Cerebral Cortex November 2015;25:4213–4225
doi:10.1093/cercor/bhu144
Advance Access publication June 26, 2014

The Human Motor System Supports Sequence-Specific Representations over Multiple Training-Dependent Timescales

Nicholas F. Wymbs[1] and Scott T. Grafton[2]

[1]Department of Physical Medicine and Rehabilitation, Johns Hopkins Medical Institution, Baltimore, MD, USA and [2]Department of Psychological and Brain Sciences, University of California at Santa Barbara (UCSB), Santa Barbara, CA 93106-9660, USA

Address correspondence to Scott T. Grafton. Email: grafton@psych.ucsb.edu

Motor sequence learning is associated with increasing and decreasing motor system activity. Here, we ask whether sequence-specific activity is contingent upon the time interval and absolute amount of training over which the skill is acquired. We hypothesize that within each motor region, the strength of any sequence representation is a non-linear function that can be characterized by 3 timescales. We had subjects train for 6 weeks and measured brain activity with functional magnetic resonance imaging. We used repetition suppression (RS) to isolate sequence-specific representations while controlling for effects related to kinematics and general task familiarity. Following a baseline training session, primary and secondary motor regions demonstrated rapidly increasing RS. With continued training, there was evidence for skill-specific efficiency, characterized by a dramatic decrease in motor system RS. In contrast, after performance had reached a plateau, further training led to a pattern of slowly increasing RS in the contralateral sensorimotor cortex, supplementary motor area, ventral premotor cortex, and anterior cerebellum consistent with skill-specific specialization. Importantly, many motor areas show changes involving more than 1 of these 3 timescales, underscoring the capacity of the motor system to flexibly represent a sequence based on the amount of prior experience.

Keywords: motor learning, repetition suppression, sensorimotor, skill learning, SMA

Introduction

Motor learning is characterized by rapid gains followed by slow incremental improvement in performance (Bryan and Harter 1897; Crossman 1959). This pattern has been taken as evidence in support of a model that suggests sequence learning can be broken down into distinct fast and slow stages (Doyon et al. 2003; Doyon and Benali 2005). While a number of imaging studies present a role for specific and distinct brain regions at putative learning stages (Doyon et al. 2002; Penhune and Doyon 2002, 2005), evidence also suggests that learning-dependent changes might be more complex, with a given region showing a decrease in one stage and an increase in a later stage (Hlustík et al. 2004; Floyer-Lea and Matthews 2005; Xiong et al. 2009).

Upward and downward shifts in motor cortex metabolic activity over several days to weeks of training have generated seemingly conflicting accounts. For instance, one set of observations has shown that practice leads to decreases in activation extent (Xiong et al. 2009) and magnitude (Steele and Penhune 2010) in primary and secondary motor cortices. This pattern of decreasing activation with learning is suggestive of efficiency, such that with longer bouts of training, fewer neural resources are needed to express motor skills. This account is consistent with recent evidence showing that after several months to years of practice, regional metabolic activity is substantially reduced in the motor cortex of non-human primates (Picard et al. 2013). However, in contrast to these findings, there is a different set of observations that report an expansion in the extent (Karni et al. 1995; Hlustík et al. 2004) and magnitude (Floyer-Lea and Matthews 2005) of motor cortex activation. This is thought to belie a neural specialization of skilled representations, or focal expansion of task-specific neural resources, which can develop through years of physical practice (Elbert et al. 1995; Classen et al. 1998; Rosenkranz et al. 2007). Thus, within a given motor area, it is possible that these two processes (efficiency and specialization) could occur simultaneously, but over different timescales.

Previous functional imaging studies have commonly studied learning through the repeated practice of a single sequence. Thus, it is unclear whether changes of brain activity in either magnitude or extent reflect information processing that is specific to the given sequence (i.e., a representational change), or rather, a more general effect that could be due to time spent on a novel and repetitive task. To address this potential ambiguity, experimenters have typically employed random or unlearned sequence controls. However, contrasts between a random or rare control and the learned sequence can introduce additional uncertainties that are related to differences in strategy, awareness, difficulty, and kinematics (Poldrack 2000). Novel methods are therefore needed to identify sequence-specific changes of brain activity. To do this, a recent study used machine learning to classify 4 sequences learned in parallel (Wiestler and Diedrichsen 2013). After a single training session, two key observations were noted. Compared with unlearned sequences, there was a general decrease of BOLD amplitude in motor areas, irrespective of the sequence that was tested. This fits with the observation that practice leads to neural efficiency, which is reflected by decreasing blood flow or metabolism during learning. Critically, the authors also found that across a number of motor areas, the ability of the classifier to distinguish each sequence increased over time. Thus, specialization emerges even if there is an overall decline in activity. Over this short training period, it was therefore possible to observe both a general process of declining levels of brain activity and the appearance of sequence-specific patterns of activity.

In the current study, we extend these observations over a longer time horizon, highlighting the sequence-specific involvement of motor regions across 3 timescales of learning, ultimately to identify regions that express sequence-specific representations over multiple timescales. To do so, subjects learned a set of 6, visually cued, 10-element sequences, analogous to piano arpeggios, using a daily behavioral training

regimen that lasted approximately 6 weeks. We manipulated the intensity of training in order to control for non-specific familiarity effects due to the amount of time spent performing the experiment. Thus, 2 sequences were practiced extensively, 2 occasionally, and 2 rarely throughout the training regimen. This was guided by the basic prediction that the total amount of prior practice, rather than chronologic time, is the primary determinant of the magnitude and location of sequence-specific representations.

To characterize brain activity during learning, functional magnetic resonance imaging (fMRI) was acquired at 4 time points over the course of training. In order to identify sequence-specific brain activity, we used a repetition suppression (RS) design whereby BOLD signal change due to sequence repetition determines the strength to which a given sequence is represented in a given brain region. We characterized 3 learning-related patterns of RS magnitude, operationally defined as initial recruitment, skill-specific efficiency, and skill-specific specialization. These correspond to rapid increasing RS magnitude early in training, followed by slow decreasing or increasing RS magnitude with continued training. These patterns were then used to evaluate the involvement of motor regions over 3 corresponding timescales, such that any given region could show sequence-specific sensitivity for any combination of the 3 learning-related patterns.

The standard RS effect is interpreted as an overall fatigue of a population or sharpening of response due to reduction of non-essential neurons, such that in either case, the larger the difference between new and repeated events, the stronger the representational strength (Grill-Spector et al. 2006). Here, we apply RS to quantify representational strength over several time points during learning. We operationally define decreasing RS with continued training as evidence of "skill-specific" efficiency. With fewer non-essential neurons recruited over time, decreasing RS is reflective of a smaller difference between repeated events, and thus, an increasingly efficient population governing the representation of a particular sequence. We further tested whether there would be brain areas showing a slow increase of RS following extensive practice, suggestive of regional specialization. Note that "skill-specific" specialization is not simply due to a general increase in BOLD. Instead, it reflects a gradual increase in the relative difference between repeated trials, which is assumed to occur because of a stronger representation. We use the term "skill-specific" to distinguish our measurement of changing RS magnitude from that of general increases or decreases of BOLD signal that are observed whenever a task is repeated over time. Thus, our focus is on sequence-specific changes that occur with learning, rather than general changes of BOLD magnitude or extent.

In using RS, any change in activation is the relative difference between successive trials of the same sequence, so that performance between repeated trials is nearly identical. This is particularly advantageous because any change in RS over the course of learning can be disassociated from changes of BOLD magnitude related to kinematics or general task familiarity. This approach provides a novel method whereby subjects can perform at their highest possible performance level throughout training and imaging, ensuring strong test validity between how sequences were produced during training and their expression during scanning. An additional feature of the experiment was that subjects practiced the sequences under 3 different training intensities throughout the training regimen. This provided a means to examine functional contributions as defined by amount of practice independent from the overall amount of time spent on the task.

Motivated by previous imaging studies that relate learning to changes in BOLD activity, we hypothesized that at the onset of learning, many brain regions are needed to represent an unfamiliar motor sequence. With sufficient familiarity for a given sequence, this expansive recruitment is no longer needed. Finally, extensive practice should lead to an enduring representation that can be localized to motor output areas. Reframing these 3 patterns in terms of RS, there should be an initial and rapid expansion of RS throughout sensorimotor-associated cortices. Continued training should result in a pattern of skill-specific efficiency with contraction of RS magnitude occurring on a relatively moderate timescale (hundreds of trials). Further, the emergence of skill-specific specialization, denoted by an increase in RS magnitude, would occur within motor output regions and emerge over a slower timescale (thousands of trials). Finally, we predicted that any given motor region could support a skill-specific representation across more than 1 of these 3 timescales. Therefore, we tested whether skill-specific functional change within a given motor region might span any combination of the 3 proposed RS patterns.

Materials and Methods

Experiment Setup and Procedure

Twenty-two right-handed subjects (13 female, average age 24 years) volunteered with informed consent in accordance with the Institutional Review Board/Human Subjects Committee, University of California, Santa Barbara. All subjects had normal/corrected vision and no history of neurological disease or psychiatric disorders. Two subjects were excluded from the analysis: one did not complete the experiment, and the other had persistent head motion greater than 5 mm during MRI scanning.

Subjects performed a training regimen that involved the simultaneous acquisition of 6 different 10-element sequences using a discrete sequence production (DSP) task (Abrahamse et al. 2013). Subjects trained both at home on their personal laptop computers and inside an MRI scanner during the collection of event-related BOLD. Subjects began the experiment with an initial pre-training session inside the MRI scanner. Afterwards, they completed a minimum of 10 home training sessions (1 session/day) during a 14-day interval and then returned to the MRI scanner for a training scan. This pattern between home and scanner training was repeated 3 times, so that by the end of the training regimen, subjects completed at least 30 home training sessions and 3 training scan sessions.

Subjects practiced the DSP sequences with their right hand using either a laptop keyboard (home training) or a button box (scanner training) (Fig. 1). Each trial started with the presentation of a sequence identity cue, followed 2 s later by the first DSP target, which also served as the "go cue". Sequence targets were displayed with 5 horizontally presented stimuli, and responses to these followed a spatially compatible left to right mapping, such that the thumb corresponded to the leftmost stimulus and the pinky finger corresponded to the rightmost stimulus. A target highlighted in red served as the instruction for which key to press. Once the correct response was made, the next target in the sequence was immediately highlighted in red. Unlike the serial reaction time (SRT) task, there was no inter-stimulus interval between successive target instructions, allowing subjects to generate rapid, arpeggio-like motor sequences. In the event of an error, subjects were given visual feedback ("Incorrect"), and the target remained highlighted until the correct response was made. Subjects had an unlimited amount of time to produce a sequence but were instructed to produce the sequences as quickly and accurately as possible. The presentation

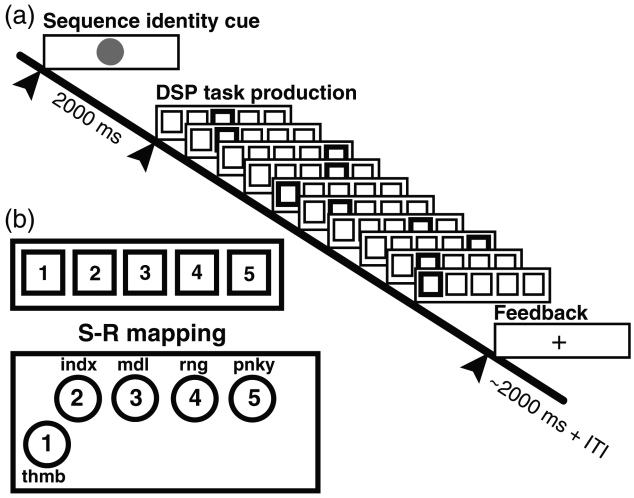

Figure 1. Trial structure and stimulus–response (S-R) mapping. (*a*) A trial began with the visual presentation of a sequence identity cue (2 s), which was followed by the initial DSP stimulus (go cue). A correct key press led to the immediate succession of DSP target stimuli. Subjects received feedback "+" signaling sequence completion and waited (0–6 s) for the next trial. (*b*) Direct S-R mapping between response device, either MRI-compatible button box (shown) or keyboard, and the right hand.

of a fixation cross "+" signified the end of a trial and remained on the screen until the onset of the next sequence identity cue. Sequences were constructed so that each stimulus–response location and corresponding finger was used twice per 10-element sequence. Sequences with repetition (e.g., "2–2"), and regularities such as trills (e.g., "3–2–3") and runs (e.g., "2–3–4") were excluded.

Sequence familiarity was manipulated during home training by presenting the 6 sequences at 3 exposure levels. During each home training session, subjects practiced 150 randomly ordered trials so that 2 sequences were extensively practiced (EXT, 64 trials/sequence), 2 were moderately practiced (MOD, 10 trials/sequence), and 2 were minimally practiced (MIN, 1 trial/sequence). The same sequence pairs were used for EXT, MOD, and MIN trials on all days and for all subjects. There was no counterbalancing of the sequences across the different exposure conditions. The sequences were of comparable difficulty, as reflected by performance during the pre-training scan session done prior to any home training (see Results).

The degree of RS was assessed separately for MIN, MOD, and EXT sequences. This was done by having subjects practice the sequences in groups of 10 trials from the same exposure level (MIN, MOD, or EXT). Within each 10 trial group (5 trials/sequence), sequence presentation order was pseudorandom so that there could be, on successive trials, the repeated presentation of the same sequence, or a non-repeat (we refer to this as a new sequence). At the end of each 10 trial group, subjects received performance feedback that reported the number of error-free sequences as well as the average time needed to complete a correct sequence. Each scan session contained a total of 300 trials, which was divided up into 5 scan runs with each run containing 6 trial groups. Because sequence production was self-paced, the number of scanned TRs varied between subject and session. In order to collect event-related fMRI data, the inter-trial interval ranged between 0 and 6 s (average of ~5 s). The number of sequence trials performed during each scan session was the same for all subjects with the exception of 2 abbreviated sessions due to technical problems. In each of these 2 cases, the scanning protocol was stopped short, so that 4 out of the normally acquired 5 runs were completed. Data from these sessions are included in the presented analysis.

During scanning, subject comfort was improved by placing padding under the knees and right arm. Head motion was minimized by inserting padded wedges between the subject and head coil of the MRI scanner. Subjects made responses using a fiber-optic response box designed with a similar configuration of buttons as those found on the typical laptop used during training (Fig. 1b). The center-to-center spacing between the top row of buttons was 20 mm, and the spacing between the top row and lower left thumb button was 32 mm. The response box was stabilized using a board so that positioning was adjustable with respect to reach and hand size.

Behavioral Apparatus

Stimulus presentation during home training was controlled via each subject's laptop computer using PsychtoolBox Version 3 (http://www.psychtoolbox.org) in conjunction with Octave 3.2.4 (an open source version similar to MATLAB). Keyboard sampling rate was set to 200 Hz. Scan sessions were controlled using a laptop computer running MATLAB version 7.1 (Mathworks). A custom fiber-optic button box and transducer handled key-press responses and response times (RTs) via a serial port (200 Hz polling latency; button box: HHSC-1 × 4-L; transducer: fORP932; Current Designs).

Imaging Procedures

A 3.0 T Siemens Trio with a 12-channel phased-array head coil was used for the acquisition of BOLD. A single-shot echo planar imaging sequence that is sensitive to BOLD contrast was used to acquire 37 slices per repetition time (TR = 2000 ms, 3 mm thickness, and 0.5 mm gap), echo time (TE) of 30 ms, flip angle of 90°, field of view (FOV) of 192 mm, and 64 × 64 acquisition matrix. Prior to the acquisition of BOLD, a high-resolution T1-weighted sagittal sequence image of the whole brain was obtained (TR = 15.0 ms; TE = 4.2 ms; flip angle = 9°, 3D acquisition, FOV = 256 mm; slice thickness = 0.89 mm, acquisition matrix = 256 × 256).

Data Analysis: Behavior

Three behavioral variables of interest were collected during scanning 1) the time elapsed from the presentation of the first DSP target and the first button press, or RT; 2) the time needed to complete the entire sequence of 10 button presses starting with the first key press, or movement time (MT); 3) error trials that included any incorrect response. For each scan session, RT and MT were grouped, preserving temporal order, into 5 bins of 10 trials. For each of these variables, effects of learning were tested by using a three-way repeated-measures ANOVA (session, trial bin, and exposure condition), so that subject was treated as a random factor. Error was grouped according to exposure condition and session and then evaluated using a two-way repeated-measures ANOVA. Because RS was used to measure systematic changes in BOLD amplitude through sequence repetition, it was also important to test whether there were any performance effects due to behavioral priming. A three-way repeated-measures ANOVA was used to measure effects due to priming, with the factors being session, exposure condition, and repetition (either new or repeated trial).

The training regimen required subjects to learn the sequences simultaneously at 3 different exposure levels (MIN, MOD, and EXT). As result, 30 training days were needed to complete approximately 200 MIN sequence trials, but only 3 days were needed to complete the same amount of EXT sequence trials. These differences in training intensity enabled the dissociation of effects related to the depth of training, measured as the actual number of practice trials performed for a given sequence, from effects related to the overall amount of time spent on task. To test whether learning is determined by training depth, performance during scanning was compared between the different exposure conditions that were matched in terms of physical practice. This allowed for direct comparisons between MIN and MOD after 140, 150, and 190 trials during scanning. For a more expanded scope, we extended this analysis to include all home and scanner training data points (see Supplementary Material).

The expression of a motor skill was characterized by 1) the appearance of predictive, rather than cued, movements and (2) the emergence of a relative plateau in MT-based performance. In order to conclude that a given finger movement is predictive, the duration between successive key presses, or inter-key interval (IKI), should be faster than a reactive movement to an unknown DSP stimulus. In this respect, predictive movements are generated using the memory of practiced motor patterns rather than being directed by the sequence stimuli. This

method is similar to generative tests for sequence knowledge commonly employed in studies using a SRT task but is different in that representation strength is quantified over the course of learning rather than as a test after training is complete. We adopted the criterion for predictive responses as follows. Using only correct sequence trials, we chose the 75th percentile of the IKIs (220 ± 65 ms) within the MIN trials from the last training scan as a threshold for labeling an IKI as predictive. This is a conservative estimate, because at this point in training, subjects had already practiced each MIN sequence for approximately 150 trials. Of note, a similar proportion of the above threshold IKIs were obtained across a range of other percentiles used as the cutoff (98–50%). The number of IKIs per trial that were faster than the threshold was then compared across exposure condition and session using a two-way repeated-measures ANOVA.

In order to estimate a time in training when a subject's performance reached a virtual plateau, we fit each subject's MT data to a double-exponential model using all correct trials collected both at home and during scanner training sessions from the EXT sequences. Using the model estimate, the MT plateau was defined as the point on the slope where the estimated MT changed less than 0.25 ms over a sliding 5-trial window. The midpoint of this window was then labeled as the time when a subject reached a relative plateau in MT performance. A similar approach was used to test whether MOD and MIN sequences had reached performance plateau.

Data Analysis: fMRI

Functional imaging data were processed and analyzed using Statistical Parametric Mapping (SPM8, Wellcome Department of Cognitive Neurology). Raw functional data were realigned, co-registered to the native T1 (using the first mean image as the base image for all functional scans), normalized to the MNI-152 template with a re-sliced resolution of 3 × 3 × 3 mm, and then smoothed with a kernel of 8 mm full-width at half-maximum.

For each subject, the BOLD response was modeled using a single design matrix with parameters estimated using the general linear model (GLM). An event-related design was used to model the expression of sequence-specific representations, with trial onset corresponding to the presentation of the sequence identity cue, 2 s prior to the presentation of the initial DSP target stimulus. It should be noted that this approach includes both the preparation and production of learned sequences. The design matrix for each subject was constructed using separate factors for each scan session (pre-training, training sessions 1–3), exposure condition (MIN, MOD, and EXT), and repetition (new or repeated trial). In order for a trial to be coded as a repeated event, the previous trial must have been the exact same sequence, and the previous trial must have been performed correctly. Those repeated trials that followed error trials, as well as the error trials themselves, were modeled using a separate column in the design matrix. In order to account for non-specific effects of session, blocking variables were included for each scan run.

Potential differences in BOLD due to MT-related kinematics were accounted for by using the MT from each trial as the trial duration for modeled events. This approach has been shown to produce an accurate estimation of the BOLD response in relationship to task duration using the GLM (Grinband et al. 2008). In a previous unreported analysis, this approach was combined with an additional column in the design matrix that also included the trial-wise MT as a covariate of non-interest. Because no appreciable differences were found between the 2 models, we report results pertaining to the former model that corrects for MT in terms of stimulus duration. Note that as MTs shorten with training, there is a greater time interval without movement between trials in later scan sessions. This in turn could interact with the estimates of RS, which compare new and repeated events. In order to control for the potential influence of the time elapsed between trials, each event was weighted by the time elapsed from the previous trial. Following center mean normalization, this column was added to the model as a covariate of non-interest. Events were convolved using the canonical hemodynamic response function and temporal derivative in SPM8. Using freely available software (Steffener et al. 2010), corresponding beta image pairs for each event type (HRF and temporal derivative) were then combined at the voxel level to form a magnitude image (Calhoun et al. 2004):

$$H = \text{sign}\,(\hat{B}_1) \times \sqrt{\hat{B}_1^2 + \hat{B}_2^2},$$

where H is the combined amplitude of the estimation of BOLD $(\hat{B}_1)$ and temporal derivative $(\hat{B}_2)$. RS magnitude images were calculated by subtracting the image for repeated trials from the image for new trials. This was done separately for each exposure condition (MIN, MOD, and EXT) and scan session. The resulting RS magnitude images were then entered into the group analysis.

Higher-level mixed-effects group analysis was performed using the full-factorial design option in SPM. A single factor (12 levels: 1 for each session and exposure condition) was used to model skill-specific longitudinal effects. Model factor levels were entered according to the cumulative amount of training trials performed rather than chronological time: pre-training (MIN/MOD/EXT), MIN during training scans 1–3, MOD during training scans 1–3, and EXT during training scans 1–3.

The main objective of the experiment was to test for the presence of temporal dynamics in the BOLD RS that might change at multiple timescales across extensive practice. To do this, contrasts for 1 main effect and 3 main temporal interactions were evaluated at the group level.

1. The main effect of RS, collapsed across all sequences, scanning sessions, and types of training intensity, was calculated using a *t*-test and corrected for multiple comparisons using family-wise error (FWE) correction ($P < 0.05$). A less conservative version of this contrast ($P < 0.001$, uncorrected) was used as a mask to constrain the search volume for the remaining contrasts.
2. To investigate the effects related to the initial formation of sequence-specific knowledge, we applied a linear subtraction of RS effects for trials performed during the pre-training scan (i.e., when the sequences are unknown or weakly known) from MRI training scan 1, when the MIN sequences were weakly known. We predicted that a broad set of motor areas would show increasing RS due to the initial formation of MIN sequence representations.
3. Efficiency models predict that motor areas will show a progressive reduction in BOLD activation with the continuation of training. This general compression of activation will lead to a decrease in RS magnitude. In order to test for potential changes of RS magnitude suggestive of skill-specific efficiency, a linear model was used to measure longitudinal decreases in RS magnitude as a function of the total number of practice trials performed. RS magnitudes for each type of sequence from training scan sessions 1–3 were weighted by the amount of prior training experience.
4. Specialization models predict that with extensive training, task-sensitive motor regions, some of which that might show efficiency effects earlier in training, will show a pattern of slowly increasing RS. In order to test for the expansion of sequence-specific representations, herein referred to as skill-specific specialization, a quadratic interaction was applied over all MIN, MOD, and EXT sequences from training scan sessions 1–3, ordered by prior exposure.
5. Evidence for more than 1 timescale within any given brain region was evaluated using a standard conjunction approach on the second-level group contrasts specified earlier (Contrasts 2–4). Using the logical "and" approach as specified by Nichols et al. (2005), any voxels identified through conjunction were required to be individually significant. Commonalities were assessed using 3 separate tests 1) between initial learning, efficiency, and specialization; 2) between initial learning and efficiency; 3) between efficiency and specialization.

Contrasts 2–4 were tested using an initial liberal threshold of $P < 0.005$, with a minimum cluster size of 10 voxels, and restricted using the mask image corresponding to the main RS effect of task obtained in Contrast 1. These were then corrected for false-positives using topological false discovery rate (FDR) correction. Only clusters surviving correction for multiple comparisons are reported in table form. Clusters larger than 100 voxels were further inspected for additional sub-maxima, and those surviving FDR correction are reported in table form. In order to reduce the possibility of redundant sub-maxima, we

selected only those peaks that did not share any voxels when applied with a spherical region of interest (ROI, 6 mm radius) centered on each peak. Significance was similarly measured for the conjunction analysis except that clusters containing more than 20 voxels were inspected for sub-maxima.

To generate unbiased effect size plots from the different contrasts, the linear and quadratic temporal interactions were repeatedly generated using a leave-one-subject-out approach. That is, the estimation of the higher-level group mixed-effects model was done using all but 1 subject (i.e., 19 instead of 20 subjects), and from this, the identified local maxima were used to extract mean beta weights from the remaining subject. Mean beta weights were extracted using a spherical ROI (6 mm radius) centered on each local maxima coordinate. This procedure iterated over each of the 20 subjects, so that the displayed amplitudes for each interaction correspond to the overall mean and SEM of the left out subjects' beta weights.

Results

Behavior Effects: Baseline Performance during Pre-training

In order to determine whether the sequences were equally challenging at the start of training, performance during the pre-training session was evaluated using repeated-measure ANOVA. There was no significant effect of condition (MIN, MOD, and EXT) for MT; however, main effects of condition were found for both RT [$F_{2,19} = 4.06$, $P < 0.05$] and error [$F_{2,19} = 6.9$, $P < 0.005$]. The effect of RT was driven by small, but significantly slower RT for the MIN sequences with respect to either MOD or EXT (MIN vs. MOD: $t = 2.44$, $P < 0.05$; MIN vs. EXT: $t = 2.22$, $P < 0.05$). In addition, the accuracy for the EXT sequences was lower than MIN or MOD during pre-training (MIN vs. EXT: $t = 3.13$, $P < 0.01$; MOD vs. EXT: $t = 2.41$, $P < 0.05$). Despite low accuracy during pre-training, this pattern of low EXT accuracy was not carried forward during training (see Results).

Behavior Effects: Initial Learning: Pre-training and Training Scan 1 MIN

To determine how performance improved during initial learning, paired *t*-test comparisons were used to evaluate differences in MT, RT, and error between the pre-training scan session and the MIN trials from training scan 1. As predicted, subjects initiated (RT: $t = 4.26$, $P < 0.0001$) and completed (MT: $t = 10.71$, $P < 0.00001$) the sequences faster after a brief period of training. However, their improvement in speed was accompanied by a decrement in accuracy (error: $t = 5.30$, $P < 0.0001$), altogether being consistent with a speed–accuracy tradeoff.

Behavior Effects: Training Scans 1–3

The impact that practice exposure (MIN, MOD, and EXT) has on MT performance during the scan test sessions was evaluated using a three-way repeated-measures ANOVA (practice exposure × session × time bin). MT shortened dramatically as a function of practice exposure [$F_{2,38} = 149.75$, $P < 0.00001$], session [$F_{2,38} = 65.56$, $P < 0.00001$] and trial bin [$F_{4,76} = 38.84$, $P < 0.00001$], indicating that performance improved as a function of number of practice trials (Fig. 2a). There was a significant interaction between practice exposure and session [$F_{4,76} = 13.45$, $P < 0.00001$]. Post hoc testing (test session 3–test session 1) revealed that this interaction in MT performance across sessions was determined by MIN and MOD improvement relative to EXT (MIN vs. EXT: $t = 4.69$, $P < 0.001$; MOD vs. EXT: $t = 6.32$, $P < 0.00001$). This effect was driven by decreasing performance gains, which occurred as a function of the overall amount of practice. Because the EXT sequences were trained at a higher intensity during home training, performance improvements from 1 scan session to the next were expected to be relatively small compared with between-session improvements for MOD and MIN. Within-session performance also improved in relation to training depth as revealed by an interaction between exposure condition and trial bin [exposure × trial bin: [$F_{8,152} = 15.69$, $P < 0.00001$]. This effect was driven by larger within-session gains for MIN relative to MOD and EXT sequences (MIN vs. MOD: $t = 7.58$, $P < 0.00001$; MIN vs. EXT: $t = 3.78$, $P < 0.001$) and suggests that MIN performance is affected by the amount of practice rather than the amount of time spent on the task in general.

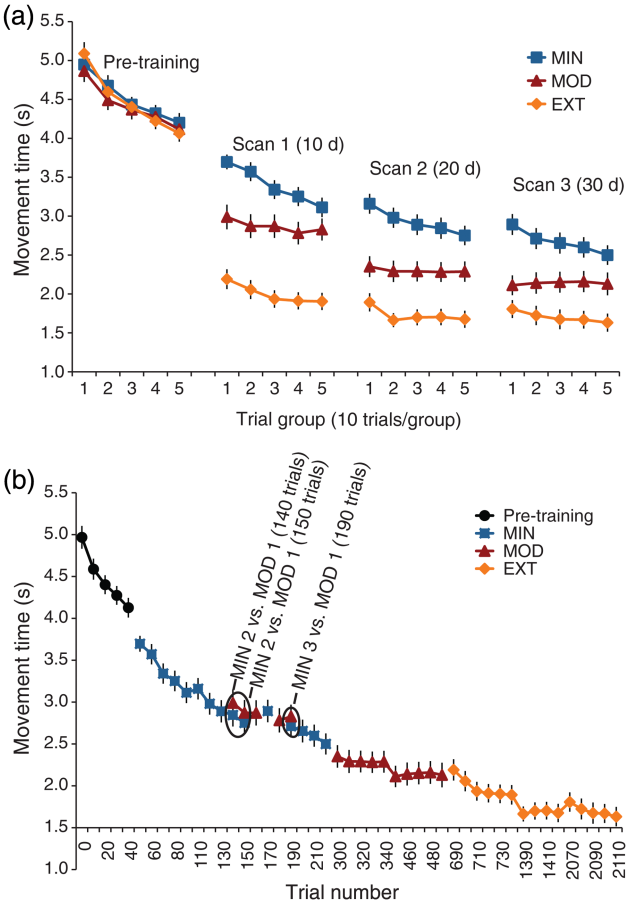

Figure 2. Performance effects during learning. (*a*) Chronologically ordered MTs from the scanning sessions reveal widely different amounts of improvement as a function of prior practice exposure. (*b*) MTs ordered to show the cumulative effect of practice. Note the similar performance for paired MOD and MIN sequences at 140, 150, and 190 trials of exposure, even though these comparisons span 2–3 weeks of chronologic time.

Although MTs became faster over the course of learning, it is notable that performance was sometimes slower at the start of a new scan session relative to the end of the previous scan session (Fig. 2a). For instance, MTs for MIN at the start of scan 3 were slower than those at the end of scan 2 ($t = -2.85$; $P = 0.01$). However, no other direct comparisons were significant (e.g., EXT start scan 3 vs. EXT end scan 2; $t = -1.70$, $P = 0.1$). It is conceivable that this slight fluctuation in performance could be due to differences between learning apparatus and

environment (home/scanner training). It is perhaps more likely that this effect is driven by the relatively small amount of training for the MIN sequences (1 trial/session) during the home training sessions. Although we did not set out to test memory decay, it is possible that such a limited exposure could lead to a drop in performance. Altogether, these small effects are unlikely to have any influence on RS magnitude effects of learning.

RT was evaluated as a secondary performance measure using a similar three-way repeated-measures ANOVA with practice exposure, session, and trial bin serving as the factors. We found effects of exposure [$F_{2,38} = 55.18$, $P < 0.00001$] and session [$F_{2,38} = 11.68$, $P < 0.0001$], indicating that subjects initiated sequences faster with increased practice exposure. An interaction between practice exposure and session further supported this perspective, with rate of improvement in RT occurring more strongly over sessions for MIN with respect to EXT (MIN vs. EXT: $t = 2.62$; $P = 0.0167$). An interaction between practice exposure and trial bin [$F_{8,152} = 3.91$, $P < 0.001$] revealed that within-session RT change differed between exposure conditions.

Error rates were compared over the course of training using a two-way repeated-measures ANOVA with practice exposure and session serving as factors. A significant main effect of practice exposure confirmed that more errors were made for MIN than either MOD or EXT sequences [$F_{2,38} = 9.54$, $P < 0.001$]. An interaction [$F_{4,76} = 2.78$, $P < 0.05$] showed that subjects gradually improved in error for MIN over time but showed somewhat of the opposite pattern for MOD and EXT, indicative of a possible change in speed–accuracy tradeoff over long-term practice.

The above-mentioned findings suggest that sequence performance might be determined entirely by the amount of prior physical practice, with no additional gains from having practiced other sequences. If true, then there should be no difference in performance between different sequence exposure conditions (e.g., MIN vs. MOD) if the amount of physical practice between conditions is matched. To test this idea, comparisons were made between MIN and MOD sequence MT data acquired during scanning. These data were matched in terms of exposure but differed substantially in the overall amount of time spent practicing the task. For instance, 190 MIN sequence trials were completed during training scan 3 whereas 190 MOD trials were completed during training scan 1. We performed paired *t*-test comparisons of MT after the completion of 3 different amounts of training (140, 150, and 190 trials). We found no difference between any of these pairs that were matched in terms of training amount (140 trials: $t = 1.43$, $P = 0.17$; 150 trials: $t = 1.43$, $P = 0.17$; 190 trials: $t = 1.24$, $P = 0.23$), indicating that performance gains are dominated by the amount of physical practice (Fig. 2b). The same conclusion was met with comparisons using either RT (140 trials, $t = -0.0124$, $P = 0.99$; 150 trials: $t = 0.7174$, $P = 0.48$; 190 trials: $t = -0.5475$, $P = 0.59$) or error (140 trials, $t = 1.6839$, $P = 0.11$; 150 trials: $t = 2.4259$, $P = 0.03$ (ns, $\alpha = 0.0167$); 190 trials: $t = 1.2789$, $P = 0.22$). This analysis suggests that time alone has little influence on performance. A similar conclusion was met when comparing across all home and scanning behavior (see Supplementary Material).

The experiment was primarily designed to test for changes in RS of fMRI BOLD activity during learning. By pairing sequences with similar practice exposure, this approach could control for longitudinal changes of movement kinematics, as measured by MT or RT. Systematic differences in MT between new and repeated events as a function of condition (MIN, MOD, and EXT) and training scan session (1–3) were evaluated using a three-way (practice exposure × session × repetition) repeated-measures ANOVA. There was a significant main effect of repetition [$F_{1,19} = 133.26$, $P < 0.0001$], revealing that MT for repeat trials was generally faster than new trials. We also found a significant interaction between condition and repetition [$F_{2,38} = 4.52$, $P < 0.05$], which was driven by greater priming for MIN compared with MOD (MIN vs. MOD: $t = 3.03$, $P < 0.01$). While significant, these relative differences between new and repeated trials were small, with the average change for MIN being 115 ms (±73 ms), 73 ms (±35 ms) for MOD, and 86 ms (±38 ms) for EXT. It is doubtful that this small difference in movement duration would have an impact on the magnitude of movement-related activity in motor areas. Further, there was no significant interaction of repetition and session, suggesting that the small differences between new and repeat trials were stationary over the course of the experiment. Thus, there should be no impact on RS measures acquired early versus late in the experiment, and thus, no influence on longitudinal changes of RS. There were no significant effects of repetition on RT.

Behavioral Effects: Predictive Sequence Movements and Performance Plateau

Subjects improved in overall sequence performance, expressed through dramatically faster MTs. By the end of the experiment, subjects could type out a complex, 10-element arpeggio in 1.6 s, with some subjects able to complete a sequence in under 900 ms. This rapidity can only be achieved if subjects no longer rely on the DSP stimuli to guide individual finger responses and strongly suggests that sequences are expressed as a unified action rather than a simple chain of stimulus–response behavior. To test for the emergence of unified sequence execution, IKIs faster than the upper 75% of the IKIs from MIN training scan 3 were defined as predictive. For each practice exposure level and scan session, the percent of predictive trials were compared in a two-way repeated-measures ANOVA (practice exposure × session). There was a main effect for training exposure [$F_{2,38} = 105.37$, $P < 0.00001$] and session [$F_{2,38} = 52.72$, $P < 0.00001$], and interaction [$F_{4,76} = 6.86$, $P < 0.0001$]. The interaction was driven by a larger within-condition increase over sessions for the MOD sequences relative to MIN and EXT (MIN vs. MOD: $t = 3.70$, $P < 0.005$; MOD vs. EXT: $t = 3.18$, $P < 0.005$). Subjects made substantial gains in predictive IKIs between training scans 1 and 2 for MOD ($t = 8.92$, $P < 0.0000001$), at which point they had practiced 200–300 trials. This can also be seen through the acceleration of learning in Figure 2b from 200 to 300 trials. Interestingly, there was a similar jump in predictive sequence movements for EXT ($t = 5.23$, $P < 0.0001$) between training scans 1 and 2. Given the amount of trials needed to reach a virtual plateau in performance (see below section), this jump in predictive IKIs likely reflects the transition from cue- to memory-guided sequence production.

In order to determine when subjects reached a virtual asymptote in performance, MTs for EXT sequences were fit to a double-exponential model. The quality of the fits across individuals were strongly supportive of the model (adjusted

$R^2 = 0.82$). Using the estimated model, the rate of MT change reached an asymptote after 915 trials (±291 trials) according to a conservative incremental rate cutoff of 0.25 ms/trial. On average, subjects reached this asymptote during their 13th home training session (±4 sessions), which was achieved between training scans 1 and 2. This indicates that the majority of subjects had reached asymptotic performance by training scan 2. To test whether asymptotic performance was achieved earlier in training, both MIN and MOD sequences were also evaluated using the same double-exponential model. The model fits were acceptable for both (MIN-adjusted $R^2 = 0.73$; MOD-adjusted $R^2 = 0.77$). However, no subjects reached a plateau in performance for either exposure condition, which is consistent with the above-mentioned finding that approximately double the amount of MOD practice trials are needed to reach a performance plateau.

Imaging Results: Main Effect of the DSP Task as Revealed through Repetition Suppression

The magnitude of RS revealed a robust sensitivity throughout the motor system whenever the same sequence was repeated. The RS response occurred over an extensive network of motor regions (Fig. 3). These regions largely overlap with motor areas that are active during sequencing compared with rest (Grafton et al. 1992, 1995; Jenkins et al. 1994). This widespread pattern of activation suggests that primary and secondary motor system regions can represent information that is needed to either retrieve or execute a specific motor sequence. This main effect collapsed across trials from all exposure conditions from all scan sessions. We used this contrast as a mask in subsequent analyses to characterize learning effects of RS.

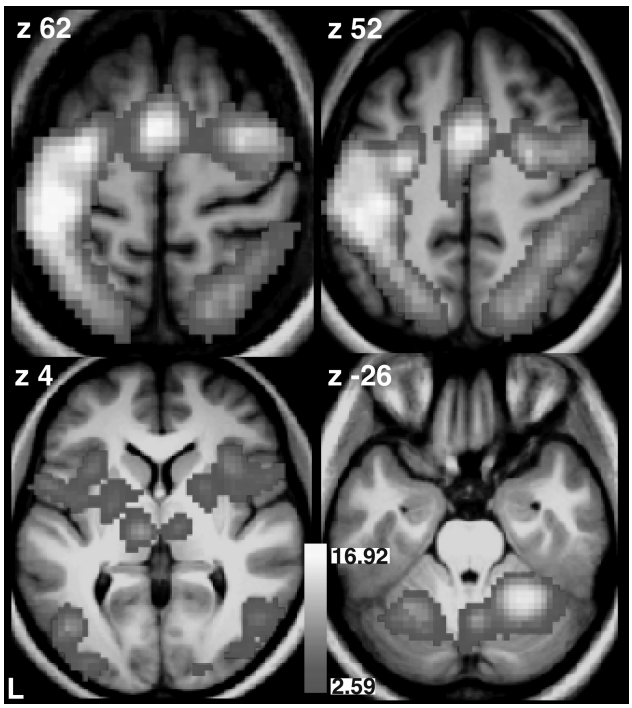

Figure 3. Main effect of RS magnitude for executing the DSP task, collapsed across all sequences, all sessions, and all training intensities. Results are shown at a whole-brain-corrected threshold of $P < 0.05$ (FWE).

The Formation of Initial Sequence Knowledge

Previous accounts of sequence learning have commonly focused on locating rapid changes of brain activation that occur over the course of a single training session. Regions involved in the expression of sequence-specific representations, rather than that which simply pertains to general task knowledge, should show increasing RS over a corresponding rapid timescale. To test for evidence of RS changes on this rapid timescale, we performed a simple linear contrast between MIN training scan 1 (110 trials/sequence) and the pre-training scan (50 trials/sequence). There was a robust increase in RS magnitude in primary and secondary cortical and subcortical motor regions, relative to the pre-training session (Fig. 4, Table 1). This suggests that motor regions, as exemplified by the premotor cortex (PMd/ bilateral PMv) and cerebellum (bilateral lobule VI), support the initial formation of specific sequence knowledge at a rapid timescale. Based on previous single-session learning studies, it does seem possible to find evidence of rapid timescale learning over the course of the pre-training session. However, the comparison of early and late pre-training RS failed to produce any results within the motor system, even at a liberal uncorrected threshold ($P < 0.01$). The lack of within-session change at the start of the experiment is unsurprising, because the sequences were unfamiliar. This is consistent with the perspective that sequence-specific representations did not solidify during the pre-training session. The lack of familiarity therefore leads to the recruitment of similar population activity from one trial to the next, regardless of whether the same sequence has been produced on successive trials. This demonstrates the sensitivity of RS to the development of sequence-specific representations.

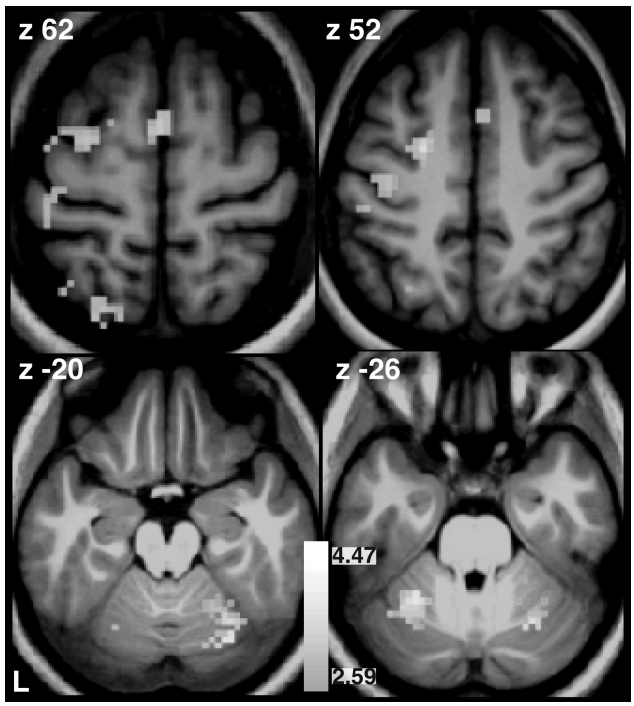

Figure 4. Initial formation of sequence knowledge. Primary and secondary cortical and subcortical motor system regions show increased RS at a rapid timescale (training session 1 MIN > pre-training session). Images are displayed at a corrected threshold using topological FDR ($q < 0.05$), and follow neurological convention (left image is left brain).

Table 1
Brain regions showing effects of the initial formation of sequence knowledge based on increasing RS

Region	Functional name	Side	MNI coordinates			Voxels	Peak *T*
			x	*y*	*z*		
Precentral gyrus	PMv	R	51	−4	43	106	4.47
Cerebellum	Lobule VI	R	30	−61	−26	117	4.35
Superior parietal gyrus	SPL	R	30	−67	46	38	4.17
Intraparietal sulcus	aIPS	L	−39	−43	40	38	4.08
Cerebellum	Lobule V	L	−24	−49	−26	76	4.07
Precentral gyrus	PMv	L	−63	2	28	30	4.01
Precentral gyrus	PMd	L	−21	−4	52	65	3.73
Postcentral gyrus	M1/S1	L	−45	−28	55	63	3.70
Superior frontal gyrus	Pre-SMA/SMA	L	0	5	58	42	3.33
Superior occipital gyrus	sLO	L	−24	−67	58	46	3.32
Intraparietal sulcus	aIPS	R	33	−49	46	20	3.18

Note: Significance for all voxels tested with a group mixed-effects analysis. Sub-maxima for clusters larger than 100 voxels are listed below the main cluster in which they are located. All effects are corrected using topological FDR, or if sub-maxima, standard FDR correction ($q < 0.05$)

Effects of Skill-Specific Efficiency

Previous evidence from skilled performance in humans (Steele and Penhune 2010; Wiestler and Diedrichsen 2013) and long-term learning in non-human primates (Picard et al. 2013) suggests that the expression of skilled movement comes through the generation of efficient neural activity in the motor system. This suggests that there should be regions that show a reduction in RS over the course of training. Skill-specific efficiency was defined as the longitudinal reduction of sequence-specific RS as a function of amount of prior practice during training scans 1–3. As predicted, there was a substantial decrease in RS magnitude as a function of prior training exposure (Fig. 5, Table 2). An expansive cluster of activation was localized to the contralateral left hemisphere cortex. Significant voxels in this cluster spanned dorsally from just rostral to the juncture of the precentral and superior frontal sulci (pre-PMd) caudally through the precentral and postcentral gyrii and into the superior parietal gyrus. Within the cluster, local sub-maxima of note included the M1, S1, the SMA, and the PMd. Extracted parameter estimates of these local peaks all revealed strong decreases in RS magnitude over the course of training. It is important to note that this decrease in RS was occurring while there was also an overall decrease in BOLD magnitude, for both new and repeated trials, as a function of practice exposure (Supplementary Fig. 2). The reduction of RS with practice therefore reflects a compression of the difference in BOLD between new and repeated events. In this respect, skill-specific efficiency emerges through the continued reduction of neural resources that are needed for the expression of a motor skill. Additional regions showing this pattern over the 3 training scans included the bilateral cerebellum (lobule VI), left ventrolateral thalamic nucleus (VL), right PMd, and the right anterior intraparietal sulcus (aIPS). The pattern of decreasing RS magnitude in these regions suggests that they are continually modified over the course of extended learning.

Effects of Skill-Specific Specialization

Skill-specific specialization was defined as a late increase of sequence-specific RS. This was tested using a quadratic function over all trials (sorted by prior training exposure) acquired in training scans 1–3. By using a quadratic, the model allows for initial declines in RS that might occur with efficiency effects, followed by later increases of RS related to specialization. Regions showing this pattern of skill-specific specialization included the left M1, SMA, right anterior cerebellum, and the right ventral premotor (PMv) cortex (Fig. 6, Table 3). Visual inspection of the RS parameter weights shows that each region expresses a complex pattern of RS over the course of training. After the initial increase of RS, there is substantial reduction of RS typically after the first 200 training trials. Critically, these regions show an eventual increase in RS magnitude, which suggests that these regions support the ongoing formation of motor sequences during slow but continued incremental gains in performance. Of note, the RS magnitude weights in these regions begin to increase between 740 and 1480 trials. This corresponds roughly to the same time that behavior has reached a performance plateau, and moreover, after there is a sharp increase in the proportion of predictive sequence movements. We also tested for simple linear increases of RS over training scans 1–3, without any significant results. This suggests that there is no area with pure specialization of function that does not also undergo prior changes related to efficiency.

Overlapping Effects of Initial Learning, Efficiency, and Specialization

Within many regions, there was a striking degree of overlap for changes taking place along different timescales of learning, as revealed through the conjunction analysis (Fig. 7, Table 4). Regions undergoing an initial sensitivity to the formation of sequence representations and then a subsequent decrease in their relative contribution (a pattern consistent with skill-specific efficiency) included the premotor cortex (bilateral PMd/PMv/SMA) and anterior cerebellum (bilateral lobule VI). The primary sensorimotor cortex (M1/S1) demonstrated further complex dynamics over learning with initial increases and decreases of RS as well as subsequent slow increases in sequence-specific representation late in learning. This overall pattern captures an interesting feature in that there is an initial expansion of sensorimotor involvement, which is then followed by the gradual concentration of sequence-specific specialization.

Comparisons with Changes of BOLD Magnitude

Much of what we know regarding functional plasticity of motor sequence learning has come from the standard GLM approach and investigation of general BOLD magnitude change over time. For the current task, this time-dependent approach is of limited value because of the dramatic changes in movement kinematics over time. That was a major reason to employ RS for detecting sequence-specific changes independent of kinematics. Nevertheless, initial learning, efficiency, and specialization contrasts were also carried out with respect to general BOLD magnitude change as an exploratory analysis. Using this approach, we found a prominent efficiency effect across the motor system, which was similar to our skill-specific efficiency results (Supplementary Fig. 3) and also consistent with previous effects of neural efficiency (Wu et al. 2004; Steele and Penhune 2010). However, we failed to detect any changes of BOLD due to initial learning or slow timescales effects. Hence, RS affords a greater sensitivity for detecting longitudinal changes of sequence representations.

Discussion

We identified brain areas supporting the expression of motor sequence representations that were acquired over 6 weeks of

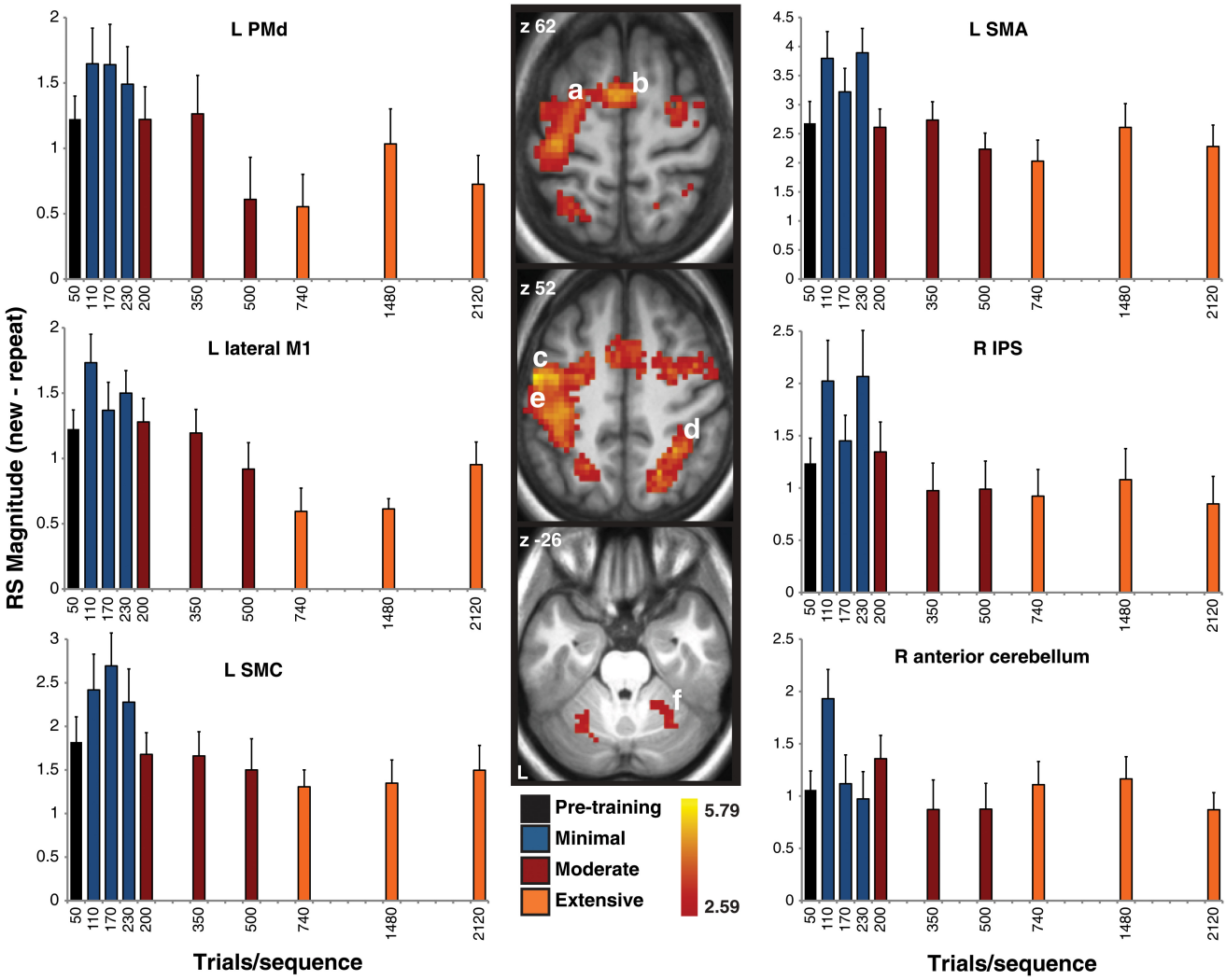

Figure 5. Effects of skill-specific efficiency. Motor system regions reflect skill-specific efficiency as evidenced by decreasing RS magnitude between 110 and 2120 practice trials. Images are displayed at a corrected threshold using topological FDR ($q < 0.05$), and follow neurological convention (left image is left brain). Bar plots are obtained from local maxima and extracted from the parameter estimates, ordered in terms of practice exposure rather than chronological time. SMA, supplementary motor area; M1, primary motor cortex; PMd, dorsal premotor cortex; SMC, somatosensory cortex; IPS, intraparietal sulcus.

training. Improving performance was shaped by the amount of deliberate practice for each sequence without strong support for generalization between sequences. The degree to which motor areas represented a particular sequence changed over 3 timescales. Importantly, by using RS, these changes were interpreted as sequence-specific and free of potential confounds related to changes in kinematics and general task familiarity. Within the first 50–100 trials of practice, when sequence knowledge was initially being formed, both primary and secondary motor areas demonstrated increasing sequence-specific activity, as assessed by RS fMRI. With continued moderate training over 100–700 trials, there was a widespread decline of sequence-specific activity in both motor and other task-related areas, a result that we label as skill-specific efficiency. With more extensive practice over 1000–2000 trials, coinciding with near asymptotic behavioral performance, a small set of motor areas including M1, PMv, SMA, and anterior cerebellum increased in their sequence-specific activity, indicative of late, skill-specific specialization. Critically, we found that many regions changed on at least 2 of these proposed timescales demonstrating that functional contributions during learning are non-monotonic and diverse across cortical regions.

Efficiency effects usually reflect a reduction of absolute BOLD magnitude over time. Here, we consider skill-specific efficiency by testing for a reduction in RS magnitude that is independent of overall reductions of BOLD. Skill-specific efficiency was prominent in the contralateral primary sensorimotor cortex and premotor regions including the PMd/PMv, SMA/CMA, as well as the posterior parietal cortex. These regions also show the general pattern of decreasing BOLD activation as sequence learning progresses, which has been previously observed both at fast (Wu et al. 2004; Floyer-Lea and Matthews 2005; Steele and Penhune 2010) and slower timescales over several months to years of training (Jäncke et al. 2000; Picard et al. 2013). Our results based on RS fMRI sharpen a general model of efficiency by demonstrating that there is a decrease that is sequence-specific.

Visual comparison of the parameter estimates from different regions indicates that there is considerable diversity of how efficiency can be expressed. For example, there was an early sign

of efficiency (110–170 trials) within lobule VI of the anterior cerebellum, which might be related to the reduction of online error correction (Doyon and Benali 2005; Grafton et al. 2008). On the other hand, the left PMd, which is consistently involved during sequence learning (Hardwick et al. 2013), demonstrated a slower appearance of skill-specific efficiency (350–500 trials). This more sustained decay might reflect a longer time to optimize the visuomotor task, which is in line with the initial increase in predictive sequence behavior. This regional variability in the time course of efficiency undermines any ability to use changes of RS within any single region to predict transitions between learning stages.

Decreasing RS could be due to either overall population fatigue or sharpening due to the reduction of non-essential neurons (Grill-Spector et al. 2006). We speculate that there is an overall reduction in non-essential neurons, leaving just those neurons that are needed to express a particular motor sequence skill. This suggests that over moderate timescales of practice, the motor system resolves the initial cost of sustaining a newly acquired motor skill through the gradual emergence of a reduced set of optimally tuned neuronal populations and elimination of non-essential populations.

We also found that a subset of regions demonstrated an increase in RS magnitude at a slower timescale. Increasing RS magnitude implies that there is an increase in representational strength as the sequences were practiced beyond a point of asymptotic performance. Because of the closely matched performance on repeated trials with the RS paradigm, these increases are unlikely to be due to faster rates of movement. Skill-specific specialization occurred in the left sensorimotor cortex, SMA, right PMv, and the ipsilateral (right) anterior cerebellum. These regions are all known to shape corticospinal output (Dum and Strick 2005).

The left sensorimotor cortex, encompassing both banks of the central sulcus, demonstrated substantial changes along fast, moderate, and slow timescales (Fig. 7). Previous studies

Table 2
Brain regions showing evidence of RS-based neural efficiency

Region	Functional name	Side	MNI coordinates			Voxels	Peak *T*
			x	*y*	*z*		
Precentral gyrus	M1 (4a)	L	−51	−7	52	1598	5.79
Postcentral gyrus	S1 (BA2)	L	−51	−22	43		5.71
Precentral gyrus	M1 (4a)	L	−45	−13	46		5.51
Precentral gyrus	PMd	L	−42	−4	49		5.30
Cingulate gyrus	CMA	R	12	5	49		5.12
Superior frontal gyrus	SMA	L	0	2	58		5.06
Precentral gyrus	M1 (4p)	L	−36	−25	49		4.99
Postcentral gyrus	S1 (BA2)	L	−45	−28	52		4.68
Intraparietal sulcus	aIPS	L	−33	−40	43		4.64
Inferior occipital gyrus		R	45	−67	−8	119	5.21
Intraparietal sulcus	aIPS	R	30	−46	49	388	5.07
Superior parietal gyrus	SPL	R	21	−58	52		5.00
Intraparietal sulcus	aIPS	R	36	−37	46		4.91
Precentral gyrus	PMd	R	27	−4	55	281	4.59
Occipital pole		L	−15	−91	−11	37	4.40
Parietal operculum		L	−54	−19	19	34	4.10
Thalamus	VL	L	−15	−13	4	46	3.98
Inferior occipital gyrus		L	−42	−79	−2	63	3.88
Cerebellum	Lobule VI	L	−21	−64	−23	35	3.85
Cerebellum	Lobule VI	R	30	−58	−29	35	3.47

Note: Significance for all voxels tested with a group mixed-effects analysis. Sub-maxima for clusters larger than 100 voxels are listed below the main cluster in which they are located. All effects are corrected using topological FDR, or if sub-maxima, standard FDR correction ($q < 0.05$).

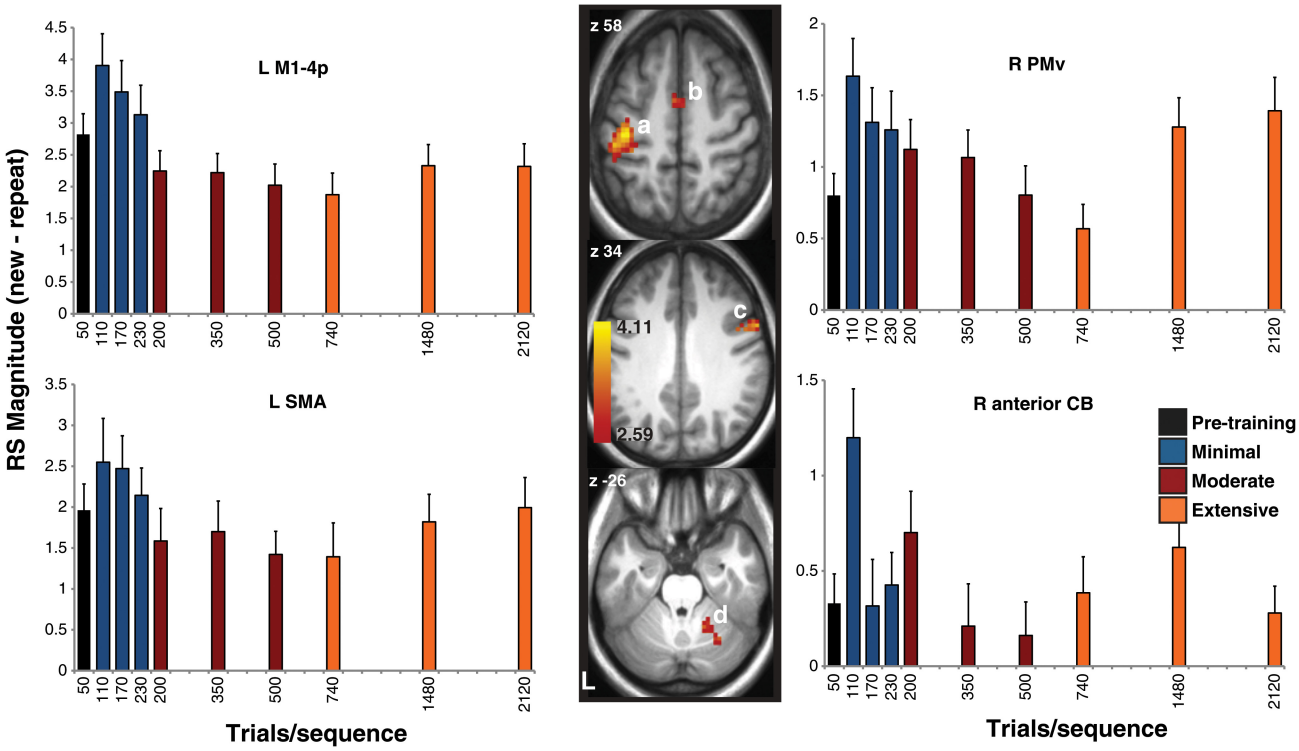

Figure 6. Effects of skill-specific specialization. Motor system regions show a quadratic RS response as a function of physical practice, with early decreases followed by late increases in RS. Images are displayed at a corrected threshold using topological FDR ($q < 0.05$), and follow neurological convention (left image is left brain). Bar plots are obtained from local maxima, ordered in terms of practice exposure rather than chronological time. M1, primary motor cortex; PMv, ventral premotor cortex; SMA, supplementary motor area; CB, cerebellum.

Table 3
Brain regions showing evidence of RS-based neural specialization

Region	Functional name	Side	MNI coordinates			Voxels	Peak *T*
			x	*y*	*z*		
Precentral gyrus	M1 (4p)	L	−36	−22	55	141	4.11
Precentral gyrus	PMv	R	60	8	34	27	3.70
Cerebellum	Lobule VI	R	24	−64	−29	39	3.42
Superior frontal gyrus	SMA	R	3	−4	52	28	3.35

Note:Significance for all voxels tested with a group mixed-effects analysis. Sub-maxima for clusters larger than 100 voxels are listed below the main cluster in which they are located. All effects are corrected using topological FDR, or if sub-maxima, standard FDR correction ($q < 0.05$).

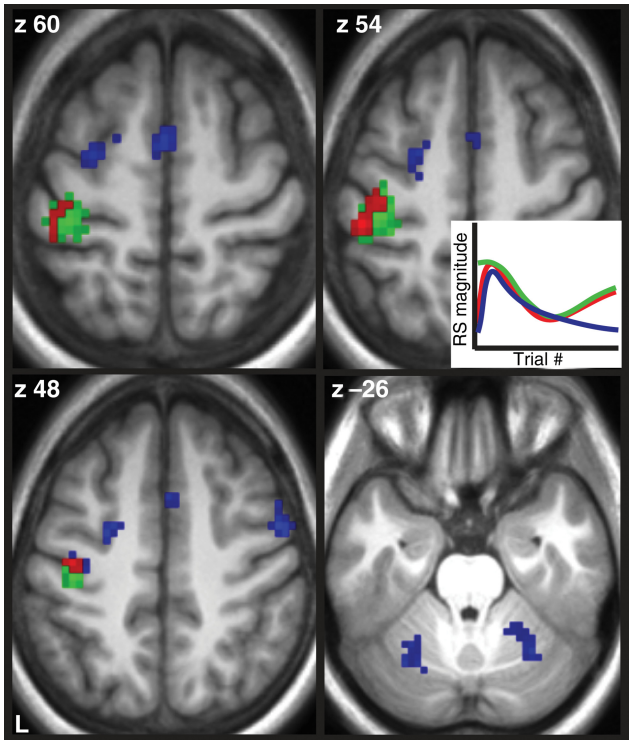

Figure 7. Conjunction of functional timescales of learning. Colors correspond to the timescale curves shown on the middle-right of the image. Red indicates initial learning, efficiency, and specialization, blue indicates initial learning and efficiency, and green indicates efficiency and specialization.

have observed increasing extent of M1 activation with long-term learning (Karni et al. 1995; Hlustík et al. 2004) as well as sequence-specific neuronal activity (Matsuzaka et al. 2007). The current study extends these results because subjects always produced the sequences as quickly as possible, without any external pacing, so there was close correspondence with what was actually learned during practice. Furthermore, by using RS, we could establish that the changes in M1 were skill-specific and not due to effects related to general task familiarity.

Other human studies investigating functional change in M1 with extensive practice have relied on comparisons between novices and experts (Jäncke et al. 2000; Krings et al. 2000) or have compared over-learned with random sequences in the macaque (Picard et al. 2013). It is interesting to note that M1 change in these long-term studies is dominated by neural efficiency. On the other hand, studies of structural change show an overall expansion of cortical volume in task-relevant structures (Gaser and Schlaug 2003; Han et al. 2009; Park et al. 2009). Further studies are clearly needed to match longitudinal functional and structural changes at the same timescales and training intensities to see how these changes might interact.

Table 4
Brain regions showing conjunction effects based on the strict overlap of different timescales of learning

Region	Functional name	Side	MNI coordinates			Voxels	Peak *T*
			x	*y*	*z*		
Overlap between initial learning, efficiency, and specialization							
Postcentral gyrus	S1	L	−45	−28	55	51	3.70
Precentral gyrus	M1	L	−39	−16	49		3.33
Overlap between initial learning and efficiency							
Precentral gyrus	PMv	R	51	−4	43	49	4.38
Precentral gyrus	PMd	R	42	−1	46		3.14
Postcentral gyrus	S1	L	−45	−28	55	60	3.70
Precentral gyrus	M1	L	−39	−16	49		3.33
Cerebellum	Lobule VI	L	−24	−64	−23	24	3.49
Precentral gyrus	PMd	L	−21	−4	52	54	3.49
Cerebellum	Lobule VI	R	30	−58	−26	28	3.36
Superior frontal gyrus	SMA	L	0	5	58	32	3.33
Overlap between efficiency and specialization							
Precentral gyrus	M1	L	−36	−22	55	131	4.11
Postcentral gyrus	S1	L	−42	−28	58		3.84

Note: Significance for all voxels tested with a group mixed-effects analysis. Individual time effects are corrected using topological FDR, or if sub-maxima, standard FDR correction ($q < 0.05$).

Our locus of skill-specific specialization in M1 overlapped with the posterior section of M1, referred to as either area 4p or "new M1" (Geyer et al. 1996; Rathelot and Strick 2006). Unlike 4a, the pyramidal neurons of 4p make far more direct monosynaptic connections with motoneurons (Rathelot and Strick 2006, 2009). One hypothesis for the evolution of 4p is that it allows greater manual dexterity without the reliance on central pattern generators or motor primitives from within the spinal cord, which may be critical for sculpting novel muscle synergies (Rathelot and Strick 2009). Furthermore, neuroimaging studies suggest that 4p is involved in execution of abstract motor behavior (Binkofski et al. 2002; Sharma et al. 2008) and that activation of 4p is critical in recovery from stroke (Ward et al. 2003; Sharma et al. 2009).

The SMA is modulated at both fast and moderate timescales, with activation increasing following the completion of relatively small amounts of practice (Grafton et al. 1992, 2002; Bischoff-Grethe et al. 2004; Floyer-Lea and Matthews 2005). We too observed that the pre-SMA/SMA demonstrated both an initial rapid increase in RS magnitude as well as skill-specific efficiency, which emerged following the completion of approximately 200 trials. Importantly, we found that a caudal portion of pre-SMA/SMA demonstrated an effect of skill-specific specialization (Fig. 6). It is curious to note that despite strong support from non-human primate research (Tanji and Shima 1994; Shima and Tanji 2000), direct evidence for the involvement of the SMA in the expression of skilled motor sequences is lacking. Other long-term studies failed to observe slow timescale effects of SMA (Hlustík et al. 2004; Lehéricy et al. 2005), which could be due to the fact that previous experiments confounded internally generated movement by clamping performance during imaging.

The anterior cerebellum showed evidence of skill-specific efficiency at a moderate timescale (Fig. 5), as well as specialization on a slow timescale (Fig. 6). Evidence for the involvement of cerebellum in long-term learning is supported by cross-sectional studies that report increased gray matter differences in experts relative to controls (Gaser and Schlaug 2003; Han

et al. 2009; Park et al. 2009). Further, a recently updated model raises the notion that the cerebellum is continually active during the expression of sequence skills, in order to update forward-model parameters (Penhune and Steele 2012). Given the interactivity between motor cortical areas and cerebellum during learning (Grafton et al. 2008; Steele and Penhune 2010; Galea et al. 2011), our results suggest that increased motor representations late in training reflect both a motor plan in M1 (including SMA) and a forward model in the cerebellum.

Sequence-specific activity also increased in the right PMv. Although little direct evidence suggests a role for PMv in long-term learning, it is known to be readily involved in the representation, decoding, and implementation of sequential behavior (di Pellegrino et al. 1992; Grafton et al. 1996, 1997). It is also involved in the prediction of perceptual sequence patterns (Schubotz and von Cramon 2003). Similar to known properties of the medial premotor cortex (Shima and Tanji 2000), the PMv also appears to support the structure or organization of sequences.

Previous studies (Doyon et al. 2002; Lehéricy et al. 2005; Coynel et al. 2010) and related models of sequence learning (Hikosaka et al. 2002; Doyon and Benali 2005) suggest that the basal ganglia are differentially engaged over the course of sequence learning. While we did find robust engagement of the basal ganglia throughout the task (Fig. 3), we failed to find any modulation of RS magnitude as a function of physical practice, even at more liberal uncorrected thresholds. Our results suggest that the basal ganglia are consistently engaged in the expression of skill-specific representations throughout motor learning. This does not exclude the possibility that these structures are involved in learning within the current context; however, they do suggest that the basal ganglia are not critical for the storage of sequence representations. This observation is consistent with an account that the basal ganglia are modulated by kinematics but not necessarily the storage of long-term motor sequence information (Desmurget and Turner 2010; Turner and Desmurget 2010).

There are some potential limitations of our study that merit future explanation. We did not directly test whether the observed timescales of changing brain activity corresponded to particular cognitive operations. Specifically, there was no control over implicit or explicit awareness. Like a musician trying to learn a new arpeggio, our subjects always knew they were learning sequences. The behavioral data show the participants could predict a sequence after approximately 250 trials. Thus, over the course of the next 1850 trials, changes in the brain were unlikely to be based on state transitions such as the emergence of awareness. Furthermore, we cannot link the current results to studies demonstrating the important role of sleep on memory consolidation in the first few days of practice. We assume that over the course of 6 weeks of practice, the impact of initial memory consolidation can be discounted, and that the overall effect of sleep would only be relevant for analyzing individual differences of learning rates.

Supplementary Material

Supplementary Material can be found at http://www.cercor.oxfordjournals.org/ online.

Funding

This material is based on work supported by the US National Institutes of Health (grant: P01 NS044393).

Notes

The authors thank Robert S. Turner, Peter Strick, F. Greg Ashby, and Konrad Koerding for helpful discussions. *Conflict of Interest*: None declared.

References

Abrahamse EL, Ruitenberg MFL, De Kleine E, Verwey WB. 2013. Control of automated behavior: insights from the discrete sequence production task. Front Hum Neurosci. 7(82):1–16.

Binkofski F, Fink GR, Geyer S, Buccino G, Gruber O, Shah NJ, Taylor JG, Seitz RJ, Zilles K, Freund H-J. 2002. Neural activity in human primary motor cortex areas 4a and 4p is modulated differentially by attention to action. J Neurophysiol. 88(1):514–519.

Bischoff-Grethe A, Goedert KM, Willingham DT, Grafton ST. 2004. Neural substrates of response-based sequence learning using fMRI. J Cogn Neurosci. 16(1):127–138.

Bryan WL, Harter N. 1897. Studies in the physiology and psychology of telegraphic language. Psychol Rev. 4:27–53.

Calhoun VD, Stevens MC, Pearlson GD, Kiehl KA. 2004. fMRI analysis with the general linear model: removal of latency-induced amplitude bias by incorporation of hemodynamic derivative terms. Neuroimage. 22:252–257.

Classen J, Liepert J, Wise SP, Hallett M, Cohen LG. 1998. Rapid plasticity of human cortical movement representation induced by practice. J Neurophysiol. 79:1117–1123.

Coynel D, Marrelec G, Perlbarg V, Pélégrini-Issac M, Van de Moortele P-F, Ugurbil K, Doyon J, Benali H, Lehéricy S. 2010. Dynamics of motor-related functional integration during motor sequence learning. Neuroimage. 49(1):759–766.

Crossman ERFW. 1959. A theory of the acquisition of speed-skill. Ergonomics. 2:153–166.

Desmurget M, Turner RS. 2010. Motor sequences and the basal ganglia: kinematics, not habits. J Neurosci. 30(22):7685–7690.

di Pellegrino G, Fadiga L, Fogassi L, Gallese V, Rizzolatti G. 1992. Understanding motor events: a neurophysiological study. Exp Brain Res. 91(1):176–180.

Doyon J, Benali H. 2005. Reorganization and plasticity in the adult brain during learning of motor skills. Curr Opin Neurobiol. 15:161–167.

Doyon J, Penhune V, Ungerleider LG. 2003. Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning. Neuropsychologia. 41:252–262.

Doyon J, Song AW, Karni A, Lalonde F, Adams MM, Ungerleider LG. 2002. Experience-dependent changes in cerebellar contributions to motor sequence learning. Proc Natl Acad Sci USA. 99:1017–1022.

Dum RP, Strick PL. 2005. Frontal lobe inputs to the digit representations of the motor areas on the lateral surface of the hemisphere. J Neurosci. 25(6):1375–1386.

Elbert T, Pantev C, Wienbruch C, Rockstroh B, Taub E. 1995. Increased cortical representation of the fingers of the left hand in string players. Science. 270:305–307.

Floyer-Lea A, Matthews PM. 2005. Distinguishable brain activation networks for short-and long-term motor skill learning. J Neurophysiol. 94(1):512–518.

Galea JM, Vazquez A, Pasricha N, de Xivry J-JO, Celnik P. 2011. Dissociating the roles of the cerebellum and motor cortex during adaptive learning: the motor cortex retains what the cerebellum learns. Cereb Cortex. 21(8):1761–1770.

Gaser C, Schlaug G. 2003. Brain structures differ between musicians and non-musicians. J Neurosci. 23(27):9240–9245.

Geyer S, Ledberg A, Schleicher A, Kinomura S, Schormann T, Bürgel U, Klingberg T, Larsson J, Zilles K, Roland PE. 1996. Two different areas within the primary motor cortex of man. Nature. 382:805–807.

Grafton ST, Arbib MA, Fadiga L, Rizzolatti G. 1996. Localization of grasp representations in humans by positron emission tomography. 2. Observation compared with imagination. Exp Brain Res. 112(1):103–111.

Grafton ST, Fadiga L, Arbib MA, Rizzolatti G. 1997. Premotor cortex activation during observation and naming of familiar tools. NeuroImage. 6(4):231–236.

Grafton ST, Hazeltine E, Ivry R. 1995. Functional mapping of sequence learning in normal humans. J Cogn Neurosci. 7(4):497–510.
Grafton S, Hazeltine E, Ivry R. 2002. Motor sequence learning with the nondominant left hand. Exp Brain Res. 146(3):369–378.
Grafton ST, Mazziotta JC, Presty S, Friston KJ, Frackowiak RS, Phelps ME. 1992. Functional anatomy of human procedural learning determined with regional cerebral blood flow and PET. J Neurosci. 12 (7):2542–2548.
Grafton ST, Schmitt P, Van Horn J, Diedrichsen J. 2008. Neural substrates of visuomotor learning based on improved feedback control and prediction. Neuroimage. 39(3):1383–1395.
Grill-Spector K, Henson R, Martin A. 2006. Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn Sci. 10(1): 14–23.
Grinband J, Wager TD, Lindquist M, Ferrera VP, Hirsch J. 2008. Detection of time-varying signals in event-related fMRI designs. Neuroimage. 43(3):509–520.
Han Y, Yang H, Lv Y-T, Zhu C-Z, He Y, Tang H-H, Gong Q-Y, Luo Y-J, Zang Y-F, Dong Q. 2009. Gray matter density and white matter integrity in pianists' brain: a combined structural and diffusion tensor MRI study. Neurosci Lett. 459(1):3–6.
Hardwick RM, Rottschy C, Miall RC, Eickhoff SB. 2013. A quantitative meta-analysis and review of motor learning in the human brain. Neuroimage. 67:283–297.
Hikosaka O, Nakamura K, Sakai K, Nakahara H. 2002. Central mechanisms of motor skill learning. Curr Opin Neurobiol. 12(2):217–222.
Hlustík P, Solodkin A, Noll DC, Small SL. 2004. Cortical plasticity during three-week motor skill learning. J Clin Neurophysiol. 21(3): 180–191.
Jäncke L, Shah NJ, Peters M. 2000. Cortical activations in primary and secondary motor areas for complex bimanual movements in professional pianists. Brain Res Cogn Brain Res. 10(1–2):177–183.
Jenkins IH, Brooks DJ, Nixon PD, Frackowiak RS, Passingham RE. 1994. Motor sequence learning: a study with positron emission tomography. J Neurosci. 14(6):3775–3790.
Karni A, Meyer G, Jezzard P, Adams MM, Turner R, Ungerleider LG. 1995. Functional MRI evidence for adult motor cortex plasticity during motor skill learning. Nature. 377(6545):155–158.
Krings T, Töpper R, Foltys H, Erberich S, Sparing R, Willmes K, Thron A. 2000. Cortical activation patterns during complex motor tasks in piano players and control subjects. A functional magnetic resonance imaging study. Neurosci Lett. 278(3):189–193.
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 USA. 102(35):12566–12571.
Matsuzaka Y, Picard N, Strick PL. 2007. Skill representation in the primary motor cortex after long-term practice. J Neurophysiol. 97(2): 1819–1832.
Nichols T, Brett M, Andersson J, Wager T, Poline JB. 2005. Valid conjunction inference with the minimum statistic. Neuroimage. 25(3): 653–666.
Park IS, Lee KJ, Han JW, Lee NJ, Lee WT, Park KA, Rhyu IJ. 2009. Experience-dependent plasticity of cerebellar vermis in basketball players. Cerebellum. 8:334–339.
Penhune VB, Doyon J. 2005. Cerebellum and M1 interaction during early learning of timed motor sequences. Neuroimage. 26:801–812.
Penhune VB, Doyon J. 2002. Dynamic cortical and subcortical networks in learning and delayed recall of timed motor sequences. J Neurosci. 22:1397–1406.
Penhune VB, Steele CJ. 2012. Parallel contributions of cerebellar, striatal and M1 mechanisms to motor sequence learning. Behav Brain Res. 226(2):579–591.
Picard N, Matsuzaka Y, Strick PL. 2013. Extended practice of a motor skill is associated with reduced metabolic activity in M1. Nat Neurosci. 16(9):1340–1347.
Poldrack RA. 2000. Imaging brain plasticity: conceptual and methodological issues—a theoretical review. Neuroimage. 12(1): 1–13.
Rathelot J-A, Strick PL. 2006. Muscle representation in the macaque motor cortex: an anatomical perspective. Proc Natl Acad Sci USA. 103(21):8257–8262.
Rathelot J-A, Strick PL. 2009. Subdivisions of primary motor cortex based on cortico motoneuronal cells. Proc Natl Acad Sci USA. 106 (3):918–923.
Rosenkranz K, Williamon A, Rothwell JC. 2007. Motorcortical excitability and synaptic plasticity is enhanced in professional musicians. J Neurosci. 27:5200–5206.
Schubotz RI, von Cramon DY. 2003. Functional-anatomical concepts of human premotor cortex: evidence from fMRI and PET studies. NeuroImage. 20:S120–S131.
Sharma N, Jones PS, Carpenter TA, Baron J-C. 2008. Mapping the involvement of BA 4a and 4p during motor Imagery. NeuroImage. 41 (1):92–99.
Sharma N, Simmons LH, Jones PS, Day DJ, Carpenter TA, Pomeroy VM, Warburton EA, Baron J-C. 2009. Motor imagery after subcortical stroke: a functional magnetic resonance imaging study. Stroke. 40 (4):1315–1324.
Shima K, Tanji J. 2000. Neuronal activity in the supplementary and presupplementary motor areas for temporal organization of multiple movements. J Neurophysiol. 84(4):2148–2160.
Steele CJ, Penhune VB. 2010. Specific increases within global decreases: a functional magnetic resonance imaging investigation of five days of motor sequence learning. J Neurosci. 30(24): 8332–8341.
Steffener J, Tabert M, Reuben A, Stern Y. 2010. Investigating hemodynamic response variability at the group level using basis functions. Neuroimage. 49(3):2113–2122.
Tanji J, Shima K. 1994. Role for supplementary motor area cells in planning several movements ahead. Nature. 371:413–416.
Turner RS, Desmurget M. 2010. Basal ganglia contributions to motor control: a vigorous tutor. Curr Opin Neurobiol. 20(6):704–716.
Ward NS, Brown MM, Thompson AJ, Frackowiak RSJ. 2003. Neural correlates of motor recovery after stroke: a longitudinal fMRI study. Brain. 126(11):2476–2496.
Wiestler T, Diedrichsen J. 2013. Skill learning strengthens cortical representations of motor sequences. Elife. 2:e00801.
Wu T, Kansaku K, Hallett M. 2004. How self-initiated memorized movements become automatic: a functional MRI study. J Neurophysiol. 91(4):1690–1698.
Xiong J, Ma L, Wang B, Narayana S, Duff EP, Egan GF, Fox PT. 2009. Long-term motor training induces changes in regional cerebral blood flow in both task and resting states. Neuroimage. 45(1): 75–82.