase in pre-supplementary motor area and caudate–putamen activity between the post- and pre-training sessions of the tapping task. This increase was specific to the subjects that generalized their timing acuity from the perceptual to the motor context. These results emphasize the central role of the cortico-basal ganglia circuit in the generalization of timing abilities between tasks.

Key words: **fMRI; motor practice; motor timing; SMA; time perception**

Significance Statement

Intensive training in a task can lead to improvements in other behaviors when the neural resources are shared between conditions. Hence, the learning generalization strategy is now actively used in interventions to improve timing behaviors across tasks. Here we show that timing precision enhancement after interval discrimination training can be transferred as a decrease in temporal variability during a tapping task in a subgroup of subjects. Crucially, the generalization from perceptual to motor timing increased activity in the pre-supplementary motor area and caudate–putamen in that subgroup. These findings support the notion that magnified recruitment occurs in the cortico-basal ganglia circuit when an acquired perceptual timing ability is transferred to a motor timing task.

Introduction

The human brain can flexibly quantify time across complex perceptual and motor behaviors such as musical appreciation and execution. These behaviors demand the development of sophisticated skills to extract the beat or isochronous pulse of intricate musical patterns and produce predictive movements entrained in the beat (Honing and Merchant, 2014; Mendoza and Merchant, 2014; Lenc et al., 2021). Hence, temporal learning and processing are critical elements of human intelligence that have been investigated for decades (Treisman, 1963; Herholz and Zatorre, 2012; Ayala et al., 2017; Balasubramaniam et al., 2021). The classical view from experimental psychology of a common clock for timing across sensory and motor tasks (Kristofferson, 1980; Ivry and Hazeltine, 1995; Gibbon et al., 1997) has been replaced by imaging and neurophysiological studies supporting the idea of a partially distributed neural timing circuit that has two elements (Rao et al., 1997; Jantzen et al., 2002; Macar et al., 2006; Coull et al., 2008; Wiener et al., 2010; Merchant et al., 2013b). The first element is the core timing network, integrated by key areas of the motor system, namely, the supplementary motor areas (SMA-proper and pre-SMA), the cerebellum, and the cortico-thalamic-basal ganglia circuit (Merchant et al., 2014a, 2015a; Merchant and Bartolo, 2018; Tanaka et al., 2021). This core timing network is involved in

Received Dec. 21, 2020; accepted Nov. 21, 2023.

Author contributions: I.S.-M., L.C., and H.M. designed research; I.S. and H.M. performed research; I.S.-M., L.C., and H.M. analyzed data; H.M. wrote the paper.

We thank Jennifer Coull, Víctor de Lafuente, and Juan Fernández for their fruitful comments on the manuscript and Jessica Gonzalez Norris for proofreading the manuscript. We also thank Luis Prado, Raúl Paulín, Erick Pasaye, Juan Ortiz, Leopoldo González, Luis Aguilar, and Alejandro de León, for their technical assistance. This work was supported by Consejo Nacional de Ciencia y Tecnología (CONACYT) Grant CONACYT: A1-S-8430, Programa de Apoyo a Proyectos de Investigacion e Inovacion Tecnologica UNAM-DGAPA-PAPIIT IN201721, and SECITI 2342 to H.M. L.C. is partially funded by CONACYT (C1782) and Programa de Apoyo a Proyectos de Investigacion e Inovacion Tecnologica UNAM-DGAPA-PAPIIT (AG200117, IN204720). I.S.-M. is a doctoral student from Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM) and the recipient of CONACYT Fellowship 298046. The National Laboratory for MRI is supported by CONACYT and UNAM.

The authors declare no competing financial interests.

Correspondence should be addressed to Hugo Merchant at hugomerchant@unam.mx.

https://doi.org/10.1523/JNEUROSCI.3191-20.2023

temporal processing in a wide range of perceptual and motor timing behaviors on the scale of hundreds of milliseconds, including visual, auditory, and tactile stimuli and a variety of motor effectors (Wiener et al., 2010; Merchant et al., 2013a; Merchant and Averbeck, 2017). The second element is represented by areas selectively engaged in the specific behavioral requirements of a task (Buhusi and Meck, 2005; Coull et al., 2011; Harrington et al., 2011). These task-dependent areas interact with the core timing system to produce the characteristic pattern of performance variability of a specific timing paradigm (Merchant et al., 2008b, 2013a).

The notion of a core timing network has been also supported by experiments that evaluate learning and generalization of timing (Bueti and Buonomano, 2014). The hypothesis behind these studies is that the learning-based improvements in temporal processing within a particular task will transfer to another timing behavior if they share trained neural circuit resources. Normally, learning transfer is quantified as an increase in time precision when comparing temporal performance in the generalization task between post-training and pre-training sessions. This strategy is common in the artificial neural network literature. Namely, after training a recurrent neural network in a condition with specific input–output rules, the network is tested in other conditions to determine generalization capabilities due to common neural weights and shared internal dynamics (Laje et al., 2018; Pérez and Merchant, 2018; Bi and Zhou, 2020; Merchant and Pérez, 2020). Thus, robust temporal generalization, measured from intensive training in time discrimination, has been documented as an increase in timing acuity across auditory frequencies (Wright et al., 1997; Karmarkar and Buonomano, 2003), sensory modalities (Nagarajan et al., 1998; Westheimer, 1999; Bartolo and Merchant, 2009), stimulus locations (Nagarajan et al., 1998), and sensory to motor timing tasks, which is particularly relevant to the present study (Meegan et al., 2000; Planetta and Servos, 2008; Fabio et al., 2011). These findings strongly support the existence of a multimodal and multi-context core timing network (Merchant et al., 2008a; Wiener et al., 2010; Merchant and Yarrow, 2016).

For this study, we recruited 65 healthy human subjects who underwent intensive interval discrimination training for a week and performed pre- and post-training sessions of a synchronization–continuation tapping task inside an fMRI scanner. We found that half of the participants showed learning gains in the precision of interval discrimination. These gains were transferred to the temporal execution of a motor task with initial tapping synchronization to a metronome, followed by a self-driven rhythmic response. In addition, we found groups of subjects behaving as non-learners (NL) and covert rhythmic-skill (CR) learners. Furthermore, we tested 29 additional subjects that performed the synchronization–continuation tapping task inside an fMRI scanner for two sessions separated by a week but with no training in the time discrimination task. This group served as a control. Then, we focused on the change in hemodynamic responses associated with the transfer of learning from perceptual to motor timing and compared them with the brain activation profiles of the NL, CR learners, and the control populations in the post- versus pre-training sessions.

Materials and Methods

Experimental design

For the current research, a pre-training/training/post-training intervention was implemented (Wright et al., 1997; Bartolo and Merchant, 2009). In the first session, subjects performed the synchronization–continuation task (SCT; pre-training) within the MRI scanner. Later that day, subjects started their first training session of a 7-day interval discrimination task (IDT) training program. On the 7th day, after completing the IDT, subjects performed the second SCT session (post-training) inside the MRI scanner (Fig. 1*A*). The experimental group performed the fully experimental paradigm and was formed by 69 right-handed healthy subjects (42 women) with a mean age of 27 years (age range, 20–34 years), no record of neurological or psychiatric disorders, and normal or corrected-to-normal vision. On the other hand, the control group, made up of 28 subjects (16 women) with a mean age of 27 years (age range, 20–34 years), only performed the two SCT sessions (1 week apart) with no IDT training. All subjects gave written informed consent for the study protocol, which was approved by the bioethics research committee of the Institute of Neurobiology, UNAM. The study was performed in accordance with the ethical principles of the Declaration of Helsinki.

Both tasks were programmed using MATLAB R2013a and the Psychtoolbox library (Brainard, 1997). A Dell XPS Intel Core i5 laptop with Windows 7 was used to run the tasks. During the IDT, all participants were seated comfortably on a chair facing the 15-in laptop in a quiet experimental room; only the space bar and left and right arrow keys were unlocked. We employed empty intervals, which were delimited by a 3.77×3.77 cm^2 gray square that flashed at the center of a black screen. Each marker was displayed for 33 ms (screen resolution was $1{,}366 \times 768$ pixels, and the refresh rate was 60 Hz). As mentioned before, the SCT was performed inside the scanner and presented using a pair of video goggles, which were binocular LED screens with diopter correction (VisualSystem, NordicNeuroLab). The subjects' responses were registered using a handheld response collection device (ResponseGrip, NordicNeuroLab). At the beginning of each trial, subjects were instructed to fixate on an isometric white cross (1.2 cm) that appeared at the center of the black screen. After a variable period (1.2–2.4 s), a 3.77×3.77 cm^2 gray square was presented in sequence as a metronome with an isochronous interstimulus interval of 850 ms.

Tasks and training

IDT

Subjects had to discriminate between two intervals and determine which one had the longest duration (Fig. 1*A*,*B*). One of the intervals had a constant duration of 850 ms (base interval), as it has been shown that it has a wide time generalization profile compared to the commonly used intervals in the hundreds of milliseconds (Bartolo and Merchant, 2009). The comparison intervals were selected pseudo-randomly without repetition from the following eight values, 566, 666, 783, 816, 883, 916, 1,033, and 1,330 ms, which were carefully calculated to maximize the threshold boundaries (Merchant et al., 2008b). We use the term "repetition" to refer to the subsequent random presentation of the eight comparison intervals. The first standard or comparison intervals were presented randomly. The first interval was presented three consecutive times, whereas the last interval was presented only once (Fig. 1*B*). To measure the response time, subjects were asked to press and hold the spacebar at the start of each trial. Then, subjects had to release the spacebar and press the left or right arrow key to indicate whether the first or second interval was longer, respectively. All actions were performed with the right hand. At the end of each trial, feedback was given on whether the response was correct or incorrect. During a training session, the subjects completed 4 blocks of 10 repetitions (320 total trials, with a duration of ~60 min per training session for 7 d).

SCT

Subjects were lying down inside the scanner with the video goggles comfortably adjusted and were instructed to entrain to the visual metronome by pressing a button with their right index finger. The visual metronome stopped after nine synchronized taps (synchronization epoch), and the subjects were required to continue pressing the button for another 12 taps (continuation epoch), trying to maintain the same beat (Fig. 1*C*). The fixation cross was present during both epochs. In the continuation epoch, once the subject pressed the button for the 12th time, the fixation cross disappeared, and the mean inter-tap interval (ITI) was calculated and presented to the subject as feedback for 2 s. Afterward, the screen was completely black for 10 s (inter-trial interval), and then the white

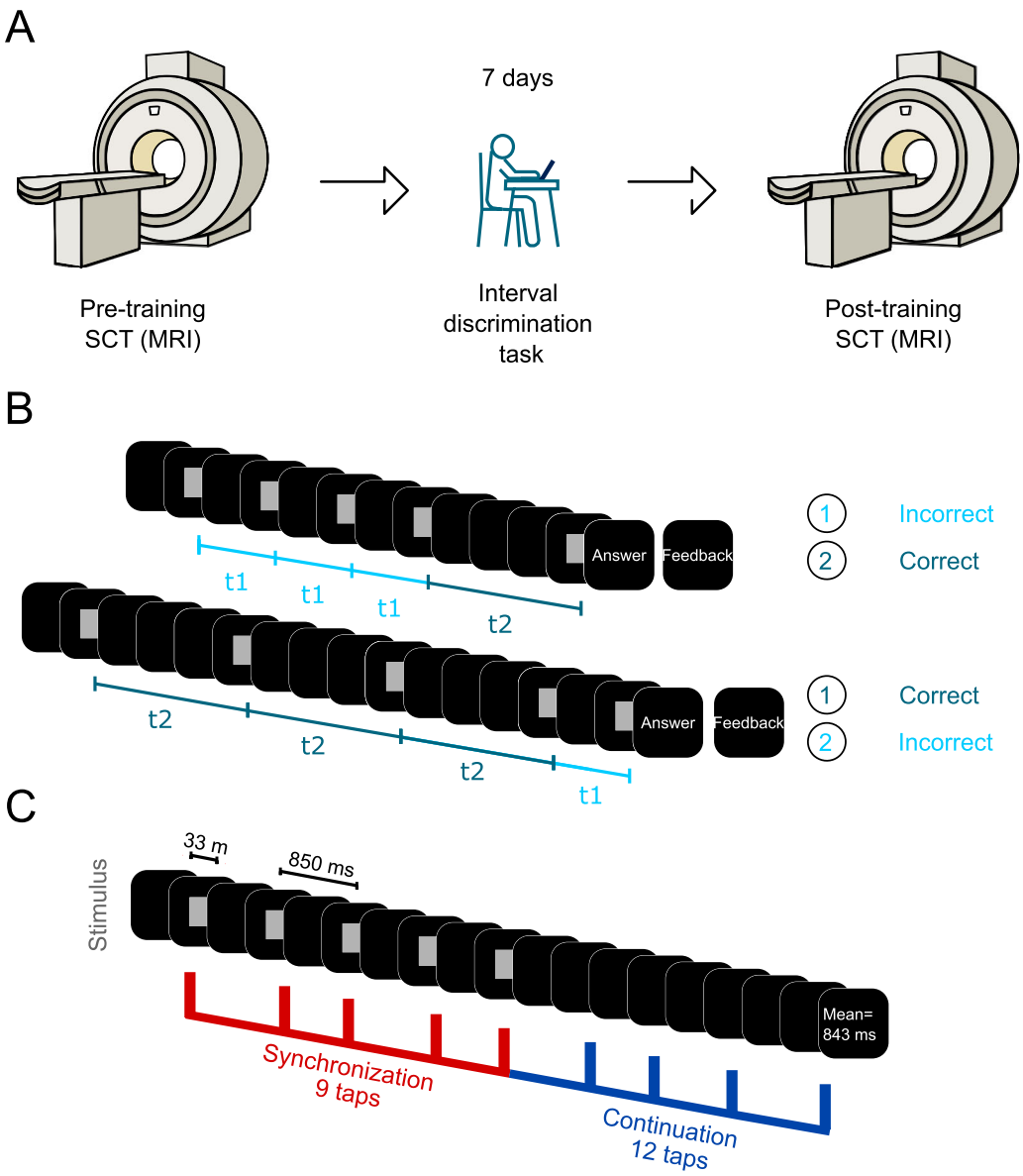

Figure 1. Experimental paradigm. ***A***, Timeline of the SCT pre- and post-training sessions and the IDT training. First, subjects ($n = 65$) performed three blocks of the SCT inside the MRI scanner; afterward, subjects underwent a 1-week training in the IDT. Finally, subjects performed the second SCT session. ***B***, IDT. Subjects were presented with two intervals of different durations and had to choose which one was the longest. The first interval was always presented three times. t1 corresponds to the shortest interval, while t2 refers to the longest interval. Intervals were chosen randomly across trials. The base interval was 850 ms, while the comparison intervals were selected from a list of four shorter and four longer intervals with respect to the base interval. ***C***, SCT. A visual metronome (gray square) was presented in the middle of a black screen with an interstimulus interval of 850 ms. Subjects had to synchronize to the metronome, pressing a button for 9 taps, after which they had to continue pressing the button for 12 taps. Feedback was provided as the mean ITI generated for both epochs. Subjects were instructed to accurately produce intervals of 850 ms.

cross appeared again, signaling the start of the next trial. If the asynchronies (time between the visual cue and response) were greater than ±425 ms, the trial was excluded from the behavioral and image analyses. Three runs were performed per SCT session, the first one with 20 trials and the rest with 16 trials. Each run lasted ~10 min. Total scanning time was about 60 min, including functional, anatomical, and diffusion images.

MRI acquisition

Images were acquired in the National Laboratory for Magnetic Resonance Imaging at UNAM, using a 3.0 T Philips Achieva TX system equipped with a 32-channel head coil. A gradient echo echo-planar imaging sequence was performed to acquire T2*-weighted images (TR = 2 s, TE = 30 ms; voxel resolution = $2 \times 2 \times 4$ mm^3). A total of 32 axial slices comprised each EPI volume. The size of the volume allowed us to scan the entire cerebrum and most of the cerebellum (below lobule VIIB). Five dummy volumes were acquired at the beginning of the run for T1 equilibration. In addition, a three-dimensional spoiled gradient-recalled echo sequence was used to obtain high-resolution T1-weighted images with a 1 mm^3 resolution (TR = 8.15 ms, TE = 3.75 ms; image matrix = $256 \times 256 \times 176$), which was used for image registration purposes.

Statistical analysis

Behavioral data

IDT. The method of constant stimuli was used to estimate the daily thresholds (Getty, 1975). The difference threshold was computed from the psychometric curve, where the probability of long-interval discrimination was plotted as a function of the comparison interval (Merchant et al., 2008b; Méndez et al., 2014). A logistic function was fitted to the data,

and the threshold corresponded to half the subtraction of the interval at 0.75 and that at 0.25 of the probability of answering long (Fig. 2*A*). The threshold was computed for each of the four blocks per day. Then, we plotted the threshold across the 7 d of training and fitted a power function ($y = Ax^B$, where y, threshold; A, the intercept; x, training days; B, first polynomial coefficient). The learning criteria consisted of (1) a significant main effect of training in an ANOVA using threshold as the dependent variable and training session as the factor and (2) a significant fit of the power function with a negative slope, implying an improvement in the subject's ability to discriminate the stimuli.

SCT. The first four trials of run 1 of the SCT were not included in the analysis to obtain data from a steady behavioral response. In total, 48 trials were analyzed (3 runs of 16 trials each). The synchronization epoch included nine taps and eight ITIs, but the first tap and ITI were discarded. The continuation epoch consisted of 12 taps and 12 ITIs, but the last ITI was not included in the analysis. In addition, trials were not further analyzed when asynchronies were above ±425 ms (half the duration of the interstimulus interval) or a single ITI was longer than 850 ms ± 400 ms. Hence, we obtained an uneven number of ITIs and tapping times per subject. Consequently, a bootstrap resampling method (10,000 iterations) was carried out to get homogenous data across subjects.

For each subject, we compared the following SCT performance measures between pre- and post-sessions: asynchronies, constant error, and temporal variability. Asynchronies were the time difference between tap and stimulus onsets and thus were computed only for the synchronization epoch. Constant error was the average difference between ITIs and the instructed interval. Temporal variability was defined as the standard deviation of ITIs. We also computed the temporal variance ratio (TVR), which is the ratio of the ITI variance of the pre-session divided by the post-session variance. Therefore, a TVR value below one corresponds to an increase, whereas a value above one corresponds to a decrease in temporal variability in the post-session with respect to the pre-session. Constant error and temporal variability were calculated separately for the synchronization and continuation epochs.

Asynchrony values were presented as phases with respect to the beat onset times over the instructed interval. Asynchronies were transformed from milliseconds (a_i) to angular units in radians (θ_i) with the equation $\theta_i = (2\pi \times a_i)/T_i$, where T_i corresponded to 850 ms, the target interval. Circular statistics were used to summarize the distribution of the relative

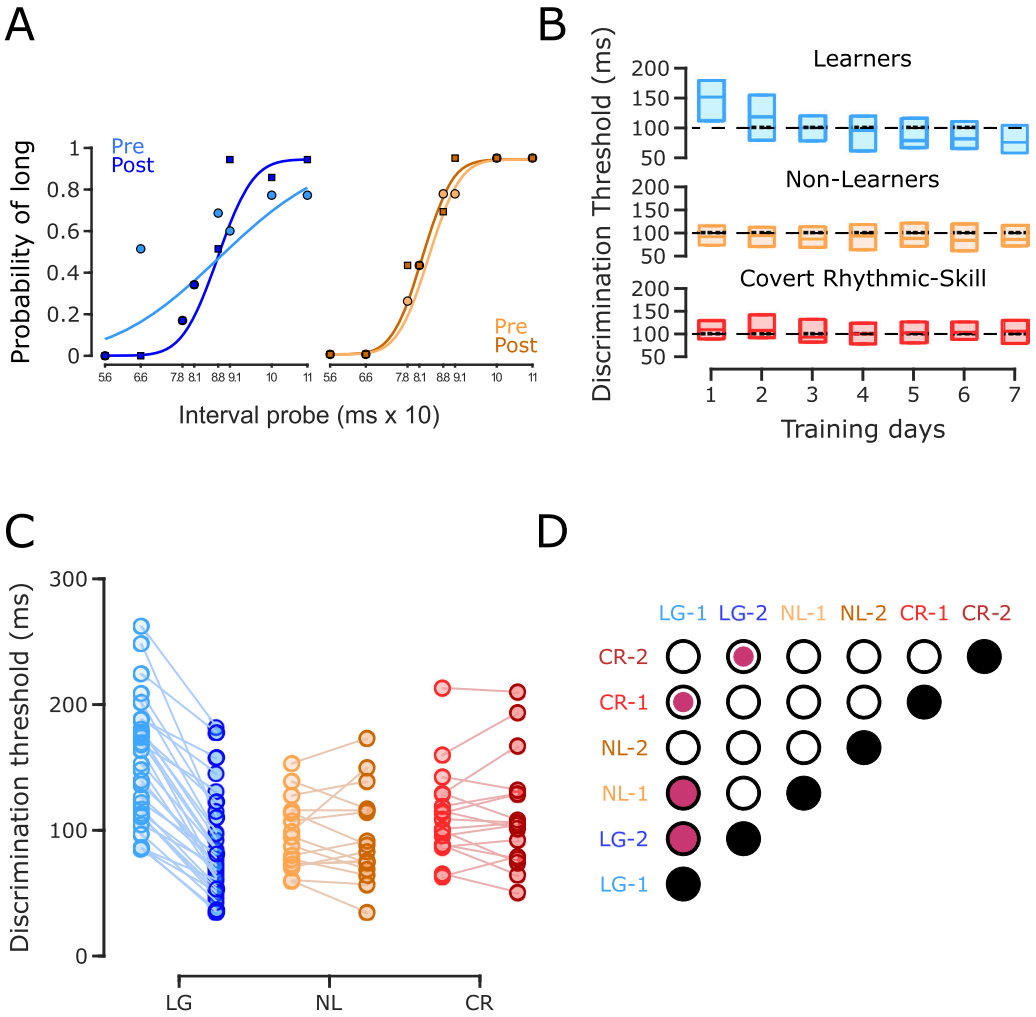

Figure 2. IDT behavioral performance. ***A***, Psychometric functions. The first and last psychometric function (one block) for a subject with a significant reduction of its discrimination threshold (left panel) and a subject without a significant reduction of the same measure (right panel). ***B***, Group discrimination threshold as a function of training day. Interquartile boxplot of the discrimination threshold for subjects that significantly improved their performance in the IDT (learners [LG], $n = 32$; top) and subjects without any improvement during the IDT, which were subdivided into NL ($n = 16$) and CR subjects ($n = 17$), based on their pre- and post-SCT performance (Fig. 3*A*). The black dotted line is at 100 ms for visual reference. ***C***, Discrimination threshold across subjects. Discrimination thresholds for the first (light colors) and last (dark color) IDT training sessions connected by a line for each subject of the three groups. ***D***, Matrix of statistical differences between groups. Each element of the matrix corresponds to the pairwise comparison (t test) between the three groups for the initial (1) and final (2) training sessions in the IDT. Empty circles depict nonsignificant effects, small filled circles indicate a significant effect at $p < 0.05$, and large filled circles indicate a significant effect at $p \leq 0.005$. Note the large decrease in discrimination thresholds for LG subjects and the slight increase in NL and CR subjects in the last IDT session with respect to the initial IDT session.

phases on the unit circle using the mean resultant vector, which has two parameters: the length R (dimensionless ranging from 0 to 1) and angle (given in radians from 0 to 2π). R equal to 0 means phases in asynchronies that are uniformly distributed along the whole inter-onset interval, whereas an R value of 1 indicates identical phases (Fig. 3*A*). A vector angle of 0 means a perfect temporal alignment between tap and stimulus, while positive and negative angles indicate that the tap followed (positive asynchronies) or preceded (negative asynchronies) the stimulus, respectively (Gámez et al., 2018). Asynchronies were analyzed with MATLAB Circular Statistics Toolbox. The Rayleigh test was used to assess unimodality with the null hypothesis of a uniform distribution around the circle. Differences between sessions and groups were determined using the Harrison–Kanji test, which is a parametric two-way ANOVA for circular data, using session (pre-training and post-training) as the within-subjects factor and group as the between-subjects factor.

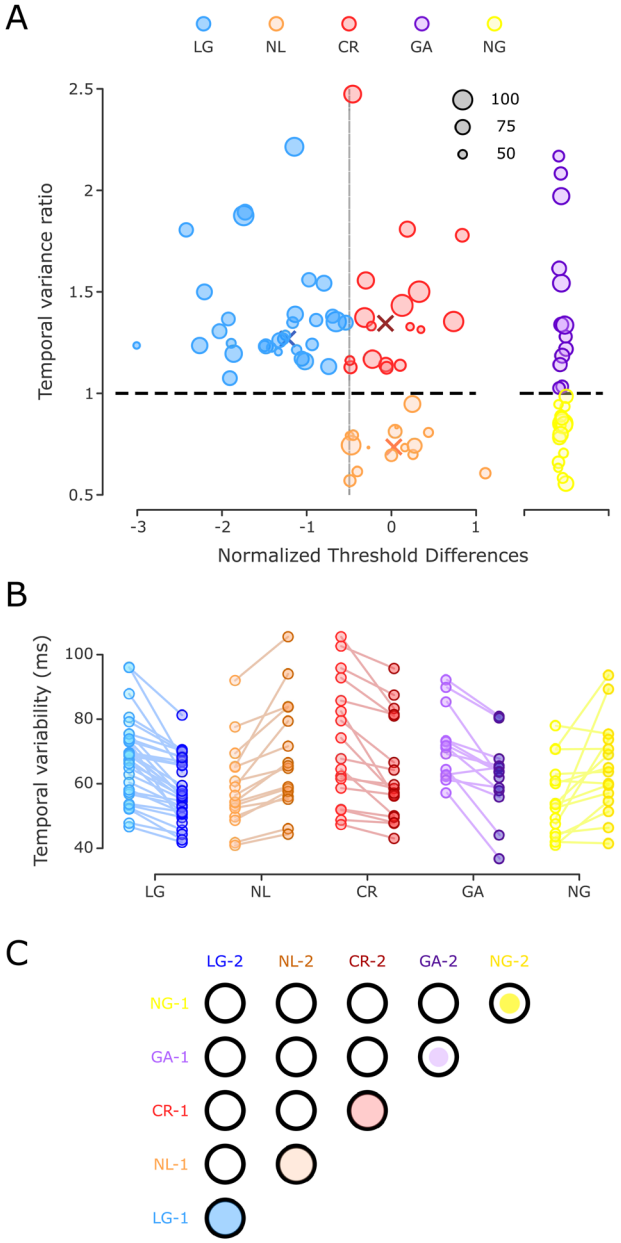

Figure 3. Temporal variability in the SCT and clustering of subjects. ***A***, *Left*, A *K*-means cluster classification identified three groups, plotting the TVR as a function of the normalized threshold differences. Blue, orange, and red dots correspond to learners with generalization, NL, and CR learners, respectively. The diameter of the circles was determined by the initial inter-tap variability (in ms). The higher the variability, the bigger the circle (black circles serve as a size guide). The horizontal black dotted line corresponds to a TVR of 1 (no IDT learning). Colored X symbols mark the centroid assigned to each group. *Right*, Two groups of control subjects were identified: a group with a significant decrease in temporal variability in the post-training SCT session, GA ($n = 13$, purple circles), and another with no changes in temporal variability between sessions, called NG ($n = 15$, yellow circles). ***B***, Temporal variability during the pre- and post-training SCT across subjects. Changes in temporal precision are depicted for the five groups using the color code in A with light color for pre-training and dark color for post-training. ***C***, Matrix of statistical differences between groups. Each element of the matrix corresponds to the pairwise comparison between the five groups for the pre-training (1) and post-training (2) sessions in the SCT. Notation as in Figure 3*D*.

Three-way repeated-measures ANOVAs were carried out using constant error and temporal variability as dependent variables, session (pre- and post-training) and epoch (synchronization and continuation epochs) as the within-subjects factor, and group [learners, NL, CR learners, gainers (GA), and non-gainers (NG); see below] as the between-subjects factor.

Post hoc paired t tests were used to assess differences between groups and sessions. Routines for statistical analysis were written using MATLAB R2013a. The statistical level to reject the null hypothesis was $\alpha = 0.05$. The Greenhouse–Geisser test in the repeated-measures ANOVAs was used to correct probability levels from deviations in sphericity.

Behavioral clustering

In Figure 3*A*, we plotted the TVR of the continuation epoch of the SCT as a function of the normalized threshold difference (Z score) between the first and last days of training in the IDT. The former is a measure of temporal generalization, with values above one indicating a decrease in temporal variability in the post-training session with respect to the pre-training session. The latter is a measure of temporal learning, with values below zero indicating an increase in temporal acuity because of a decrease in the discrimination threshold after the daily intensive training.

We found the two expected groups of subjects based on previous studies (Meegan et al., 2000; Planetta and Servos, 2008). First, we identified a group of learners with a negative threshold difference statistically different from zero (see also the above learners' criteria) with a concomitant time generalization effect, where the TVR was larger than one, and a significant effect of the session (permutation test). This group was called learners with generalization (LG, $n = 32$; Fig. 3*A–C*, blue dots). Second, we found a group of subjects with no learning, a threshold difference that was not statistically different from zero, and no time generalization (TVR below one). We named this group NL ($n = 16$; Fig. 3*A–C*, orange dots). Notably, we also found a third group, which we called CR learners ($n = 17$; Fig. 3*A–C*, red dots). The subjects in this group were NL with a TVR that exhibited a significant decrease in time variability during the post-training session. We ran K-means clustering with a "city block distance" metric and $k = 1{:}5$. With $k = 3$, we obtained the lowest Bayesian information criterion of −24.1776. The centroids in x and y coordinates (normalized threshold difference and TVR, respectively) were centroid 1 = [−1.2, 1.2], centroid 2 = [−0.3, 0.7], and centroid 3 = [0.1, 1.3], which correspond to the LG, NL, and CR groups, respectively (Fig. 3*A*).

fMRI data analysis

Preprocessing. Pre- and post-training functional imaging data were analyzed using the Oxford Centre for Functional MRI of the Brain Software Library v5.0 (FSL). All EPI volumes were time and motion corrected. All images were resampled to a 2 mm isotropic voxel size and spatially smoothed using an isotropic Gaussian kernel of 6 mm full-width half-maximum to increase their signal-to-noise ratio. A low-frequency filter was adjusted to the data for any physiological drift (high-pass filter of 100 s). In parallel, fMRIprep was used to perform the quality check of the images and calculate possible confounding factors. Nuisance regressors included six motion parameters, estimated by MCFLIRT motion correction (Jenkinson et al., 2002); the first six aCompCor physiological noise regressors (Behzadi et al., 2007); and framewise displacement (Power et al., 2012).

First-level analysis. An event-related analysis was carried out (Woolrich et al., 2001). Three regressors were used to model the synchronization, continuation, and feedback epochs. All tapping responses were modeled as one event of the correspondent epoch. Each regressor was convolved with a double gamma function that accounted for the hemodynamic response function.

Statistical parametric maps derived from the general linear model were created for each subject during task performance. T statistics were calculated and then transformed into Z score maps. A first-level analysis was run for each of the 93 subjects to define patterns of activation as compared to the baseline.

Second-level analysis. Contrast parameter estimate (COPE) maps of the first-level analysis for the first and second runs were averaged across trials for each session and subject. These COPEs were utilized to perform the rest of the second- and third-level analyses (unless indicated otherwise; Woolrich et al., 2004).

As an initial step, the mean group activation for all subjects was calculated for the synchronization and continuation epochs of the pre-training session. This was done to elucidate the common areas activated for both cognitive processes (Fig. 7). For multiple comparisons corrections, we used random field theory to calculate the minimum cluster size with a cluster-forming threshold Z score of 2.57 ($p = 0.005$) and a family-wise error rate (FWER) of 0.05. These were the parameters utilized to correct for multiple comparisons for all second- and third-level analyses.

Additionally, we ran a mixed model to test for differences between sessions for each behavioral group. The COPE values of the first-level analysis were used as dependent variables, the session (pre-training vs post-training) was modeled as a fixed variable, and each subject was used as the random variable between sessions (COPE∼Session + (Session | ID). Synchronization and continuation epochs were analyzed separately using the average of all trials. Models were run using Voxel: Mass-Univariate Voxelwise Analysis of Medical Imaging Data for R (https://github.com/angelgar/voxel). All statistical maps were corrected for multiple comparisons using AFNI's 3dClustSim to estimate the minimum cluster size with a threshold Z score of 2.57 and an FWER of 0.05 (Eklund et al., 2016; Cox et al., 2017).

Multi-voxel pattern analysis

Selection of regions of interest. The first step was to identify the voxels whose activity could accurately classify the three main groups of subjects (LG, NL, and CR). We did this using the difference in COPE values between the pre- and post-training sessions from the second-level analysis of the continuation epoch. Only the voxels with significant mean activation (depicted in Fig. 7) were included in the analysis. The searchlight function from the CosmoMVPA toolbox (MATLAB) was employed to identify the voxels with the best classification performance, and split-half correlation was used to identify statistical significance (Haxby et al., 2001). Areas of interest included the bilateral SMA, bilateral pre-SMA, bilateral caudate, bilateral putamen, left insula, left intraparietal lobule, left thalamus, right V1 and V4 cortex, right dorsolateral prefrontal cortex, and right cerebellum. From the voxels of the highest significance, we created cubes of three voxels per side. A total of 45 cubes were created along the aforementioned areas, carefully preventing overlapping.

Support vector machine. With a support vector machine (SVM), we classified the three groups of subjects. We used all the possible combinations of the 45 ROIs to construct models with four ROIs (four partial ROIs using all possible permutations from the total 45 ROIs = 1,48,995 iterations). The SVMs were carried out using 10-fold cross-validation, and the models with an accuracy above 60% were identified. The proportion of times that an area was present in these high-accuracy models was estimated. We ended up with 15 areas persistently included in the high-accuracy models. Additionally, as a part of the validation process, we ran 10,000 SVM iterations, randomly selecting 16 out of the 32 LG subjects and 16 out of the 17 CR subjects to compare the accuracy of our model with all the data and an equal number of subjects across groups. Finally, another 10,000 iterations were run, but we randomly assigned the group labels to calculate the power of classification at chance levels.

Results

Behavioral data

We tested 65 healthy subjects in a protocol that consisted of 2 SCT sessions inside the MRI scanner, which were conducted 7 d apart. Between one session and the other, the subjects performed intensive IDT training for more than an hour (Fig. 1*A*). During the SCT, a visual metronome (gray square) was presented in the middle of a black screen with an interstimulus interval of 850 ms. The subjects had to synchronize to the metronome (synchronization epoch) by pressing a button for nine taps. Afterward, they had to continue pressing the button for 12 taps without a metronome (continuation epoch; see Materials and Methods, Fig. 1*C*). During the IDT, subjects were presented with two intervals and had to choose which one was longer. The first interval was always presented three times and could be either the base interval (850 ms) or the comparison interval. Comparison intervals were selected from a list of four shorter and four longer intervals with respect to the base interval (Fig. 1*B*).

The first goal of this study was to determine whether intensive practice improved interval discrimination performance. Some subjects showed an increasing psychometric function slope as training progressed (Fig. 2*A*, left panel), which is consistent with a decrease in the discrimination threshold due to training. Other subjects, however, failed to exhibit this progressive increase (Fig. 2*A*, right panel). Thus, two general groups of subjects were observed: learners [LG; $n = 32$; power regression: $m = -0.2779$, $R^2 = 0.9926$, $p < 0.0001$ (Fig. 2*B*, top panel)], who met the learning criteria (see Materials and Methods), and NL ($n = 33$; power regression: $m = 0.0203$, $R^2 = 0.0482$, $p = 0.6361$ (Fig. 2*B*, middle and bottom panels)], who did not. As described below, we identified two NL subgroups based on their performance during the SCT pre- and post-training sessions: the proper NL and the CR subjects. Both showed a low initial interval discrimination threshold, an unchanged threshold as a function of the training day (Fig. 2*B*), and a slight increase in this parameter between the initial and final training sessions (Fig. 2*C*,*D*). In contrast, the LG group showed larger initial discrimination thresholds and a dramatic increase in interval acuity with training (Fig. 2*C*,*D*).

The next step was to determine whether improved performance in the time perception task could be accompanied by a gain in the SCT. The hypothesis was that the LG group would decrease the temporal variability of their ITIs during the post-training session of the SCT, whereas the NL group would show similar temporal variability between the post- and pre-training sessions. Consequently, in Figure 3*A*, we plotted the normalized difference in the discrimination threshold between the last and first days of IDT training against the TVR of the continuation epoch of the SCT (see Materials and Methods). We chose continuation variability over synchronization data as the changes between sessions were greater in the former (see below). TVR is a measure of temporal generalization (the pre-training/post-training ITI variance ratio), with values above one indicating a decrease in temporal variability in the post-training with respect to the pre-training. Complex intersubject differences are evident in Figure 3, and an iterative K-means clustering determined the existence of more than two groups. Values from one to five

were tested to find the best number of groups for our data. The smaller Bayesian information criterion (−24.1776) indicated that three was the most parsimonious clustering index, as follows: (1) 32 subjects who reduced their discrimination threshold and ITI variability (LG with a cluster centroid $x = -1.2801$, $y = 1.2359$); (2) 16 subjects who did not reduce their discrimination threshold or ITI variability (NL with a cluster centroid in $x = 0.0182$, $y = 0.7380$); and (3) 17 subjects who did not reduce their discrimination threshold but were able to reduce their ITI variability (CR with a cluster centroid in $x = -0.1173$, $y = 1.3486$).

Since the CR group showed a gain in temporal variability in the SCT between the pre- and post-training sessions without learning effects on the IDT, we hypothesized that these subjects experienced operational learning from practicing the SCT twice. Consequently, we ran a control group of subjects that executed the SCT in the MRI scanner twice, 1 week apart but with no IDT training. As predicted, approximately half of the subjects ($n = 15$; called NG) showed no changes in temporal variability across the two SCT sessions, whereas the other half [$n = 13$; GA] showed a statistically significant decrease in temporal variability in the second session (Fig. 3*A–C*), probably due to the benefit of practicing the SCT in the first session.

Once the three groups were defined, we tested for statistically significant differences between groups during the IDT. A two-way repeated-measures ANOVA on the interval discrimination threshold showed significant main effects for session ($F_{2, 62} = 51.786$, $p < 0.0001$) and for session × group interaction ($F_{2, 62} = 63.289$, $p \leq 0.0001$) (Fig. 2*D*). Post hoc paired t tests showed a significant threshold reduction for the LG group ($t_{31} = 13.8507$, $p < 0.0001$). No significant changes were found for the NL ($t_{15} = 0.1284$, $p = 0.8995$) or CR ($t_{16} = -0.0734$, $p = 0.9424$) groups (see the significant matrix in Fig. 2*D*). Additionally, pre-training discrimination thresholds were significantly high for the LG group compared to those of the NL ($t_{46} = 4.2605$, $p < 0.0001$) and CR groups ($t_{47} = 2.8884$, $p = 0.0058$). In addition, there were not significant differences between NL and CR ($t_{31} = -1.5069$, $p = 0.1420$; Fig. 2*D*). Hence, these results confirm the existence of a large group of learners with a higher initial discrimination threshold that is reduced after a week of intense interval discrimination training. On the other hand, the NL and CR subjects started training with a significantly low initial discrimination threshold, indicating high baseline discriminant capabilities that could account for their inability to improve their performance after training due to a floor effect.

The SCT is an explicit timing task that allowed us to gather the following behavioral parameters: asynchronies, constant error, and temporal variability (see Materials and Methods for definitions). As all participants (experimental and control) performed both sessions of the SCT, the following analyses were carried out to compare them.

The mean asynchronies (see Materials and Methods) during the synchronization epoch were plotted as relative phases on the unit circle across groups and sessions (Fig. 4). We found that the mean resultant was close to one, indicating a consistent synchronization to the metronome across all groups [Rayleigh test for the pre-training asynchronies: $z = 23.7601$, $p < 0.0001$ (LG); $z = 14.6747$, $p < 0.0001$ (NL); $z = 15.0941$, $p < 0.0001$ (CR); $z = 12.0283$, $p < 0.0001$ (GA); $z = 12.8848$, $p < 0.0001$, (NG). Post-training asynchronies: $z = 28.4821$, $p < 0.0001$ (LG); $z = 14.8011$, $p < 0.0001$ (NL); $z = 15.3140$, $p < 0.0001$ (CR); $z = 11.7674$, $p < 0.0001$ (GA); $z = 14.5736$, $p < 0.0001$ (NG)]. In addition, the five groups in the pre- and post-training sessions showed negative mean circular asynchronies [the one-sample

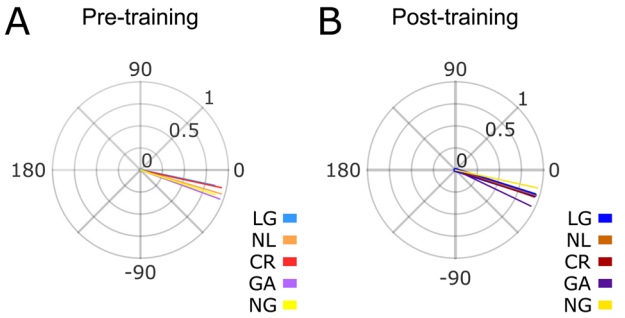

Figure 4. Asynchronies. Asynchronies for the synchronization epoch for the five groups during the pre- (***A***) and post- (***B***) training sessions. No significant differences were found between groups or sessions.

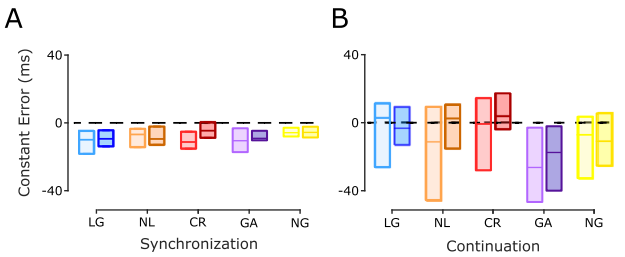

Figure 5. SCT constant error. ***A***, Synchronization epoch. ***B***, Continuation epoch. Interquartile box plots of the constant error calculated for the five groups during the pre- and post-training sessions (color code as in Fig. 3*C*). The horizontal black dotted line is a reference at zero constant error. No significant session × group interactions were found in the two epochs.

mean angle test was significantly different from zero for all groups and sessions ($p < 0.05$)] except for the LG group in the pre-training session, reflecting strong predictive behavior in all subjects in both SCT sessions. Finally, a Harrison–Kanji test (two-way ANOVA for circular data) on the asynchronies showed no significant main effect for the group ($F_{4, 88} = 0.8536$, $p = 0.4931$) or session ($F_{1, 88} = 1.2374$, $p = 0.2675$). These findings suggest that the predictive mechanisms behind consistent and negative mean asynchronies are not influenced by the generalization of a time discrimination task or exclusively being exposed to an initial SCT session.

Next, we compared the constant error, a measure of timing accuracy. A three-way ANOVA, with constant error as the dependent variable, did not show a significant main effect for group × session × epoch ($F_{4, 88} = 0.1223$, $p = 0.8850$). These results support the notion of an accurate estimation of the interval during the synchronization and continuation epochs across groups and sessions, with no performance differences between them (Fig. 5*A*,*B*). Hence, the interval discrimination learning did not generalize to changes in tapping accuracy across groups of subjects or SCT epochs.

The temporal variability showed that interval discrimination training differentially modifies the tapping precision of the subjects during the SCT (Fig. 3*D*,*E*). A three-way ANOVA (epoch, session, and group) with temporal variability as the dependent variable showed a significant group × session × epoch interaction ($F_{4, 88} = 2.9789$, $p = 0.0233$) and a group × session interaction ($F_{4, 88} = 17.541$, $p < 0.0001$). Time variability changes were larger but not limited to the internal guided epoch of the task (continuation), as previously reported (Meegan et al., 2000; Planetta and Servos, 2008). Indeed, post hoc paired t tests showed during the

synchronization epoch a significant reduction of temporal variability between the pre- and post-training sessions for the LG ($t_{31} = 2.9929$, $p = 0.0054$) and CR ($t_{16} = 2.2858$, $p = 0.0362$) groups, while the NL group ($t_{15} = -0.9346$, $p = 0.3648$) exhibited no statistically significant changes between sessions (Fig. 6*C*). Importantly, paired *t* tests for the continuation epoch showed that the LG ($t_{31} = 7.9742$, $p < 0.0001$) and CR ($t_{16} = 4.6920$, $p = 0.0002$) groups presented a significant reduction in their time variability between sessions (Fig. 6*D*). In contrast, the NL group ($t_{15} = -5.3636$, $p < 0.0001$) displayed a statistically significant increase in the variance of their performance.

The initial temporal variability showed significant group differences during the SCT continuation epoch ($F_{4, 88} = 2.53$, $p = 0.0457$). *Post hoc* paired *t* tests showed lower inter-tap variability for NL subjects compared to LG ($t_{46} = -2.3103$, $p = 0.0254$), CR ($t_{31} = -2.4004$, $p = 0.0226$), and GA ($t_{27} = -2.3554$, $p = 0.0260$) subjects, but no significant differences with NG subjects ($t_{30} = -1.0381$, $p = 0.3075$). These results suggest that the NL group had a preexistent tapping skill with low initial temporal variability.

We did not find significant correlations between the level of musical skill and the clustering of the three main groups. Every subject responded to a questionnaire regarding their musical training, sports practice, and videogame experience. Our sample had basic elementary school musical training but did not include professional musicians. Mean hours of weekly practice and years of training are included in Table 1. Although there is a bias for higher scores in the NL group, no statistically significant effects of group were found for the tested skills (Kruskal–Wallis tests using musical training, sports practice, and videogame experience as dependent variables and groups as factors).

Overall, these results support the hypothesis that intense IDT training is an important mechanism to effectively reduce time variability in the motor tapping task, especially during the continuation epoch, probably due to timing generalization. In addition, there were clear individual differences, with not only a large group of subjects learning during the IDT and generalizing to the SCT but also a group of non-learning IDT subjects with no changes in the SCT and a group that showed no learning during the IDT but did show procedural changes during the two sessions of the SCT (CR). Finally, the subjects in the control task without IDT training were also clustered into individuals with and without practice effects in the SCT continuation epoch. Next, we measured the changes in the BOLD signal between sessions across the experimental and control groups.

fMRI data

Our first approach to the functional imaging data was to determine the brain areas involved in the SCT before IDT training using whole-brain analysis. All subjects (experimental and controls, $n = 93$) were grouped together, and the mean activation was calculated for both the synchronization and continuation

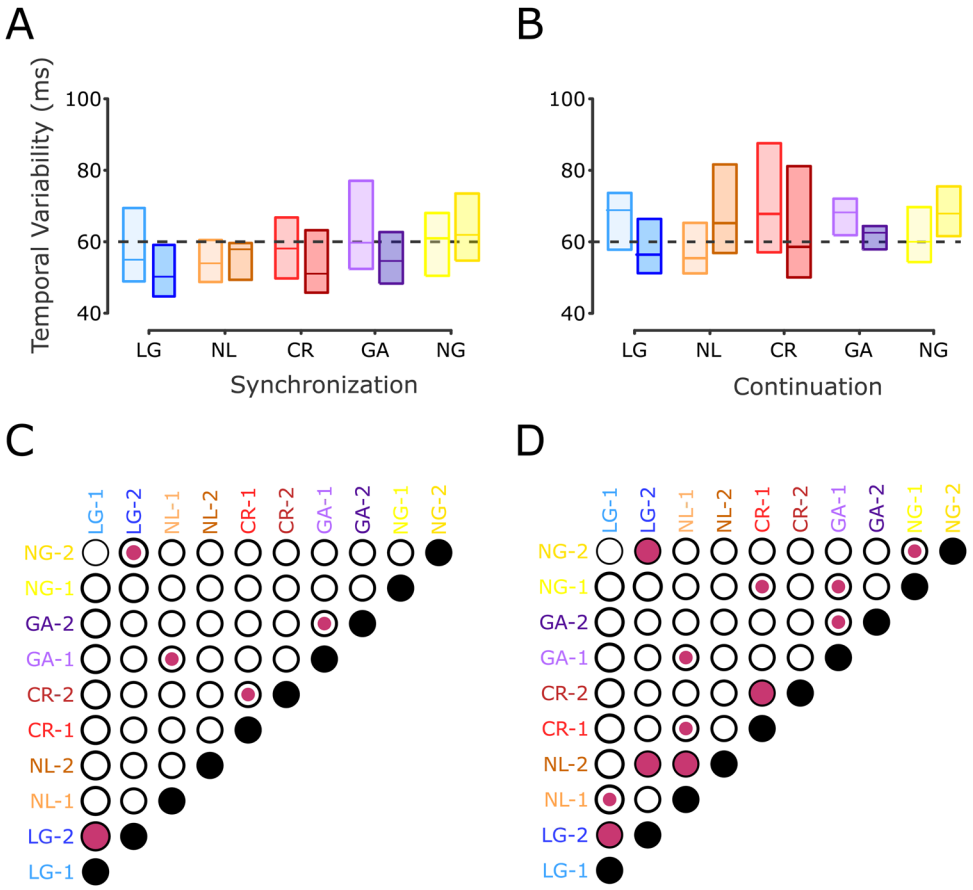

Figure 6. SCT temporal variability. ***A***, Interquartile box plots of the temporal variability during the synchronization epoch for the five groups during the pre- and post-training sessions (color code as in Fig. 3*C*). The horizontal black dotted line is a visual reference at 60 ms. ***B***, Temporal variability as in ***A*** for the continuation epoch. ***C***,***D***, Matrix of statistical differences between groups for ***A*** and ***B***, respectively. Each element of the matrix corresponds to the pairwise comparison (*t* test) between the five groups for the pre-training (1) and post-training (2) sessions in the SCT. Notation as in Figure 3*D*. Note the large decrease in temporal variability in the post-training for the LG, CR, and GA groups, especially during the continuation epoch

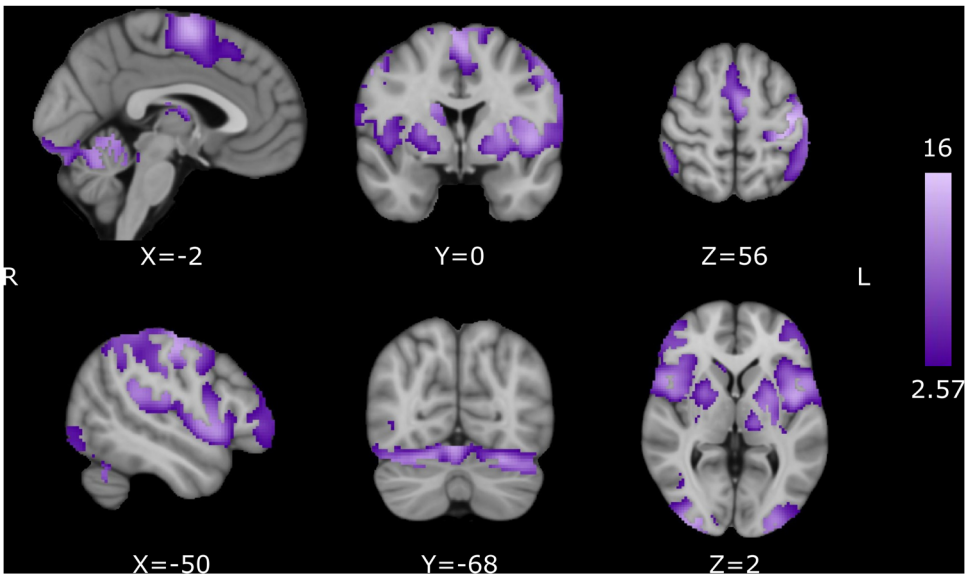

Figure 7. Pre-training activation for the SCT. Common activation for the synchronization and continuation epochs. Activation maps are displayed as *Z* scores with a threshold of $p < 0.005$, cluster-level FWE $p < 0.05$, overlaid on the MNI template.

Table 1. Musical, sports, and video game experience for LG, NL, and CR groups

Group	% of subjects with musical training	Years of practice	Hours of weekly training	% of subjects practicing sports	Years of training	Hours of weekly training	% of subjects with videogame experience	Hours of weekly practice
LG	41.9	1	2.98	90.3	4.54	7.12	48.38	1.9
NL	66.7	2.16	6.43	93.3	6.47	7.83	66.66	11.63
CR	47.1	1.83	4.97	100	7.08	8	58.82	8.17

Percentage of subjects within each group with musical, sports, and video game experience, years of practice, and hours of weekly practice.

epochs. The areas that showed a statistically larger activation during the SCT with respect to the rest of the conditions included the bilateral SMA, bilateral pre-SMA, left M1, left S1, bilateral dorsal prefrontal cortex, bilateral ventromedial prefontal cortex, left *planum temporale*, bilateral putamen, bilateral BA44 (Broca's area), bilateral insula, bilateral visual cortices, bilateral cerebellar lobules I–VI, and the left Crus I (Fig. 7, Table 2). These results indicate that the execution of motor timing tasks relies on a cortico-basal ganglia circuit, as well as a cerebellar circuit, which are key elements of the core timing system.

Subsequently, a third-level mixed model analysis was performed between the pre- and post-training sessions for each group. This model included separate subject-intercepts across sessions, improving statistical sensitivity. Notably, the NL group showed an increase in activity during the post- versus pre-training session in the bilateral pre-SMA and left caudate–putamen during the continuation epoch (Fig. 8*A*, Table 3). In addition, the CR group showed a post-training > pre-training cluster in the left caudate–putamen that might mediate CR motor learning, while the NL group exhibited a pre-training > post-training cluster in the right cerebellar cortex VI that could internalize rhythmic timing experience in this group before the experiment. No changes in hemodynamic signal were found between sessions during the synchronization epoch across groups or in the continuation of the controls. These results suggest that the successful transfer of learning from interval discrimination to more precise, internally driven tapping in the LG group depends on a greater engagement of the pre-SMA and caudate–putamen.

To further scrutinize the role of these areas in the mechanisms of learning and generalization, motor practice, and the preexistent activation patterns before IDT training, we carried out a decoding procedure in which the resulting ROIs from a whole-brain searchlight were used to train an SVM

Table 2. Activation locations for the SCT

Volume (mm^3)	Region	*X*	*Y*	*Z*	*Z* score
6,63,680	Left M1	−40	−20	54	15.2416
	Left SMA	−4	−4	62	13.6882
	Left BA 44	−46	2	4	12.9426
	Right pre-SMA	6	−4	70	12.9316
	Left putamen	24	2	6	9.5294
	Left substantia nigra	−8	−20	−12	3.3573
436,864	Right cerebellum V	12	−52	−20	15.4306
	Right cerebellum VI	22	−56	−24	15.4025
	Right cerebellum vermis	6	−54	−10	14.5481
	Left cerebellum VI	−26	−64	−22	13.1883
	Right V3	30	−92	2	12.3649
	Left V3	−28	−98	−6	11.2594
308,480	Right BA 44	52	14	4	12.0211
	Right dPMC	56	−2	46	11.4111
	Right IFG	42	42	−4	8.6029
101,184	Right *planum temporale*	60	−32	22	10.1519
	Right IPL	54	−42	−56	7.8393
77,376	Left IFG	−44	52	14	7.4687
36,736	Left thalamus	−16	−16	16	9.2133
	Right thalamus	14	−6	12	7.8774
32,128	Right putamen	24	−2	10	9.6642
	Right GP	18	0	−2	7.4482

Co-occurrent activity during both the synchronization and the continuation epochs during the pre-training session (threshold at $p < 0.05$, cluster-level FWE at $p < 0.005$). Spatial coordinates are in mm according to MNI 152 space. V3, visual cortex 3; IPL, inferior parietal lobule; IFG, inferior frontal gyrus; M1, primary motor cortex; dPMC, dorsal prefrontal cortex; SMA, supplementary motor area; BA, Brodmann area.

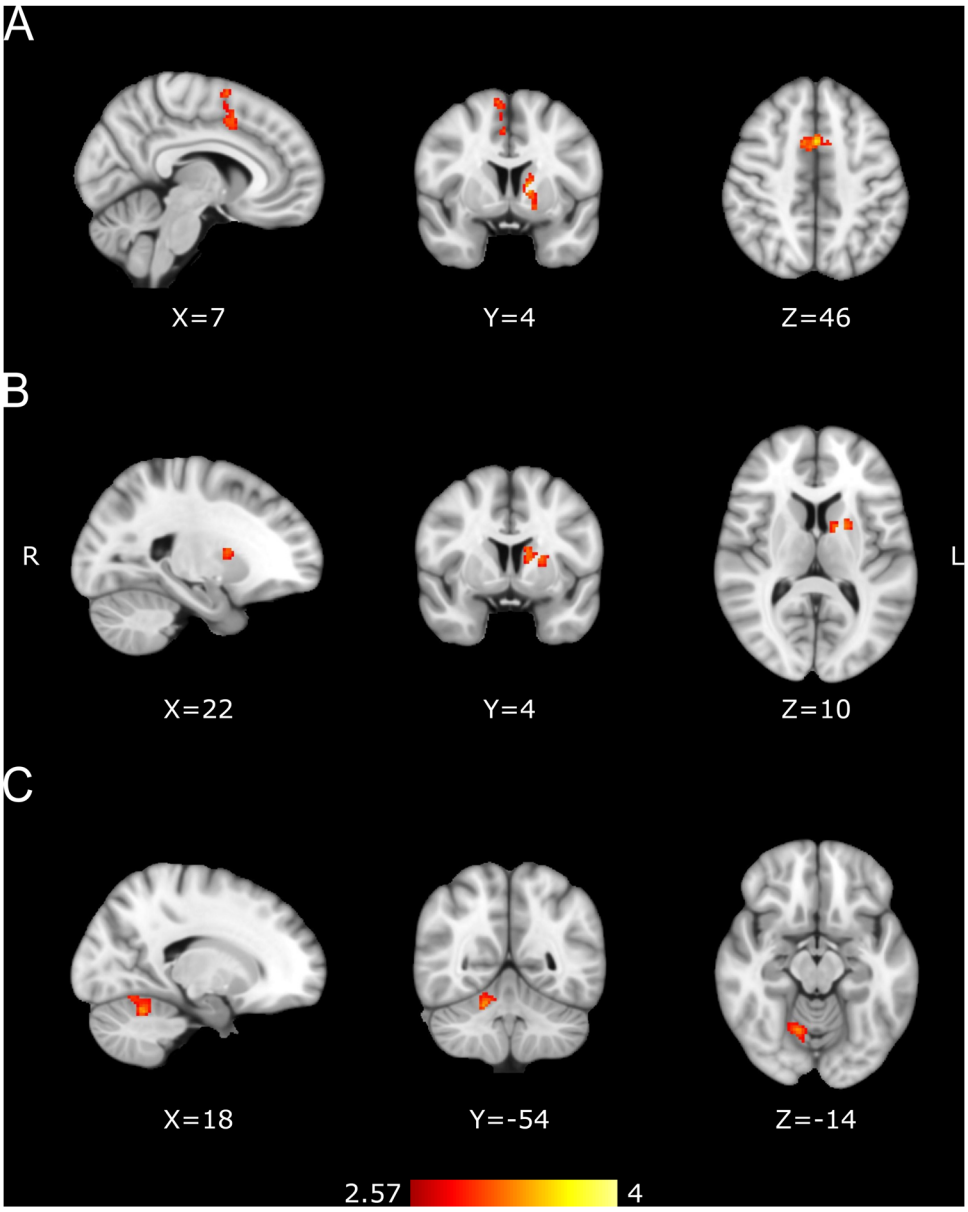

Figure 8. Functional changes between pre- and post-training identified with mixed model analyses. ***A***, LG group post-training > pre-training. Contrast map for 32 subjects. ***B***, CR group post-training > pre-training. Contrast map for 17 subjects. ***C***, NL group pre-training > post-training. Contrast map for 16 subjects. BOLD changes seen during the continuation epoch. Activation maps are displayed as *Z* scores with a threshold of $p < 0.005$, cluster-level FWE $p < 0.05$, overlaid on the MNI template.

Table 3. Activation locations for the pre-training versus post-training contrast

Groups	Volume (mm^3)	Region	*X*	*Y*	*Z*	*Z* score
LG post-training > pre-training	14,720	Pre-SMA-L	1	9	45	3.87
		Pre-SMA-R	9	5	65	3.48
	9,344	Left caudate	−11	5	5	3.6
		Left GP	−13	3	1	3.16
CR post-training > pre-training	6,144	Left putamen	−19	3	9	3.24
		Left caudate	−11	3	15	3.23
NL pre-training > post-training	9,280	Right cerebellum VI	19	−55	−25	3.47

Post-training > pre-training for the LG and CR groups and pre- > post-training for NL group during the continuation epoch (threshold at $p < 0.05$, cluster-level FWE at $p < 0.005$). Coordinates are in MNI space, expressed in mm. SMA, supplementary motor area.

algorithm to classify the three main subject groups: LG, CR, and NL. We identified 15 ROIs with robust classification power that include voxels in the SMA, pre-SMA, putamen, caudate, globus pallidus (GP), and cerebellum (Fig. 9*A*,*C*; see Materials and Methods), in accordance with the above mixed model results. Importantly, the decoding accuracy using all subjects (LG, 32; CR, 17; and NL, 16; Fig. 9*B*, black dot) and using 16 subjects per group (randomly picked, 10,000 iterations; Fig. 9*B*, green dot) was close to 70%, quite above the chance level (theoretical = 33.3%; actual with random permutations; Fig. 9*B*, pink dot). Therefore, these results support the hypothesis that the cortico-basal ganglia circuit and the cerebellum are deeply involved in defining the behavioral properties of the three groups of subjects.

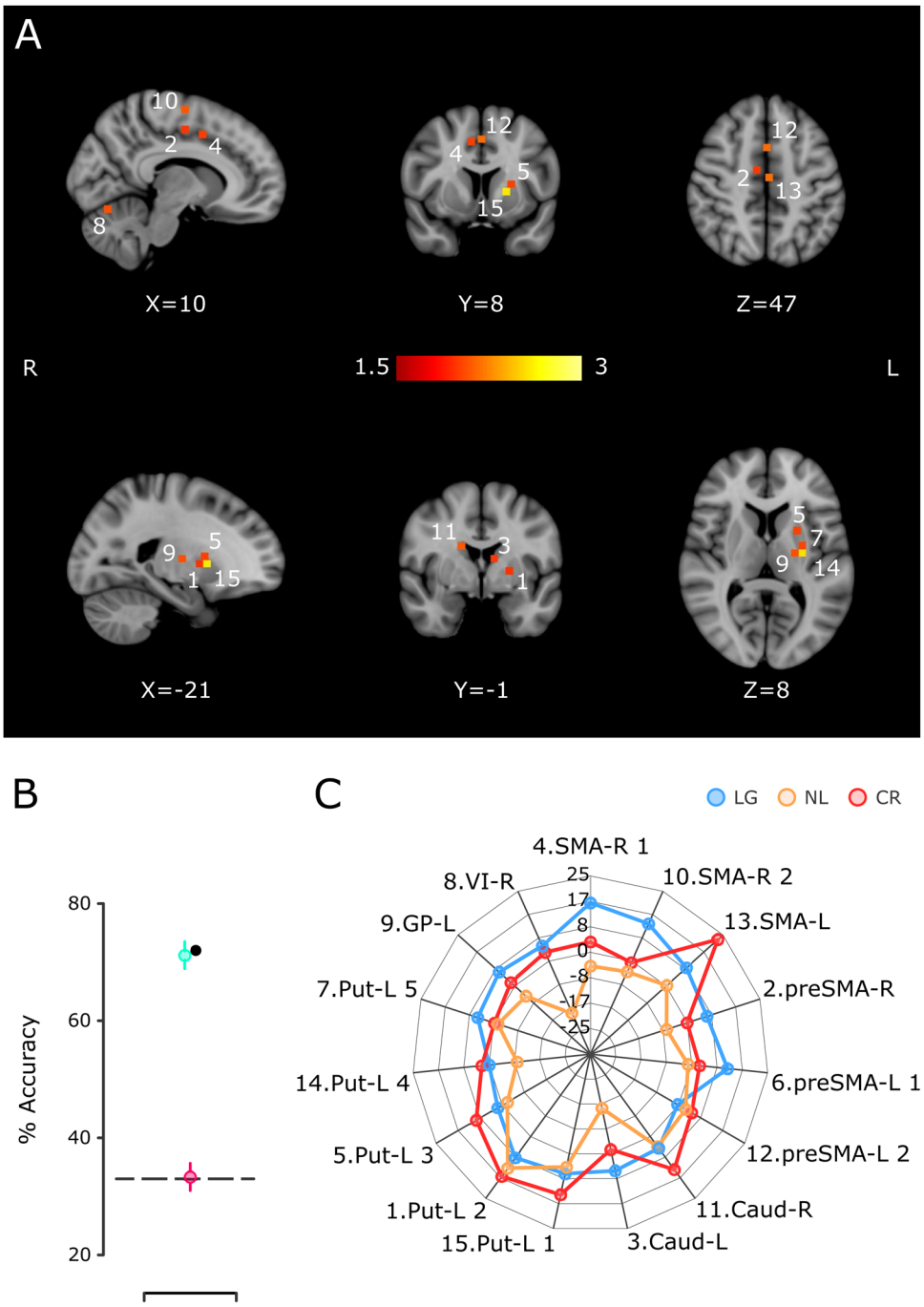

Figure 9. SVM classification of the LG, CR, and NL groups using the delta of activation between the pre- and post-training sessions. ***A***, Location of the 15 ROIs that produced larger accuracy in the classification of the three groups. The color scale corresponds to the log probability that each ROI participated (alongside three other randomly selected ROIs) in a significant classification iteration. ***B***, Percentage of the group's classification accuracy using the fifteen ROIs of ***A*** in the SVM. The black dot corresponds to the accuracy using an SVM with the total number of subjects per group, namely, 32 LG, 16 NL, and 17 CR. The green circle corresponds to the mean (bar -SEM; 10,000 iterations) of accuracy when 16 LG subjects were randomly selected. The pink circle depicts the mean accuracy (bar + SEM; 10,000 iterations) when the classification algorithm was run using random group labels, and the dotted line represents the theoretical chance level. ***C***, The radar plot that shows the mean COPE value for the continuation epoch per group for the ROIs in ***A***. The number at the beginning of the label corresponds to the numeric label in panel ***A***. Caud, caudate; GP, *globus pallidus*; Put, putamen; SMA, supplementary motor area; VI, cerebellum VI lobule.

Discussion

The present research examined changes in neural activity associated with the transfer of learning from perceptual to motor timing and compared them with the hemodynamic response of NL subjects, CR subjects, and a group of control subjects that performed the tapping task but did not undergo intensive time discrimination training. Our study supports four conclusions. First, intense training in an IDT produced an increase in the acuity of time perception in the LG group. Second, there is a strong correspondence between the reduction of the discrimination threshold in the IDT and the reduction of temporal variability of the produced intervals, mainly during the internally driven epoch of the SCT in the LG group. Third, initial interval discrimination performance accounted for the lack of learning in NL and CR

subjects, evidencing a floor effect on time perception. Finally, functional changes occurred in the bilateral pre-SMA and left caudate–putamen when there was a learning transfer from time discrimination to tapping precision during the continuation epoch. Thus, an increase in activity in the cortico-basal ganglia circuit was observed between the post-training and pre-training sessions only in the LG group. In addition, we found a post-training > pre-training increase in the left caudate–putamen in the CR group and a pre-training > post-training increase in the right cerebellar cortex VI of NL subjects. The decoding analysis corroborated the role of these areas in defining the behavioral profile of the three groups of subjects. No BOLD increases were found between sessions for the control groups, and no session changes were found for any of the groups during the synchronization epoch.

Our psychophysical results revealed that a learning process occurred during the 7 consecutive days of intensive training in the IDT. Notably, the learning function of the present study is similar to the time course of learning for auditory, visual, and somatosensory interval discrimination (Kristofferson, 1980; Wright et al., 1997; Nagarajan et al., 1998; Westheimer, 1999; Karmarkar and Buonomano, 2003). All these experiments included intensive daily training for five or more days. With this protocol, learning is characterized by an increase in time perception acuity and occurs mainly during an initial rapid improvement stage that lasts for 2 or 3 d, followed by a slower improvement phase that spans the remaining sessions. Nevertheless, important individual differences are also evident in these studies, with a proportion of participants showing no ability to learn and a decrease in their temporal precision during time perception training. In our case, the LG and NL groups allowed us not only to investigate not only the generalization rules of timing from perception to production but also to contrast the neural circuits involved in learning transfer versus those involved in the SCT practice during the pre- and post-training sessions.

The SCT is a prototypical paradigm that contains an initial tapping synchronization epoch, where subjects are entrained to an isochronous metronome, followed by an internally driven continuation epoch (Wing, 2002; Repp, 2005). Thus, a natural question is whether learning generalization from time perception was present in either or both SCT epochs. Performance in this tapping task can be characterized in terms of precision (temporal variability), accuracy (constant error), and predictability (asynchronies specific to the synchronization epoch; Zarco et al., 2009; Gámez et al., 2018; Yc et al., 2019). Importantly, the performance gain in temporal precision in our visual IDT was transferred as an increase in timing precision mainly during the internally driven period of the SCT. No generalization was observed in the timing accuracy or predictability during both SCT epochs. Furthermore, the increase in timing precision of the continuation epoch due to learning transfer is evident in the LG group but not in the NL group. Therefore, these results revealed a specific mechanism for time generalization in LG subjects: intensive training produced a more robust neural representation of an interval and a concomitant increase in perceptual acuity for this duration. In turn, the improved neural representation of the interval is transferred as an increase in temporal precision when subjects access this neural signal to produce internally driven rhythmic movements.

In the present study, we found a specific increase in activation in the bilateral pre-SMA and left caudate–putamen between the post- and pre-training sessions for the LG group. These results suggest that the cortico-basal ganglia circuit is responsible for the learning transfer from time discrimination to the internally driven epoch of the tapping task. Many functional imaging studies have documented the fundamental role of medial premotor areas (SMA-proper and pre-SMA) and the putamen in timing, showing increments in activity in diverse interval- and beat-based timing tasks (Rao et al., 1997; Bengtsson et al., 2005; Jantzen et al., 2007; Karabanov et al., 2009; Coull et al., 2013; Konoike et al., 2015; Mendoza et al., 2018). Our findings indicate that the pre-SMA and caudate–putamen are critical nodes for the generalization of an acquired timing precision skill. The main question, then, is how this could be achieved. The learning generalization literature concurs with the principle of lack of generalization in the time domain, where the learned gain in temporal precision does not transfer for durations differing for >50% of the trained interval. Indeed, in a previous study, we found that training in an interval reproduction task produced a Gaussian generalization function, with large generalization for closely neighboring untrained intervals and no generalization for intervals distant from the trained duration (Bartolo and Merchant, 2009). Therefore, these observations suggest the existence of neural circuits that are tuned to a specific time length. In fact, interval-tuned cells have been recorded in pre-SMA/SMA (Mita et al., 2009; Merchant et al., 2013b; Crowe et al., 2014; Gámez et al., 2019), putamen (Bartolo et al., 2014), caudate, and cerebellum (Kunimatsu et al., 2018). Furthermore, different functional imaging studies have demonstrated the existence of a topographic representation of time where neuronal units selectively respond to specific durations, generating chronotopic maps on the surface of the human medial premotor and posterior parietal areas (Protopapa et al., 2019; Harvey et al., 2020). Thus, the increase in timing precision of an interval during learning and generalization may depend on an increase in the density of neurons tuned to this interval within the chronotopic map, a decrease in the width of the tuning function of these cells, and/or a concomitant change in the precision of neural population signals across areas of the core timing circuit (Bueti et al., 2008; Merchant et al., 2014b, 2015b; Sohn et al., 2019).

Strong individual differences in the pattern of IDT learning and the changes in tapping precision during the two SCT sessions required us to cluster subjects into three groups. The heterogeneous learning generalization profile suggests that subjects possess different abilities to process temporal information due to genetic or learned factors. We suggest that these factors modulate how the core timing network encodes and predicts timing events and how this timing network dynamically interacts with sensory and cognitive areas to define the precision in timing performance across timing paradigms. The fact that only LG subjects showed activity changes in the cortico-basal ganglia circuit, which is associated with the transfer of learning from sensory to motor timing, supports the notion that the ability to benefit from intervention protocols depends on the individual properties of the core timing network. Hence, the proposed plastic changes in the interval tuning of the medial premotor area and caudate–putamen during our time generalization can be a functional signature of the LG group. Furthermore, the hemodynamic changes in these areas were specific to the continuation epoch of the SCT, where the behavioral gains were greater. On the other hand, we found a cerebellar increase in activity during the pre-training with respect to the post-training in the NL group. Thus, the preexisting tapping ability of NL subjects might be due to previous rhythmic timing skills stored in this area, which is consistent with the role of the cerebellum in timing and motor learning (Tanaka et al., 2021).

Finally, in the CR group, we found a separate set of voxels in the caudate–putamen that showed increased activity during the post-training with respect to the pre-training session. This suggests that the caudate–putamen is engaged in determining the practice effects of timed tapping in two sessions (Lehéricy et al., 2005; Bosnell et al., 2011; Toyomura et al., 2015). Notably, the lack of coordinated activation of the pre-SMA and caudate–putamen in CR and GA subjects, even when they showed a gain in tapping precision in the second SCT session, supports the notion of a specific engagement of the cortico-basal ganglia circuit in the learning transfer from perceptual to motor timing.

References

Ayala YA, Lehmann A, Merchant H (2017) Monkeys share the neurophysiological basis for encoding sound periodicities captured by the frequency-following response with humans. Sci Rep 7:16687.

Balasubramaniam R, Haegens S, Jazayeri M, Merchant H, Sternad D, Song J-H (2021) Neural encoding and representation of time for sensorimotor control and learning. J Neurosci 41:866–872.

Bartolo R, Merchant H (2009) Learning and generalization of time production in humans: rules of transfer across modalities and interval durations. Exp Brain Res 197:91–100.

Bartolo R, Prado L, Merchant H (2014) Information processing in the primate basal ganglia during sensory-guided and internally driven rhythmic tapping. J Neurosci 34:3910–3923.

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

Bengtsson SL, Ehrsson HH, Forssberg H, Ullén F (2005) Effector-independent voluntary timing: behavioural and neuroimaging evidence. Eur J Neurosci 22:3255–3265.

Bi Z, Zhou C (2020) Understanding the computation of time using neural network models. Proc Natl Acad Sci U S A 117:10530–10540.

Bosnell RA, Kincses T, Stagg CJ, Tomassini V, Kischka U, Jbabdi S, Woolrich MW, Andersson J, Matthews PM, Johansen-Berg H (2011) Motor practice promotes increased activity in brain regions structurally disconnected after subcortical stroke. Neurorehabil Neural Repair 25:607–616.

Brainard DH (1997) The psychophysics toolbox. Spat Vis 10:433–436.

Bueti D, Buonomano DV (2014) Temporal perceptual learning. Timing Time Percept 2:261–289.

Bueti D, Walsh V, Frith C, Rees G (2008) Different brain circuits underlie motor and perceptual representations of temporal intervals. J Cogn Neurosci 20:204–214.

Buhusi CV, Meck WH (2005) What makes us tick? Functional and neural mechanisms of interval timing. Nat Rev Neurosci 6:755–765.

Coull JT, Cheng R-K, Meck WH (2011) Neuroanatomical and neurochemical substrates of timing. Neuropsychopharmacology 36:3–25.

Coull JT, Davranche K, Nazarian B, Vidal F (2013) Functional anatomy of timing differs for production versus prediction of time intervals. Neuropsychologia 51:309–319.

Coull JT, Nazarian B, Vidal F (2008) Timing, storage, and comparison of stimulus duration engage discrete anatomical components of a perceptual timing network. J Cogn Neurosci 20:2185–2197.

Cox RW, Chen G, Glen DR, Reynolds RC, Taylor PA (2017) FMRI clustering in AFNI: false-positive rates redux. Brain Connect 7:152–323.

Crowe DA, Zarco W, Bartolo R, Merchant H (2014) Dynamic representation of the temporal and sequential structure of rhythmic movements in the primate medial premotor cortex. J Neurosci 34:11972–11983.

Eklund A, Nichols TE, Knutsson H (2016) Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. Proc Natl Acad Sci U S A 113:7900–7905.

Fabio RPD, Merchant H, Bartolo R, Tuite P (2011) Temporal discrimination learning for treatment of gait dysfunction in Parkinson's disease: a feasibility study using single subject design. Res Rev Parkinsonism 1:8–11.

Gámez J, Mendoza G, Prado L, Betancourt A, Merchant H (2019) The amplitude in periodic neural state trajectories underlies the tempo of rhythmic tapping. PLoS Biol 17:e3000054.

Gámez J, Yc K, Ayala YA, Dotov D, Prado L, Merchant H (2018) Predictive rhythmic tapping to isochronous and tempo changing metronomes in the nonhuman primate. Ann N Y Acad Sci 1423:396–414.

Getty DJ (1975) Discrimination of short temporal intervals: a comparison of two models. Percept Psychophys 18:1–8.

Gibbon J, Malapani C, Dale CL, Gallistel C (1997) Toward a neurobiology of temporal cognition: advances and challenges. Curr Opin Neurobiol 7:170–184.

Harrington D, Castillo G, Fong C, Reed J (2011) Neural underpinnings of distortions in the experience of time across senses. Front Integr Neurosci 5:32.

Harvey BM, Dumoulin SO, Fracasso A, Paul JM (2020) A network of topographic maps in human association cortex hierarchically transforms visual timing-selective responses. Curr Biol 30:1424–1434.e6.

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science (New York, N.Y.) 293:2425–2455.

Herholz SC, Zatorre RJ (2012) Musical training as a framework for brain plasticity: behavior, function, and structure. Neuron 76:486–502.

Honing H, Merchant H (2014) Differences in auditory timing between human and nonhuman primates. Behav Brain Sci 37:557–558; discussion 577–604.

Ivry RB, Hazeltine RE (1995) Perception and production of temporal intervals across a range of durations: evidence for a common timing mechanism. J Exp Psychol Hum Percept Perform 21:3–18.

Jantzen KJ, Oullier O, Marshall M, Steinberg FL, Kelso JAS (2007) A parametric fMRI investigation of context effects in sensorimotor timing and coordination. Neuropsychologia 45:673–684.

Jantzen KJ, Steinberg FL, Kelso JAS (2002) Practice-dependent modulation of neural activity during human sensorimotor coordination: a functional magnetic resonance imaging study. Neurosci Lett 332:205–209.

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 17:825–841.

Karabanov A, Blom O, Forsman L, Ullén F (2009) The dorsal auditory pathway is involved in performance of both visual and auditory rhythms. NeuroImage 44:480–488.

Karmarkar UR, Buonomano DV (2003) Temporal specificity of perceptual learning in an auditory discrimination task. Learn Mem 10:141–147.

Konoike N, Kotozaki Y, Jeong H, Miyazaki A, Sakaki K, Shinada T, Sugiura M, Kawashima R, Nakamura K (2015) Temporal and motor representation of rhythm in fronto-parietal cortical areas: an fMRI study. PLoS One 10:e0130120.

Kristofferson AB (1980) A quantal step function in duration discrimination. Percept Psychophys 27:300–306.

Kunimatsu J, Suzuki TW, Ohmae S, Tanaka M (2018) Different contributions of preparatory activity in the basal ganglia and cerebellum for self-timing. Elife 7:e35676.

Laje R, Agostino PV, Golombek DA (2018) The times of our lives: interaction among different biological periodicities. Front Integr Neurosci 12:10.

Lehéricy S, Benali H, Van de Moortele P-F, Pélégrini-Issac M, Waechter T, Ugurbil K, Doyon J (2005) Distinct basal ganglia territories are engaged in early and advanced motor sequence learning. Proc Natl Acad Sci U S A 102:12566–12571.

Lenc T, Merchant H, Keller PE, Honing H, Varlet M, Nozaradan S (2021) Mapping between sound, brain and behaviour: four-level framework for understanding rhythm processing in humans and non-human primates. Philos Trans R Soc Lond B Biol Sci 376:20200325.

Macar F, Coull J, Vidal F (2006) The supplementary motor area in motor and perceptual time processing: FMRI studies. Cogn Process 7:89–94.

Meegan DV, Aslin RN, Jacobs RA (2000) Motor timing learned without motor training. Nat Neurosci 3:860–862.

Méndez JC, Pérez O, Prado L, Merchant H (2014) Linking perception, cognition, and action: psychophysical observations and neural network modelling. PLoS One 9:e102553.

Mendoza G, Méndez JC, Pérez O, Prado L, Merchant H (2018) Neural basis for categorical boundaries in the primate pre-SMA during relative categorization of time intervals. Nat Commun 9:1098.

Mendoza G, Merchant H (2014) Motor system evolution and the emergence of high cognitive functions. Prog Neurobiol 122:73–93.

Merchant H, Averbeck BB (2017) The computational and neural basis of rhythmic timing in medial premotor cortex. J Neurosci 37:4552–4564.

Merchant H, Bartolo R (2018) Primate beta oscillations and rhythmic behaviors. J Neural Transm 125:461–470.

Merchant H, Bartolo R, Pérez O, Méndez JC, Mendoza G, Gámez J, Yc K, Prado L (2014a. *Neurobiology of interval timing* (Merchant H, de Lafuente V, ed), pp 143–154. Queretaro, Mexico: Springer.

Merchant H, Bartolo R, Pérez O, Méndez JC, Mendoza G, Gámez J, Yc K, Prado L (2014b) Neurophysiology of timing in the hundreds of milliseconds: multiple layers of neuronal clocks in the medial premotor areas. Adv Exp Med Biol 829:143–154.

Merchant H, Grahn J, Trainor L, Rohrmeier M, Fitch WT (2015a) Finding the beat: a neural perspective across humans and non-human primates. Philos Trans R Soc Lond, B, Biol Sci 370:20140093.

Merchant H, Harrington DL, Meck WH (2013a) Neural basis of the perception and estimation of time. Annu Rev Neurosci 36:313–336.

Merchant H, Pérez O (2020) Estimating time with neural networks. Nat Mach Intell 2:492–493.

Merchant H, Pérez O, Bartolo R, Méndez JC, Mendoza G, Gámez J, Yc K, Prado L (2015b) Sensorimotor neural dynamics during isochronous tapping in the medial premotor cortex of the macaque. Eur J Neurosci 41: 586–602.

Merchant H, Pérez O, Zarco W, Gámez J (2013b) Interval tuning in the primate medial premotor cortex as a general timing mechanism. J Neurosci 33:9082–9096.

Merchant H, Yarrow K (2016) How the motor system both encodes and influences our sense of time. Curr Opin Behav Sci 8:22–27.

Merchant H, Zarco W, Bartolo R, Prado L (2008a) The context of temporal processing is represented in the multidimensional relationships between timing tasks. PloS One 3:e3169.

Merchant H, Zarco W, Prado L (2008b) Do we have a common mechanism for measuring time in the hundreds of millisecond range? Evidence from multiple-interval timing tasks. J Neurophysiol 99:939–949.

Mita A, Mushiake H, Shima K, Matsuzaka Y, Tanji J (2009) Interval time coding by neurons in the presupplementary and supplementary motor areas. Nat Neurosci 12:502–507.

Nagarajan SS, Blake DT, Wright BA, Byl N, Merzenich MM (1998) Practice-related improvements in somatosensory interval discrimination are temporally specific but generalize across skin location, hemisphere, and modality. J Neurosci 18:1559–1570.

Pérez O, Merchant H (2018) The synaptic properties of cells define the hallmarks of interval timing in a recurrent neural network. J Neurosci 38:4186–4199.

Planetta PJ, Servos P (2008) Somatosensory temporal discrimination learning generalizes to motor interval production. Brain Res 1233:51–57.

Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE (2012) Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. NeuroImage 59:2142–2154.

Protopapa F, Hayashi MJ, Kulashekhar S, van der Zwaag W, Battistella G, Murray MM, Kanai R, Bueti D (2019) Chronotopic maps in human supplementary motor area. PLoS Biol 17:e3000026.

Rao SM, Harrington DL, Haaland KY, Bobholz JA, Cox RW, Binder JR (1997) Distributed neural systems underlying the timing of movements. J Neurosci 17:5528–5535.

Repp BH (2005) Sensorimotor synchronization: a review of the tapping literature. Psychon Bull Rev 12:969–992.

Sohn H, Narain D, Meirhaeghe N, Jazayeri M (2019) Bayesian computation through cortical latent dynamics. Neuron 103:934–947.e5.

Tanaka M, Kunimatsu J, Suzuki TW, Kameda M, Ohmae S, Uematsu A, Takeya R (2021) Roles of the cerebellum in motor preparation and prediction of timing. Neuroscience 462:220–234.

Toyomura A, Fujii T, Kuriki S (2015) Effect of an 8-week practice of externally triggered speech on basal ganglia activity of stuttering and fluent speakers. NeuroImage 109:458–468.

Treisman M (1963) Temporal discrimination and the indifference interval. Implications for a model of the "internal clock". Psychol Monogr 77: 1–31.

Westheimer G (1999) Discrimination of short time intervals by the human observer. Exp Brain Res 129:121–126.

Wiener M, Turkeltaub P, Coslett HB (2010) The image of time: a voxel-wise meta-analysis. NeuroImage 49:1728–1740.

Wing AM (2002) Voluntary timing and brain function: an information processing approach. Brain Cogn 48:7–30.

Woolrich MW, Behrens TEJ, Beckmann CF, Jenkinson M, Smith SM (2004) Multilevel linear modelling for FMRI group analysis using Bayesian inference. NeuroImage 21:1732–1747.

Woolrich MW, Ripley BD, Brady M, Smith SM (2001) Temporal autocorrelation in univariate linear modeling of FMRI data. NeuroImage 14: 1370–1386.

Wright BA, Buonomano DV, Mahncke HW, Merzenich MM (1997) Learning and generalization of auditory temporal-interval discrimination in humans. J Neurosci 17:3956–3963.

Yc K, Prado L, Merchant H (2019) The scalar property during isochronous tapping is disrupted by a D2-like agonist in the nonhuman primate. J Neurophysiol 121:940–949.

Zarco W, Merchant H, Prado L, Mendez JC (2009) Subsecond timing in primates: comparison of interval production between human subjects and rhesus monkeys. J Neurophysiol 102:3191–3202.