Symbolic Representations in Motor Sequence Learning

J. Bo[1,*], **S.J. Peltier**[2], **D.C. Noll**[2], and **R. D. Seidler**[1,3]

[1] School of Kinesiology, University of Michigan, Ann Arbor, MI 48109-1109 USA

[2] Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI 48109-1109 USA

[3] Department of Psychology, University of Michigan, Ann Arbor, MI 48109-1109 USA

Abstract

It has been shown that varying the spatial versus symbolic nature of stimulus presentation and response production, which affects stimulus-response (S-R) mapping requirements, influences the magnitude of implicit sequence learning (Koch & Hoffman, 2000). Here, we evaluated how spatial and symbolic stimuli and responses affect the neural bases of sequence learning. We selectively eliminated the spatial component of stimulus presentation (spatial versus symbolic), response execution (manual versus vocal), or both. Fourteen participants performed the alternating serial reaction time task under these conditions in an MRI scanner, with interleaved acquisition to allow for recording of vocal response reaction times. Nine regions of interest (ROI) were selected to test the hypothesis that the dorsolateral prefrontal cortex (DLPFC) was preferentially engaged for spatially cued conditions and cerebellum lobule HVI, crus I & II were associated with symbolically cued learning. We found that left cerebellum lobule HVI was selectively recruited for symbolic learning and the percent signal change in this region was correlated with learning magnitude under the symbolic conditions. In contrast, the DLPFC did not exhibit selective activation for learning under spatial conditions. The inferior parietal lobule exhibited increased activation during learning regardless of the condition, supporting its role in forming an abstract representation of learned sequences. These findings reveal different brain networks that are flexibly engaged depending on the conditions of sequence learning.

Keywords

sequence learning; spatial; symbolic; fMRI; cerebellum

Introduction

Motor sequence learning has been studied extensively with neuroimaging techniques. While it is clear that motor, perceptual, and cognitive systems are engaged during this type of learning, the specific contributions of different brain regions are not well understood (cf. Ashe et al. 2006). The serial reaction time (SRT) task has been widely used to investigate implicit sequence learning (Nissen & Bullemer, 1987). In the typical SRT task, stimuli appear successively at different spatial locations in a repeating sequence, and participants respond by pressing a

corresponding key. Previous studies have used variants of the task that manipulate the degree of required stimulus-response (S-R) mapping, as illustrated in Figure 1.

It has been reported that varying the spatial versus symbolic nature of stimulus presentation and response production may influence implicit sequence learning (c.f. Koch & Hoffman, 2000). Roberston et al. (2001) reported that the dorsolateral prefrontal cortex (DLPFC) plays a specific role in spatial sequence learning. They found that spatially cued sequence learning was abolished following low frequency repetitive transcranial magnetic stimulation (rTMS) applied to the DLPFC, while symbolically cued sequence learning was not affected. However, Schwarb & Schumacher (2009) have recently argued that the DLPFC plays a general role in spatial mapping rather than specifically for sequence learning. The authors orthogonally manipulated spatial sequence structure (sequence versus random) and spatial S-R compatibility (i.e., spatial correspondence between stimuli and response fingers). They found a significant main effect of S-R compatibility but not sequence learning in the DLPFC (Schwarb & Schumacher, 2009). Thus, it is not clear whether the DLPFC makes a specific contribution to spatial sequence learning per se.

Recently, Spencer & Ivry (2009) reported that patients with cerebellar lesions were impaired in learning sequences when the responses were symbolically cued, but had intact learning under direct cueing (i.e. participants directly touched the spatially-arranged stimuli), implicating a role for this structure in maintaining stimulus-response relationships during learning. However, given that the patients in this study had diffuse bilateral degeneration, the laterality and localization of this function within the cerebellum remains unclear. Using a delayed stimulus-response task, Balsters & Ramnani (2008) argued that the cerebellum lobule HIV and crus I & II encode symbolic representations of action. They found higher cerebellar activation when movements were cued with geometrical symbols compared to when movements were directly cued with fingers. Thus, it is possible that the cerebellum lobule HIV and crus I & II play a similar role in symbolic sequence learning.

Based on the previous literature, there is evidence to suggest that varying networks are engaged for spatial (DLPFC) and symbolic (cerebellum) representations of implicit sequence learning. However, previous studies have varied stimulus-response mapping demands by either changing the mode of stimulus presentation or response production, but none have investigated both manipulations in combination. Thus, it remains unclear how these factors map onto the brain regions engaged during sequence learning. In the current fMRI study, we selectively eliminated the spatial component of stimulus presentation (spatial versus symbolic), response execution (manual versus vocal), or both to examine the neural networks that encode spatial versus symbolic motor sequence representations. We selected nine regions of interest to test the hypotheses that the cerebellum HVI and crus I & II would be preferentially engaged for symbolically cued learning, while the DLPFC would be specifically engaged for spatially cued learning (see Figure 1).

In addition, experiments using spatially (i.e., stimulus locations indicate the appropriate motor response) or symbolically cued stimuli (i.e., colors or letters indicate the appropriate motor response) have shown that supplementary motor area (SMA), sensorimotor cortex, putamen and inferior parietal cortex are engaged for implicit sequence learning independently of the stimulus cueing characteristics (Grafton et al., 1995; Grafton et al., 1998; Hazeltine et al., 1997). Thus, consistent with this literature, we expected that the sensorimotor cortex, SMA, putamen and inferior parietal lobule (IPL) would be activated for sequence learning across all experimental conditions in the current study.

Methods

Participants

Fourteen right-handed (determined by self-report and the Edinburgh handedness inventory (Oldfield, 1971), mean age = 21.4 years (±2.5), 6 males) individuals participated in this study. The mean score for the Edinburgh handedness inventory was 0.87±0.12. All the experimental procedures were approved by the Institutional Review Board of the University of Michigan. Individuals gave their informed consent and were paid for their participation.

Procedure

Sequence learning task—On the first day of testing, participants lay supine in a mock scanner with a mirror adjusted to allow visualization of a video screen. They performed two runs of training. Each run consisted of four, 10-trial conditions (i.e. spatial-manual, spatial-vocal, symbolic-manual and symbolic-vocal). Each trial was made up of 8 randomly generated elements. Each condition began with a 10-second instruction and a 20-second resting period during which the participant visually fixated on the center of the screen. Then, the 8-element sequence was presented element by element every 1000 ms, and the participants were instructed to make the appropriate response as soon as they saw the stimulus. After each 8-second stimulus-response period, participants looked at the fixation cross for 4 seconds. There were four experimental conditions (Figure 2): 1): *spatial-manual*, four visual stimulus boxes were presented on the screen with one of four letters appearing in one of the stimulus boxes. The positions of these four letters were fixed to the four stimulus boxes. Participants were instructed to press one of four corresponding buttons as fast as possible on a key-press device with the fingers of their right hand. The letters A, B, C and D were mapped onto the index, middle, ring and little fingers of the right hand. 2): *spatial-vocal*, participants were asked to make the vocal responses: *ant, bear, cat,* and *dog* when seeing the stimulus letters, *A, B, C,* and *D* respectively. We chose to use word responses instead of letter reading because Hartman et al. (1989) found more consistent implicit sequence learning with categorical rather than letter reading responses. The stimulus presentation was identical to that of condition 1, except that participants responded with spoken words instead of manually pressing the response buttons. 3): *symbolic-manual,* this condition involved only one centrally located stimulus box. Participants were asked to make the appropriate finger response corresponding to the letter appearing in the centrally located visual stimulus box. 4): *symbolic-vocal*, participants made the same vocal responses *ant, bear, cat,* and *dog* when the letters appeared in the single visual stimulus box. Stimuli were controlled by a PC using custom software written in E-Prime ™ version 1.0 (Psychology Software Tools, Pittsburgh) and were presented on a mirror mounted on the mock head coil, reflecting the video projection screen at the rear of the scanner. This training with randomly presented stimuli familiarized participants with the task and the scanning environment.

On the second day of testing, participants were positioned in a 3.0 Tesla GE MRI scanner (General Electric, Waukesha, WI) and performed five alternating serial reaction time (ASRT) (Howard et al., 2004) runs and two random runs (one at the beginning of the experiment and one at the end). Each run was collected as a separate fMRI run and lasted 10 minutes. Stimuli were presented through a mirror mounted on a set of specialized goggles, reflecting the video projection screen at the rear of the scanner. During the ASRT task, each trial contained an 8-element sequence in which fixed and random elements were alternated (i.e. D r B r A r C r). During the random run, each trial was made up of 8 randomly generated elements. Each condition began with a 10-second instruction and a 20-second resting period in which the participant visually fixated the center of the screen. For the two manual conditions, the instruction was "In the following condition, you need to press one of four corresponding buttons as fast as possible with one of your fingers." For the two vocal conditions, the instruction was

"In the following condition, you need to speak to the microphone with one of four corresponding words (i.e. *ant, bear, cat* and *dog*) as fast as possible." Then, the 8-element sequence was presented element by element every 1000 ms, and the participants were instructed to make the appropriate response as soon as they saw the stimulus. After each 8-second stimulus-response period, participants looked at the fixation cross for 4 seconds. The scanner remained quiet (see below) throughout the 8-second stimulus-response period to allow for recording of vocal responses and then was triggered to acquire two functional volumes within the next 4 seconds (TR = 2 seconds). This delayed interleaved scanning method (Liegeois et al., 2003) enables motion artifacts generated during speech to be avoided (Barch et al., 1999), while still allowing for detection of the resulting hemodynamic response, which typically peaks 5 to 6 seconds after an event of interest (Figure 3). After the 4-second data acquisition period, the fixation cross was replaced by the first element of the next sequence. Each complete 8-element sequence plus 4 seconds of scanning defined an individual trial. Each run consisted of four, 10-trial conditions (i.e. spatial-manual, spatial-vocal, symbolic-manual and symbolic-vocal). The order of these four conditions was randomized within a run. The same sequence was used across the four conditions for each participant.

Following the experiment, we interviewed participants to probe their awareness of the sequence. Five increasingly specific questions were asked: 1) Do you have anything to report regarding the task? 2) Did you notice anything special about the task or the materials? 3) Did you notice any regularity in the way the stimulus was moving on the screen? 4) If you responded "yes" to question 3, please answer this question: Did you attempt to take advantage of the regularities you noticed in order to anticipate subsequent targets? Is so, did this help? 5). In fact, there was some regularity to the sequences you observed. What do you think it was? Try to describe any regularity you think might have been there. During this experiment, no participant could explicitly express the structure of the sequence in the questionnaire, supporting that learning was implicit (Howard et al., 2004).

fMR Image acquisition—Images were acquired using a 3.0 Tesla GE MRI scanner at the University of Michigan's Functional Magnetic Resonance Imaging Center. The standard GE head coil was used. Each imaging session began with the acquisition of an anatomical localizer. The first scan was a sagittal T1-weighted image. The repetition time (TR) was 200ms, time echo (TE) was 5.7ms, flip angle (FA) was 90, field of view (FOV) was 220mm and the voxel size was $1 \times 1 \times 1.2$mm. This image was used to identify landmarks for the position of the subsequent scans. Functional images were then acquired using a custom single-shot gradient-echo (GRE) reverse spiral pulse sequence. Pulse sequence parameters were TR/TE/FA/FOV of 2000ms/30ms/90/220mm and a voxel size of $3.44 \times 3.44 \times 3$mm. Forty 3.0 mm thick slightly oblique axial slices (no gap) were acquired. We used an interleaved scanning method to avoid scanner noise during the vocal responses (cf. Liegeois et al., 2003) and to eliminate the image artifact that is associated with jaw movement. This interleaved method allowed for a relatively "quiet" period of 8 seconds (i.e. during the 8-second stimulus-response period) followed by acquisition of two functional volumes within the next 4 seconds (i.e. no manual or vocal responses were made) for every trial throughout the experiment. During the "quiet" period, the spiral imaging gradients were turned off to decrease the audio noise so that participants' vocal responses could be recorded, while the slice selective RF pulse was kept on, in order to preserve the magnetization effects throughout the scan. According to Birn et al. (1999), the hemodynamic response signal onsets at 2–5 seconds and typically peaks at 5 to 6 seconds. Our interleaved method measured the hemodynamic response well within the five second window after subjects made their responses. To confirm that the data were not influenced by a partial magnetization effect during the 2 TR "loud" acquisition periods, we calculated the percent signal change in the cerebellum HVI region of interest (Table 2, centroid coordinate −8, −72, −20) and plotted the signal ratio between the first and second TR during the loud period across all runs. If there was a global signal shift from the first to the second TR, we should expect that

the signals at the first TR would be constantly higher than those at the second TR. Our results (supplemental figure 1, supplemental table 1), however, showed that the signal ratio within every repetition across all seven runs was very close to 100% and was mostly within a ± 1% range of difference. Based on these results, we believe that our data were not biased by the magnetization effect.

To maintain consistency between the manual and vocal conditions, the gradients were off in the manual conditions in the same way as in the vocal conditions. Using this method, a total of 140 images within 600 seconds were acquired for each run (Figure 3).

High resolution anatomical images were also acquired using a T1- weighted gradient echo pulse sequence with the following parameters TR/TE/FA/FOV 300ms/5ms/90/220mm and a voxel size of 1mm × 1mm × 1.2mm.

Data Analysis

Behavior—We acquired reaction time (RT) and accuracy measures for the manual conditions using E-prime software. To measure the RTs in the vocal conditions, a customized Matlab program was first used to record all of the vocal responses during the experiments. After the auditory data were recorded, another customized Matlab script was used to mark the RT for each response. RT was defined as the time between stimulus appearance on the screen and the beginning of the sound signal. Accuracy was checked by playing back the auditory signals.

Median RTs for each condition within each run were computed separately for correct random and fixed trials. The median RT, instead of the mean RT, was used because it is less biased in with the face of relatively large variability. Then, the *RT difference* (random – fixed trials) was calculated by taking the difference between random and fixed trials for the same learning runs. Two-way ANOVA (3 fitting models (linear, exponential, power fitting) × 4 conditions (spatial-manual, spatial-vocal, symbolic-manual, symbolic-vocal)) on goodness-of-fit parameters (adjusted-R square, sum square of error (SSE) and root mean square of error (rmse)) was performed to select the best fitting model for the learning data (Supplemental Table S2). Then, we conducted a linear regression analysis on RT difference with learning run (i.e. runs 2 – 6) as a continuous variable, and stimulus (spatial vs. symbolic) and response (manual vs. vocal) as categorical variables to assess learning and condition differences.

fMRI—The first three volumes per run were discarded to allow for signal equilibration. FSL MCFLIRT was used for motion correction and the Brain Extraction Tool (BET) was used to strip the skull from the images. Participants who showed head motion more than 3 mm were excluded from the analysis. Statistical Parametric Mapping version 5 (SPM5) was used for subsequent fMRI analyses. We first performed slice timing correction across functional images, and co-registered the structural images with the mean functional image. Then, the structural image was normalized to MNI (Montreal Neurological Institute) space. The acquired normalization parameters were then applied to the functional images. Functional images were spatially smoothed using a full width at half-maximum 8mm Gaussian smoothing kernel.

We used a hypothesis-driven ROI approach to probe the functional activation. Nine regions were selected based on the relevant literature (listed in Table 1). We used a non-linear transformation (Duncan et al., 2000;Calder et al., 2001) to transform the coordinates from Talairach to MNI space. Then, we created spherical ROIs (r=15mm) using these activation coordinates as the center (Figure 5). Comparisons between spatial and symbolic conditions (pooling across response mode) during the learning runs (runs 2–6) were focused on DLPFC, cerebellum lobule HVI, crus I & II, pre-SMA and premotor areas. We concatenated all learning runs for this contrast. In order to clarify the role of DLPFC for spatial mapping and the role of cerebellum for symbolic mapping rather than sequence learning, the same comparison across

all runs (runs 1–7 concatenated) were performed. ROIs in SMA, sensorimotor cortex, putamen and inferior parietal cortex were selected to test sequence learning in general. To assess the learning effect over time, linear coefficients were applied to all learning runs. For example, 2, 1, 0, −1, −2 were applied to the five learning runs to examine the brain regions that were more active early in learning while the contrast −2, −1, 0, 1, 2 was used to determine the brain regions that were increasingly active later in learning.

An ANOVA analysis was performed within the ROIs to examine the interactions between stimulus presentation (spatial, symbolic) and learning runs (runs 2–6) and another ANOVA (2 response (manual, vocal) and 5 learning runs was performed to explore the interaction between response and learning. Both experimental conditions and learning runs were treated as within-subject factors with the same pooling procedures described above. For example, when examining the stimulus by run interaction, we pooled the two spatial and the two symbolic conditions together. The five learning runs were assigned linear coefficients. When areas of significant activation were observed, we extracted the signal to examine the patterns of interactions.

Each of the above models was first run on individual participants' data. Head motion parameters were treated as nuisance covariates in the models. Then, second level SPM random effects analyses were used for group analyses. All results within these ROIs were evaluated at a family-wise error corrected (FWE, Chen 1998) $P \leq 0.05$ level with a spatial threshold of ten contiguous voxels ($3.44 \times 3.44 \times 3mm \times 10$ voxels). Significant areas of activation were then localized with motor cortical areas identified as in Picard and Strick (1996) and Mayka et al. (2006). Precent signal changes between responses and the fixation period within significant activation regions were calculated to examine their relationship with behavioral performance (e.g. learning magnitude).

Besides the main ROI analyses, we also explored the effects among stimulus, response and learning runs using whole brain analysis with a relatively liberal threshold (uncorrected $P < 0.005$ level with a spatial threshold of ten contiguous voxels). These analyses were exploratory and the results are presented in the supplemental materials.

Results

Behavioral findings

Figure 4A illustrates the RT differences between sequence and random elements across all 5 learning runs. To determine which fitting model was the best to quantify the learning pattern across runs, we performed goodness-of-fit analyses for linear fitting, power fitting and exponential fitting. Two-way ANOVA (3 fitting models (linear, exponential, power fitting) $\times$ 4 conditions (spatial-manual, spatial-vocal, symbolic-manual, symbolic-vocal)) on adjusted-R square, SSE and rmse revealed a significant model effect on adjust-R ($F_{(2, 104)} = 3.94$, $P < 0.05$, $\eta^2 = 0.19$). Post-hoc pair-wise comparisons suggested that the linear model was significantly better than the exponential ($P < 0.05$) and power fits ($P < 0.01$). No difference was found between exponential and power fitting (Supplemental Table S2). Thus, a linear regression analysis was used to measure learning. Results showed a significant run effect ($F_{(4, 260)} = 43.19$, $P < 0.01$, $\eta^2 = 0.21$). None of the interactions (all $P > 0.05$), stimulus ($F_{(1,260)} = 0.79$, $P = 0.78$) or response main effects ($F_{(1, 260)} = 2.60$, $P = 0.11$) were significant, suggesting that learning occurred equally well regardless of the experimental conditions.

Then, we performed a repeated measures ANOVA (2 trial types (random vs. fixed) $\times$ 5 learning runs (run 2–6) $\times$ 4 conditions (spatial-manual, spatial-vocal, symbolic-manual, symbolic-vocal)) on RT. A significant trial type $\times$ run ($F_{(4, 260)} = 43.18$, $P < 0.01$, $\eta^2 = 0.40$) interaction and main effects of trial type ($F_{(1, 260)} = 744.03$, $P < 0.01$, $\eta^2 = 0.74$) and condition ($F_{(3, 260)}$

= 90.90, $P < 0.01$, $\eta^2 = 0.51$) were found. LSD post-hoc analysis on trial type × run revealed that RT for the random trials did not change over the learning runs (all pair-wise comparison: $P > 0.05$) while RT for the fixed trials got faster over the learning (i.e. RT for run 5, 6 was significantly faster than run 2, both $P < 0.05$). A significant main effect on trial type reflected faster RT for the fixed trials than the random trials in general. Pairwise comparisons (LSD corrected) for the condition main effect showed that the spatial-vocal and symbolic-vocal conditions had similar reaction times ($P > 0.05$) whereas the two manual conditions had longer reaction times, with symbolic-manual the longest among all conditions (all $P < 0.01$). Figure 4B illustrates the RT for random trials across the four experimental conditions.

Another repeated measures ANOVA (2 trial types × 5 learning runs × 4 conditions) on accuracy revealed the trial type × condition ($F_{(3, 260)} = 2.85$, $P < 0.05$, $\eta^2 = 0.32$) interaction and main effects of trial type ($F_{(1, 260)} = 4.34$, $P < 0.05$, $\eta^2 = 0.01$) and condition ($F_{(3, 260)} = 8.72$, $P < 0.01$, $\eta^2 = 0.09$). Post hoc analysis revealed that the accuracy for the symbolic-manual condition (M = ± STD = 89.8% ± 6.9%, 90.8% ± 7.2% for random and fixed trials respectively) was lower than all other conditions (M ± STD = 96.0% ± 5.2, 94.3% ± 7.8, 95.2% ± 5.5 for random trials in spatial-manual, spatial-vocal, symbolic-vocal; M ± STD = 94.9% ± 5.6, 93.0% ± 7.1, 94.0% ± 7.4 for fixed trials in those conditions, respectively, Figure 4C & 4D). The interaction between trial type and condition was due to the accuracy difference between random and fixed trials in the symbolic-manual condition (−0.2%) and the difference between the two trial types in the spatial-vocal condition (+1.3%).

fMRI results

Spatial versus symbolic conditions (ROIs: DLPFC, cerebellum HVI, crusI & II, preSMA, and premotor areas)—To determine the differences between using spatial and symbolic stimuli during sequence learning, we performed main effect comparisons between the two spatially cued conditions and the two symbolically cued conditions during learning (runs 2–6). We found that the left cerebellum lobule HVI showed stronger activation for the symbolically cued conditions than the spatially cued ones (Figure 6A). In addition, the percent signal change within the active region was significantly correlated with the RT difference between fixed and random trials in the symbolic conditions at the completion of sequence learning ($r = .54$, $P < .05$, Figure 6B). This effect was not evident for the spatial RT difference ($P > .05$). Figure 6C illustrates the percent signal change in both the spatial and symbolic conditions across all learning (runs 2 – 6) and random runs (runs 1 & 7). These results further support that cerebellum HVI activation was specific for symbolic sequence learning. No significant activation differences were found in the other ROIs for the spatial versus symbolic contrasts.

In order to clarify the role of the DLPFC, we performed the same ROI analyses between spatial and symbolic conditions across all runs (runs 1–7). Although it has been suggested that DLPFC plays a general role in spatial mapping rather than a specific function for sequence learning (Schwarb & Schumacher, 2009), we did not find greater activation for the spatial versus symbolic conditions in the DLPFC.

All learning versus all random (ROIs: SMA, putamen, IPL and sensorimotor areas)—Areas that were activated across the whole time course of learning regardless of condition play a role in learning the more abstract features of the sequence. Among the four ROIs, we found that only the left premotor area (BA6) had significantly higher activation in the learning runs than in the random runs. Additional results from the whole brain analyses (using a $P < .005$ uncorrected threshold) are listed in supplemental table S3.

Increased/decreased activation from early to late runs (ROIs: SMA, putamen, IPL and sensorimotor areas)—Among these ROIs, only the left inferior parietal lobule (BA 40) showed increasing activation across learning that survived the FWE corrected threshold (Table 2, Figure 7). No areas decreased activation from early to late runs in these ROIs. The results from the whole brain analysis are presented in supplemental table S4. We also did contrasts between the first and the last blocks; no significant areas were found in any of the ROIs. Supplemental table S5 lists the areas showing significant activation differences in these contrasts from the whole brain analysis.

Additional ANOVA analyses—We did not find any significant interaction effects within any of the nine ROIs. To explore the combined effects among stimulus, response and learning runs, we performed within-subject ANOVA whole brain analyses using a relatively liberal threshold (P < .005 uncorrected). In general, our results (presented in Supplemental Table S6) were consistent with the existing literature (Grafton et al., 1998; Seidler et al., 2005).

Discussion

We examined the neural networks encoding spatial, symbolic and abstract motor sequence representations by selectively eliminating the spatial component of stimulus presentation (spatial versus symbolic cueing), response execution (manual versus vocal responses), or both. Using hypothesis-driven ROI analysis, we found that left cerebellum lobule HVI was selectively recruited for symbolic learning and the percent signal change in this region was correlated with learning magnitude under the symbolic conditions. In addition, the inferior parietal lobule exhibited increased activation during learning regardless of the condition. The left premotor area (BA6) had significantly higher activation in the learning runs than in the random runs. No other significant activations were found in other contrasts within the hypothesized ROIs.

Symbolic motor sequence representations

It has been recently argued that the cerebellum encodes symbolic representations of action (Balsters & Ramnani, 2008). In the sequence learning literature, studies have documented that individuals with cerebellar lesions are impaired at SRT learning when responses are symbolically cued (Gomez-Beldarrain et al., 1998; Molinari et al., 1997; Pascual-Leone et al., 1993; Shin & Ivry, 2003). Spencer & Ivry (2009) recently demonstrated that this deficit is mitigated when responses are directly cued. The dissociation between symbolic and direct cueing provides evidence that the cerebellum plays a role in maintaining sequential stimulus-response associations. We found higher activation in the cerebellum (lobule HVI) when movements were symbolically versus spatially cued. Furthermore, there was a significant correlation between the percent signal change in this region and the learning magnitude in the symbolic conditions. The current study provides novel neuroimaging evidence to support that the cerebellum indeed contributes to sequence learning by maintaining task-related S-R mapping when movements are symbolically cued. Moreover, our results suggest that the intact directly-cued learning in cerebellar patients reported by Spencer & Ivry (2009) was not due to a compensatory mechanism, and that lobule HVI is the central site engaged in maintaining S-R mapping. When the requirements for S-R mapping are minimum (i.e. movements are directly or spatially cued), learning a new sequence does not necessarily rely on cerebellar pathways (cf. Seidler et al., 2002).

Recently, there is an increasing body of evidence suggesting segregated cognitive and motor networks involving the cerebellum (Strick et al., 2009). The motor subdivision, including lobules V, VI, VIIB and VIII, primarily connects to the cortical motor regions via dorsal parts of the cerebellar dentate nucleus and the motor thalamus. The cognitive subdivision, including

lobule VIIA (principally Crus I and Crus II), projects to the prefrontal cortex via ventral parts of the cerebellar dentate nucleus and the mediodorsal thalamus (Kelly & Strick, 2003). Using resting-state functional magnetic resonance imaging, O'Reilly et al. (2010) have also identified two distinct zones in the human cerebellum. The primary sensorimotor zone (lobules V, VI and VIII) is functionally coupled to the motor, somatosensory, visual and auditory cortices. In contrast, the supramodal zone (lobules VIIa, Crus I and II) is functionally coupled to the prefrontal and posterior parietal cortex. Our finding that lobule VI activation was correlated with the sequence learning behavior while crus I and II were not is consistent with the view that lobule VI is involved in sensorimotor networks. It is interesting that the cerebellar lobule VI activity was localized to the left hemisphere given the ipsilateral projections of this structure. This pattern is consistent with another study which reported left lobule VI activation in relation to symbolic processing (Balsters & Ramnani, 2008). Further studies are needed to determine the functional lateralization of the cerebellum.

Spatial Motor Sequence Representations

It is somewhat surprising that we did not find preferential DLPFC engagement for spatially cued learning or spatial processing in general. Robertson et al. (2001) argued that the DLPFC plays a specific role in spatially cued sequence learning while Schwarb & Schumacher (2009) reported that DLPFC was selectively activated for spatial S-R selection but not necessarily sequence learning. Thus, we created an ROI using the coordinates from their study to examine the role of the DLPFC in spatially cued sequence learning. Contradictory to previous findings, we did not find DLPFC involvement in either spatial learning or spatial mapping.

DLPFC involvement in sequence learning has been proposed in the parallel network model (Hikosaka et al.,1999). Hikosaka et al., (1999) have suggested that sequences are learned in spatial coordinates in a network involving the frontoparietal cortices (including DLPFC) and the associative regions of the basal ganglia and cerebellum, while simultaneous motor learning activates the motor cortical areas. Additionally, spatial sequences are thought to be acquired explicitly and quickly, whereas motor sequences are thought to be acquired implicitly and slowly. Thus, for the trial-and-error sequence learning task employed by this group (Hikosaka et al., 1999, 2002), it is typical to find early learning activations in the lateral prefrontal cortex. Interestingly, using a similar paradigm, Bapi et al. (2006) did not find DLPFC activation during spatially presented sequence learning. When the S-R mapping was altered by rotating the response keypad 180 degrees, they found an activated network including the cerebellum and hippocampus. Since our current study employed the ASRT paradigm, the lack of DLPFC activation during the learning runs may suggest that the role of DLPFC in spatial sequence learning may relate to explicit awareness. However, we did observe a response × run interaction in the left medial frontal gyrus (BA 9), albeit at a more liberal threshold of $P < .005$ uncorrected. This region exhibited increasing activation across the time course of learning in the manual conditions. This result supports previous findings showing DLPFC engagement during implicit sequence learning (Ashe et al., 2005; Keele et al., 2003).

Abstract Motor Sequence Representations

The IPL (BA 40) has been proposed to encode an abstract representation during implicit motor sequence learning. For example, Grafton et al. (1998) found engagement of the left IPL (BA 40) when participants changed their motor responses from finger to arm movements, supporting a role for this structure in housing abstract (i.e. effector-independent) sequence representations. In the current study, we observed significant increased activation during learning in the left IPL regardless of the condition. However, our whole brain ANOVA analyses revealed some additional patterns of activation. Left IPL increased activation across the time course of learning in the spatial conditions and decreased activation in the symbolic conditions,

while right IPL exhibited increasing activation in the manual conditions and decreasing activation in the vocal conditions. Thus, these results do not explicitly support the notion that IPL codes just the abstract features of action sequences.

The IPL (BA 40) is proposed to be a functionally and anatomically heterogeneous region that is related to multiple aspects of sensory processing and sensorimotor integration. Clower et al. (2001) has reported that IPL receives outputs from distinct subcortical systems, including the superior colliculus, hippocampus, and cerebellum. Each distinct system targets a separate subregion of the IPL, supporting the functional heterogeneity of this structure. Recently, Hattori et al. (2009) have reported that the left dorsal IPL is functionally connected to the left ventral premotor cortex, the supplementary motor area (SMA) and the right cerebellar cortex, whereas the left ventral IPL is functionally coupled to the left dorsolateral prefrontal cortex and pre-SMA. Our results from the whole brain analysis also support this functional differentiation between the left and right IPL. That is, it appears that the left IPL may encode spatial sequence representations while the right IPL, as well as other areas, may represent learning in an effector-specific fashion for manual responses.

Caveats

In the current study, we found that learning occurred equally well regardless of the experimental manipulation of spatial and symbolic information at stimulus presentation or response execution. This result differs from our recent behavioral study (Bo & Seidler 2010) in which we used the same task conditions. Our previous study revealed smaller learning effects when responses were non-spatial regardless of the mode of stimulus presentation. Similarly, Koch and Hoffman (2000) found less learning when either the responses or the stimuli were non-spatial. This discrepancy is likely due to our use of a within-subjects design in the current study, which allows for potential transfer effects among the four experimental conditions. Here, we elected to use a within-subjects design to allow us to compare the neuroimaging results without contamination by individual and experimental differences. Our previous study (Bo & Seidler, 2010) as well as Koch & Hoffmann (2000), however, employed a between-subjects design which avoided the transfer confounds but required a much larger sample size (we tested close to 100 participants for experiment 1 in Bo & Seidler 2010). In the current study we used a counterbalanced presentation of the experimental conditions across participants to limit the impact of transfer of learning on the results.

One may also question the letter cues in the spatial conditions in our current design. We manipulated spatial versus symbolic processing requirements at both the stimulus presentation and response execution stages. Without adding letters in the spatial conditions, participants would not have known how to make the vocal response in the spatial-vocal condition. In order to maintain consistency across all four conditions, it was unavoidable that we added letters in the spatial conditions.

In addition, our behavioral data suggest that the symbolic-manual condition might be the most difficult. In line with this, Koch & Hoffman (2000) found that participants exhibited less learning in conditions where spatial information was minimized (i.e. symbolic stimuli & vocal responses). Thus, the presence of spatial sequences, regardless of whether it is in the stimuli or response, appears to be an important factor for learning in SRT tasks. Given the fact that the learning outcomes across the four conditions were not significantly different in the current study, and the accuracy level for the symbolic-manual condition was also around 90%, we think it unlikely that the difficulty of the symbolic-manual condition played a large role in the activation differences reported here. One may also argue that the manual conditions were more difficult than the vocal conditions because the RTs for the manual conditions were longer. These effects are likely not an issue for the main contrasts of interest because we pooled one manual and one vocal condition for the spatial versus symbolic contrast. Thus, even if the

reaction time differences influenced the fMRI results, it is not likely that this played a role in the main findings of the current study.

Based on the whole brain analyses (see Supplemental materials), we found that left inferior parietal lobule, superior temporal gyrus, right precental gyrus and lateral sulcus exhibited increasing activation across the time course of learning in the spatial conditions and decreasing activation in the symbolic conditions. In addition, left medial frontal gyrus, the lateral sulcus, and the right superior and inferior parietal lobules exhibited increasing activation in the manual conditions and decreasing activation in the vocal conditions across the time course of learning. These exploratory results suggest that although we did not find DLPFC to be engaged for spatial learning, there were broader networks that played such a role. However, given the exploratory nature of these analyses, further studies are necessary to investigate this proposal.

Finally, the current study investigated spatial & symbolic inputs and outputs in the visual-motor (manual, vocal) domain. However, there are many more categories of stimuli and responses that we did not investigate, such as auditory inputs, whole body outputs, etc. We speculate that although there may be some domain-specific activations, these more direct versus more symbolic cues and their neural bases may be the same across other modalities (i.e. not just visuospatial). In other words, regardless of the input and output modality, we posit that the cerebellum would be more engaged for symbolic learning. This remains to be tested.

Conclusions

The current study selectively eliminated the spatial component of stimulus presentation (spatial versus symbolic), response execution (manual versus vocal), or both to examine how S-R mapping demand affects the neural bases of motor sequence learning. Contradictory to our hypothesis, we did not find evidence that DLPFC is preferentially engaged for spatial processing during sequence learning. For the symbolically cued conditions, we found specific engagement of the cerebellum HVI, providing neuroimaging evidence to support the importance of the cerebellum in maintaining S-R mapping during sequence learning. In addition, IPL showed increasing activation across learning regardless of the condition, implicating its role in more abstract sequence learning processes. These results suggest differential networks that are flexibly engaged depending on the conditions of sequence learning.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Ashe J, Lungu OV, Basford AT, Lu X. Cortical control of motor sequences. Curr Opin in Neurobiol 2006;16:213–221.

Balsters JH, Ramnani N. Symbolic representations of action in the human cerebellum. NeuroImage 2008;43:388–398. [PubMed: 18692577]

Barch DM, Sabb FW, Carter CS, Braver TS, Noll DC, Cohen JD. Overt verbal responding during fMRI scanning: empirical investigations of problems and potential solutions. Neuroimage 1999;10:642–657. [PubMed: 10600410]

Bapi RS, Miyapuram KP, Graydon FX, Doya K. fMRI investigation of cortical and subcortical networks in the learning of abstract and effector-specific representations of motor sequences. NeuroImage 2006;32:714–727. [PubMed: 16798015]

Birn RM, Bandettini PA, Cox RW, Shaker R. Event-related fMRI of tasks involving brief motion. Hum Brain Mapp 1999;7:106–114. [PubMed: 9950068]

Bo J, Seidler RD. Spatial and symbolic sequence learning in young and older adults. Exp Brain Res 2010;201:837–851. [PubMed: 19949778]

Calder AJ, Lawrence AD, Young AW. Neuropsychology of fear and loathing. Nat Rev Neurosci 2001;2:352–363. [PubMed: 11331919]

Chen JJ. P-value adjustment for multiple binary endpoints. Commun Stat-Theory Methods 1998;27:2791–2806.

Clower DM, West RA, Lynch JC, Strick PL. The inferior parietal lobule is the target of output from the superior colliculus, hippocampus, and cerebellum. J Neurosci 2001;21:6283–6291. [PubMed: 11487651]

Daselaar SM, Rombouts SA, Veltman DJ, Raaijmakers JG, Jonker C. Similar network activated by young and old adults during the acquisition of a motor sequence. Neurobio of Aging 2003;24:1013–1019.

Duncan J, Seitz RJ, Kolodny J, Bor D, Herzog H, Ahmed A, et al. A neural basis for general intelligence. Science 2000;289:457–460. [PubMed: 10903207]

Friston KJ, Frith CD, Passinghan RE, Liddle PF, Frackowiak RSJ. Motor practice and neurophysiological adaption in the cerebellum: a positron emission tomography study. Proc R Soc Lond [Bio] 1992;248:223–228.

Fortin NJ, Agster KL, Eichenbaum HB. Critical role of the hippocampus in memory for sequences of events. Nat Neurosci 2002;5:458–462. [PubMed: 11976705]

Gomez-Beldarrain M, Garcia-Monco JC, Rubio B, Pascual-Leone A. Effect of focal cerebellar lesions on procedural learning in the serial reaction time task. Exp Brain Res 1998;120:25–30. [PubMed: 9628400]

Grafton ST, Hazeltine E, Ivry R. Functional Mapping of Sequence Learning in Normal Humans. J Cogn Neurosci 1995;7:497–510.

Grafton ST, Hazeltine E, Ivry RB. Abstract and effector-specific representations of motor sequences identified with PET. J Neurosci 1998;18:9420–9428. [PubMed: 9801380]

Hartman M, Knopman DS, Nissen MJ. Implicit Learning of New Verbal Associations. J Exp Psychol Learn Mem Cogn 1989;15:1070–1082. [PubMed: 2530307]

Hattori N, Shibasaki H, Wheaton L, Wu T, Matsuhashi M, Hallett M. Discrete parieto-frontal functional connectivity related to grasping. J Neurophysiol 2009;101:1267–1282. [PubMed: 19109459]

Hazeltine E, Grafton ST, Ivry R. Attention and stimulus characteristics determine the locus of motor-sequence encoding - A PET study. Brain 1997;120:123–140. [PubMed: 9055803]

Hikosaka O, Nakahara H, Rand MK, Sakai K, Lu X, Nakamura K, et al. Parallel neural networks for learning sequential procedures. Trends Neurosci 1999;22:464–471. [PubMed: 10481194]

Hikosaka O, Nakamura K, Sakai K, Nakahara H. Central mechanisms of motor skill learning. Current Opinion in Neurobiology 2002;12:217–222. [PubMed: 12015240]

Hikosaka O, Sakai K, Miyauchi S, Takino R, Sasaki Y, Putz B. Activation of human presupplementary motor area in learning of sequential procedures: a functional MRI study. J Neurophysiol 1996;76:617–621. [PubMed: 8836248]

Howard DV, Howard JH, Japikse K, DiYanni C, Thompson A, Somberg R. Implicit sequence learning: effects of level of structure, adult age, and extended practice. Psychol Aging 2004;19:79–92. [PubMed: 15065933]

Jenkins IH, Brooks DJ, Nixon PD, Frackowiak RSJ, Passingham RE. Motor Sequence Learning - A Study with Positron Emission Tomography. J Neurosci 1994;14:3775–3790. [PubMed: 8207487]

Keele SW, Ivry R, Mayr U, Hazeltine E, Heuer H. The cognitive and neural architecture of sequence representation. Psychological Review 2003;110:316–339. [PubMed: 12747526]

Kelly RM, Strick PL. Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. J Neurosci 2003;23:8432–8444. [PubMed: 12968006]

Koch I, Hoffmann J. The role of stimulus-based and response-based spatial information in sequence learning. J Exp Psychol Learn Mem Cogn 2000;26:863–882. [PubMed: 10946368]

Lee D. Behavioral context and coherent oscillations in the supplementary motor area. JNeurosci 2004;24:4453–4459. [PubMed: 15128859]

Lee D, Quessy S. Activity in the supplementary motor area related to learning and performance during a sequential visuomotor task. JNeurophysiol 2003;89:1039–1056. [PubMed: 12574479]

Liegeois F, Baldeweg T, Connelly A, Gadian DG, Mishkin M, Vargha-Khadem F. Language fMRI abnormalities associated with FOXP2 gene mutation. Nat Neurosci 2003;6:1230–1237. [PubMed: 14555953]

Mayka MA, Corcos DM, Leurgans SE, Vaillancourt DE. Three-dimensional locations and boundaries of motor and premotor cortices as defined by functional brain imaging: A meta-analysis. Neuroimage 2006;31:1453–1474. [PubMed: 16571375]

Molinari M, Leggio MG, Solida A, Ciorra R, Misciagna S, Silveri MC, et al. Cerebellum and procedural learning: evidence from focal cerebellar lesions. Brain 1997;120:1753–1762. [PubMed: 9365368]

Nissen MJ, Bullemer P. Attentional requirements of learning: evidence from performance measures. Cogn Psychol 1987;19:1–32.

Oldfield RC. Assessment and Analysis of Handedness - Edinburgh Inventory. Neuropsychologia 1971;9:97. [PubMed: 5146491]

O'Reilly JX, Beckmann CF, Tomassini V, Ramnani N, Johansen-Berg H. Distinct and overlapping functional zones in the cerebellum defined by resting state functional connectivity. Cereb Cortex 2010;20:953–965. [PubMed: 19684249]

Pascual-Leone A, Grafman J, Clark K, Stewart M, Massaquoi S, Lou JS, et al. Procedural learning in Parkinson's disease and cerebellar degeneration. Ann Neurol 1993;34:594–602. [PubMed: 8215247]

Picard N, Strick PL. Motor areas of the medial wall: A review of their location and functional activation. Cereb Cortex 1996;6:342–353. [PubMed: 8670662]

Posner MI, Dehaene S. Attentional Networks. Trends Neurosci 1994;17:75–79. [PubMed: 7512772]

Rao SM, Binder JR, Bandettini PA, Hammeke TA, Yetkin FZ, et al. Regional cerebral blood flow durig volitional expiration in man: a comparison with volitional inspiration. J Physiol (Lond) 1993;461:85–101. [PubMed: 8350282]

Robertson EM, Tormos JM, Maeda F, Pascual-Leone A. The role of the dorsolateral prefrontal cortex during sequence learning is specific for spatial information. Cereb Cortex 2001;11:628–635. [PubMed: 11415965]

Schendan HE, Searl MM, Melrose RJ, Stern CE. An fMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. Neuron 2003;37:1013–1025. [PubMed: 12670429]

Schwarb H, Schumacher EH. Neural evidence of a role for spatial response selection in the learning of spatial sequences. Brain Res 2009;1247:114–125. [PubMed: 18976640]

Seidler RD, Purushotham A, Kim SG, Ugurbil K, Willingham D, Ashe J. Cerebellum activation associated with performance change but not motor learning. Science 2002;296:2043–2046. [PubMed: 12065841]

Seidler RD, Purushotham A, Kim SG, Ugurbil K, Willingham D, Ashe J. Neural correlates of encoding and expression in implicit sequence learning. Exp Brain Res 2005;165:114–124. [PubMed: 15965762]

Shin JC, Ivry RB. Spatial and temporal sequence learning in patients with Parkinson's disease or cerebellar lesions. J Cogn Neurosci 2003;15:1232–1243. [PubMed: 14709239]

Spencer RM, Ivry RB. Sequence Learning is Preserved in Individuals with Cerebellar Degeneration when the Movements are Directly Cued. J Cogn Neurosci 2009;7:1302–1310. [PubMed: 18752399]

Strick PL, Dum RP, Fiez JA. Cerebellum and nonmotor function. Annu Rev Neurosci 2009;32:413–434. [PubMed: 19555291]

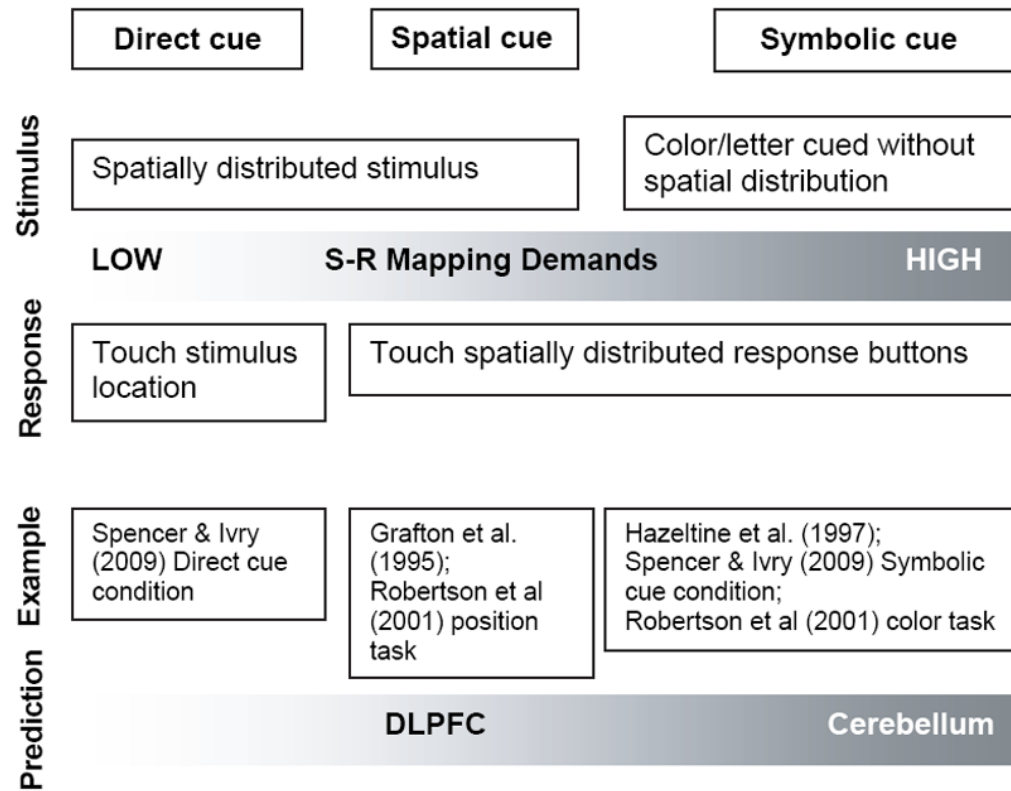Willingham DB. Implicit motor sequence learning is not purely perceptual. Mem Cognit 1999;27:561–572.

Figure 1.
The level of translation required between stimulus and response varies across conditions: 1) Directly cued (e.g. participants directly touch the stimulus on a computer), low S-R mapping demands; 2) Spatially cued (e.g. spatial location of the stimulus defines the response location on a button box), moderate S-R mapping demands; 3) Symbolically cued (e.g. letter or color of the stimuli defines the response), high S-R mapping demands.
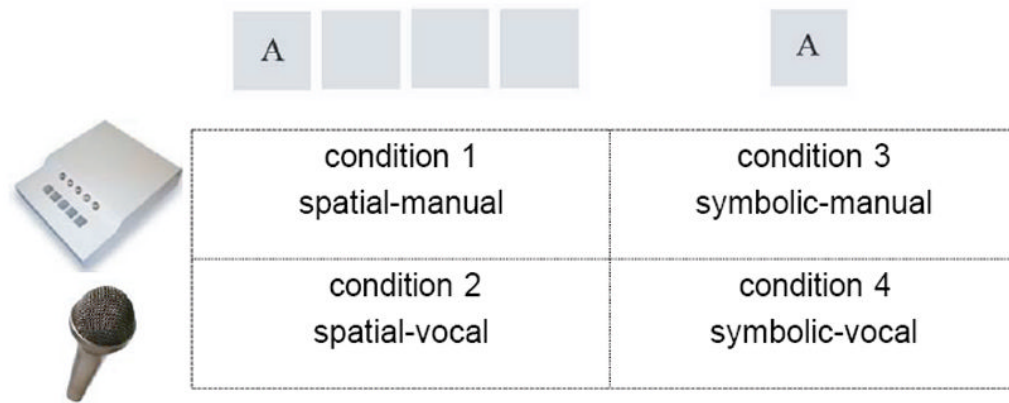
Figure 2.
Four experimental conditions of the current study: 1): spatial-manual: four visual stimulus boxes were presented on the screen with one of four letters appearing in one of the stimulus boxes. The positions of these four letters were fixed to the four stimulus boxes, i.e. "A" always appeared at the leftmost; "B" was at the second leftmost, "C" appeared at the second rightmost box and "D" was at the rightmost position. The letters A, B, C and D were mapped onto the index, middle, ring and little fingers of the right hand; 2): spatial-vocal: participants were asked to make vocal responses *ant, bear, cat, and dog* when seeing the stimulus letters, A, B, C, and D respectively. The stimulus presentation was identical to that of condition 1; 3): symbolic-manual: this condition involved only one centrally located stimulus box. The letter-finger mapping was the same as that in condition 1; 4): symbolic-vocal: participants made the same vocal responses as those in condition 2 in response to centrally located letter cues.
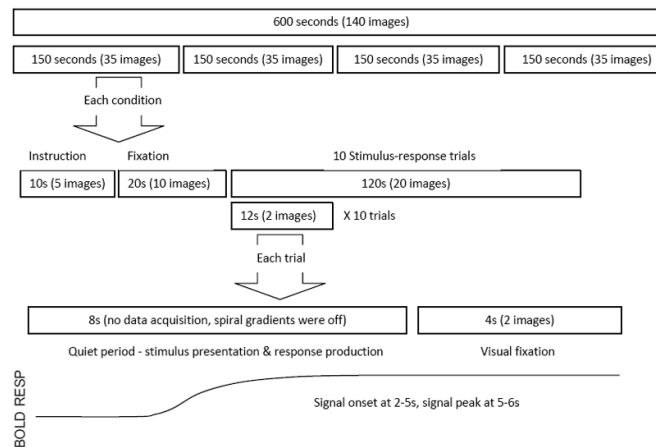
Figure 3.
Interleaved data collection method. There are four conditions randomly ordered within each run. Each condition began with a 10-second instruction, a 20-second resting period and 10 testing trials. Within each trial, the 8-element sequence was presented element by element every 1000 ms, followed by 4 seconds no response period. During the "quiet" 8-second stimulus-response period, the spiral imaging gradients were turned off to decrease the audio noise so that participants' vocal responses could be recorded, while the slice selective RF pulse was kept on in order to preserve the magnetization effects throughout the scans. Then, two functional volumes were triggered to acquire the subsequent 4 seconds of data. This delayed interleaved method detects the peak hemodynamic response which typically occurs 5 to 6 seconds after an event (Birn et al., 1999).
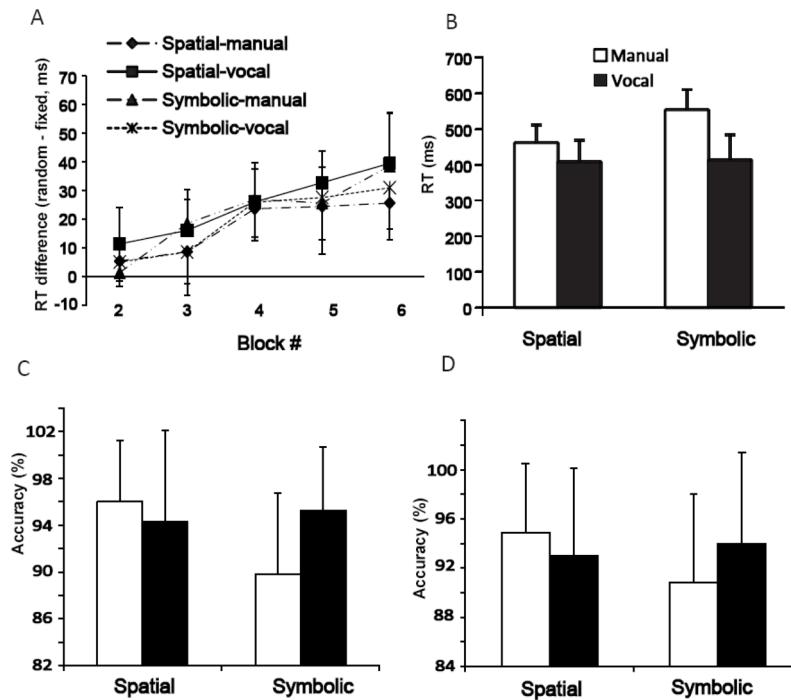
Figure 4.
A) The reaction time differences (random-fixed) in four conditions during 5 learning runs; B) The mean reaction time for random trials in four conditions; C) The mean accuracy for random trials in four conditions; D) The mean accuracy for fixed trials in four conditions. The error bars represent standard deviations.
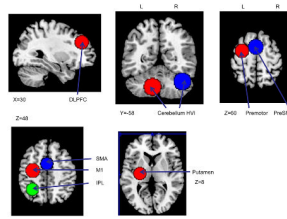
Figure 5.
Nice pre-defined ROIs: A) right DLPFC; B) left cerebellum HVI and right cerebellum HVI & crus I&II; C) left premotor and pre-SMA; D) left M1, SMA and IPL; E) left putamen
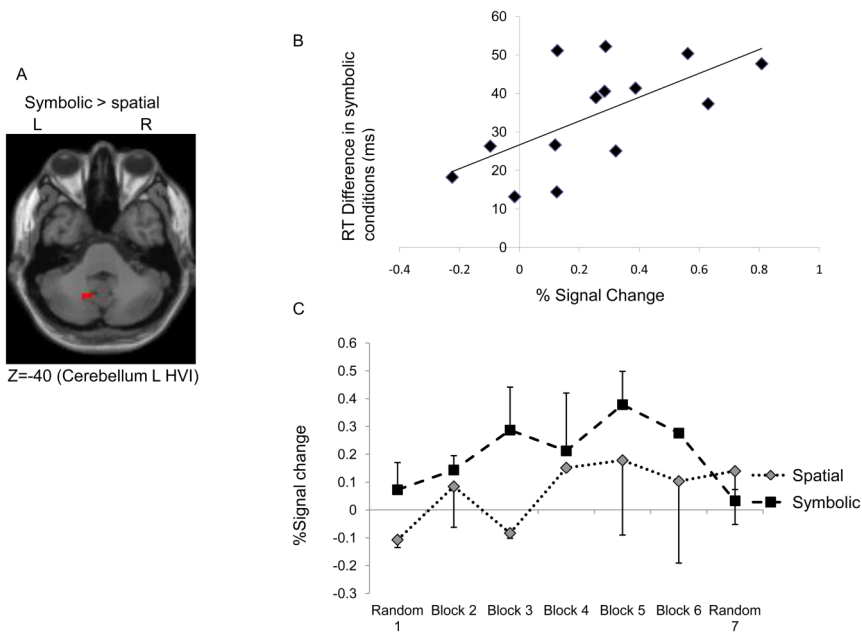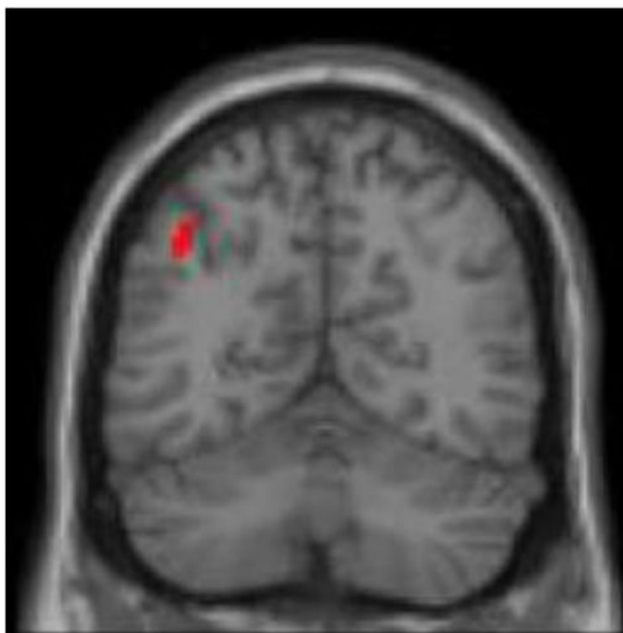
Figure 6.
A) Cerebellum HVI showed higher activation for the symbolically cued versus spatially cued conditions; B) Significant correlation between percent signal change and RT difference at run 6 in the symbolic conditions; C) Percent signal change (activation area illustrated at Figure 5A) in the spatial and symbolic conditions across all runs (runs 1–7). The error bars represent standard deviations.

Figure 7.
Left IPL showed increasing activation across all conditions during learning.

Table 1

Nine pre-defined regions of interest (ROI) to test three hypotheses: separated contrasts were performed under each subheading.

Anatomic location (BA)	Centroid coordinates (mm)			Reference
	X	Y	Z	
H1: cerebellum HVI and crus I & II would be preferentially engaged for symbolically cued learning				
Cerebellum (lobule HVI)	−8	−72	−20	Balsters & Ramnani 2008
Cerebellum (lobule HVI/crus I & II	39	−62	−31	Balsters & Ramnani 2008
L Medial superior frontal gyrus (pre-SMA, 6)	−2	6	62	Balsters & Ramnani 2008
L Superior frontal gyrus (pre-motor, 6)	−30	−2	66	
H2: DLPFC would be specifically engaged for spatially cued learning				
R Dorsal Prefrontal cortex (9)	36	39	31	Schwarb & Schumacher 2009
H3: sensorimotor cortex, SMA, putamen and inferior parietal lobule (IPL) would be activated for sequence learning across all experimental conditions				
L Precentral Gyrus (M1, 4)	−30	−13	45	Seidler et al., 2005
L Medial Frontal Gyrus (SMA, 6)	−1	1	57	Grafton et al. 1995
L Putamen	−27	−15	12	Grafton et al. 1995
L inferior parietal lobule (40)	−31	−52	40	Grafton et al. 1998

Table 2

Significant Activation within ROIs for different comparisons (FWE, $p \leq .05$)

Anatomic location (BA)	Talariach coordinates (mm)			Cluster size # of voxels	Z
	X	Y	Z		
Symbolic > Spatial					
Cerebellum L HVI	−12	−62	−40	48	3.21
Increased from early to late learning					
Parietal Areas					
L Inferior parietal lobule (40)	−42	−60	44	55	3.71
Abstract learning (sequence > random)					
Frontal Areas					
L Superior frontal gyrus (pre-motor, 6)	−20	−12	64	35	3.03