Supplemental Information Table of Contents

Supplemental Figures

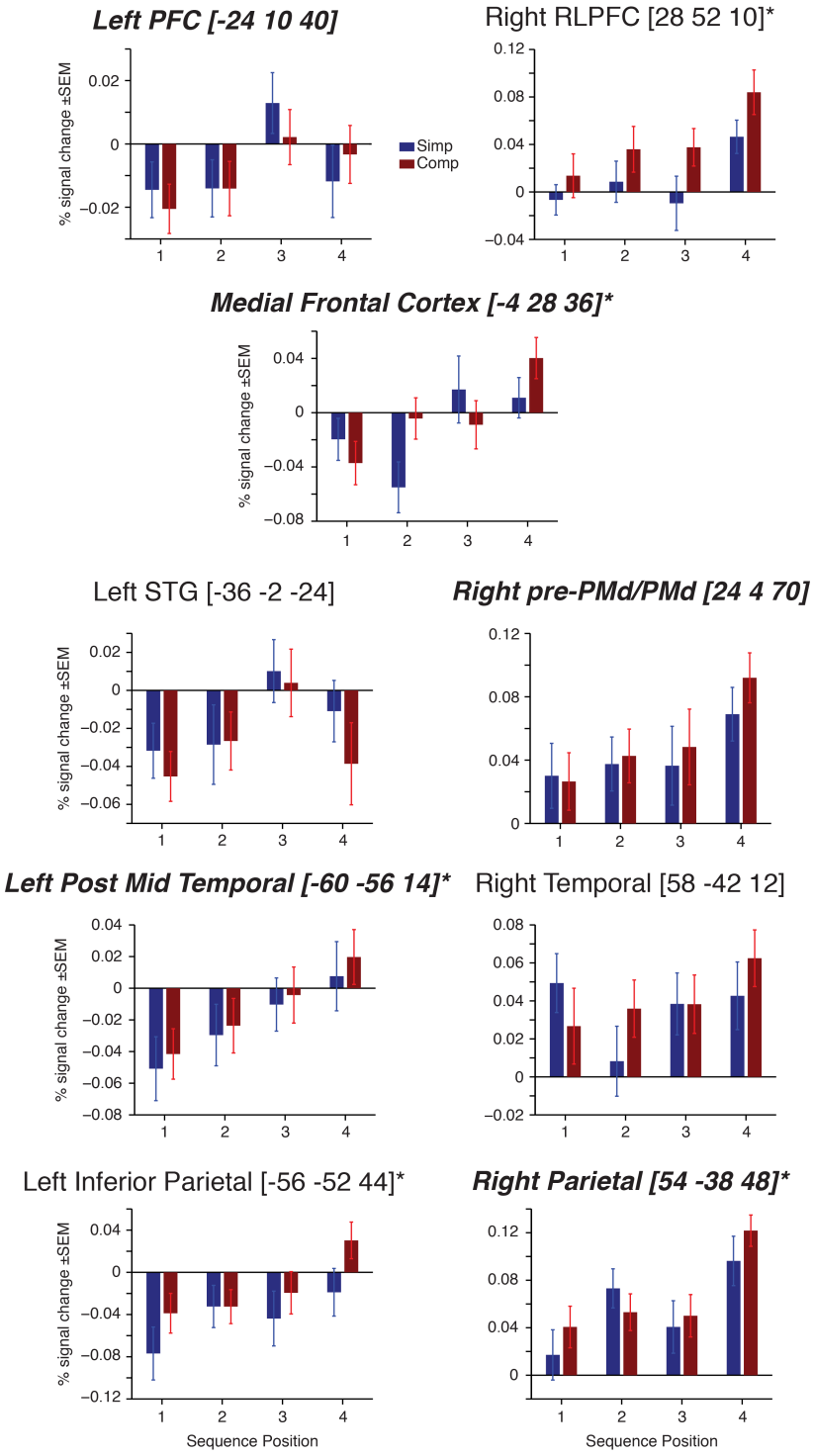

Figure S1, Related to Figure 2 Parametric Ramp > Baseline contrast peak ROIs. ***Bold and italic*** indicates clusters that overlapped with the Unique Ramp > Baseline contrast. Asterisk indicates significant main effect of Position. Only three regions; the medial frontal, parietal, and lateral anterior temporal cortices; both overlap with the Unique Ramp > Baseline contrast and have a significant main effect of Position ($F_{3,81}$'s > 2.9, P's < 0.05). All ROIs were spheres with an 8 mm radius around the peak coordinates listed.

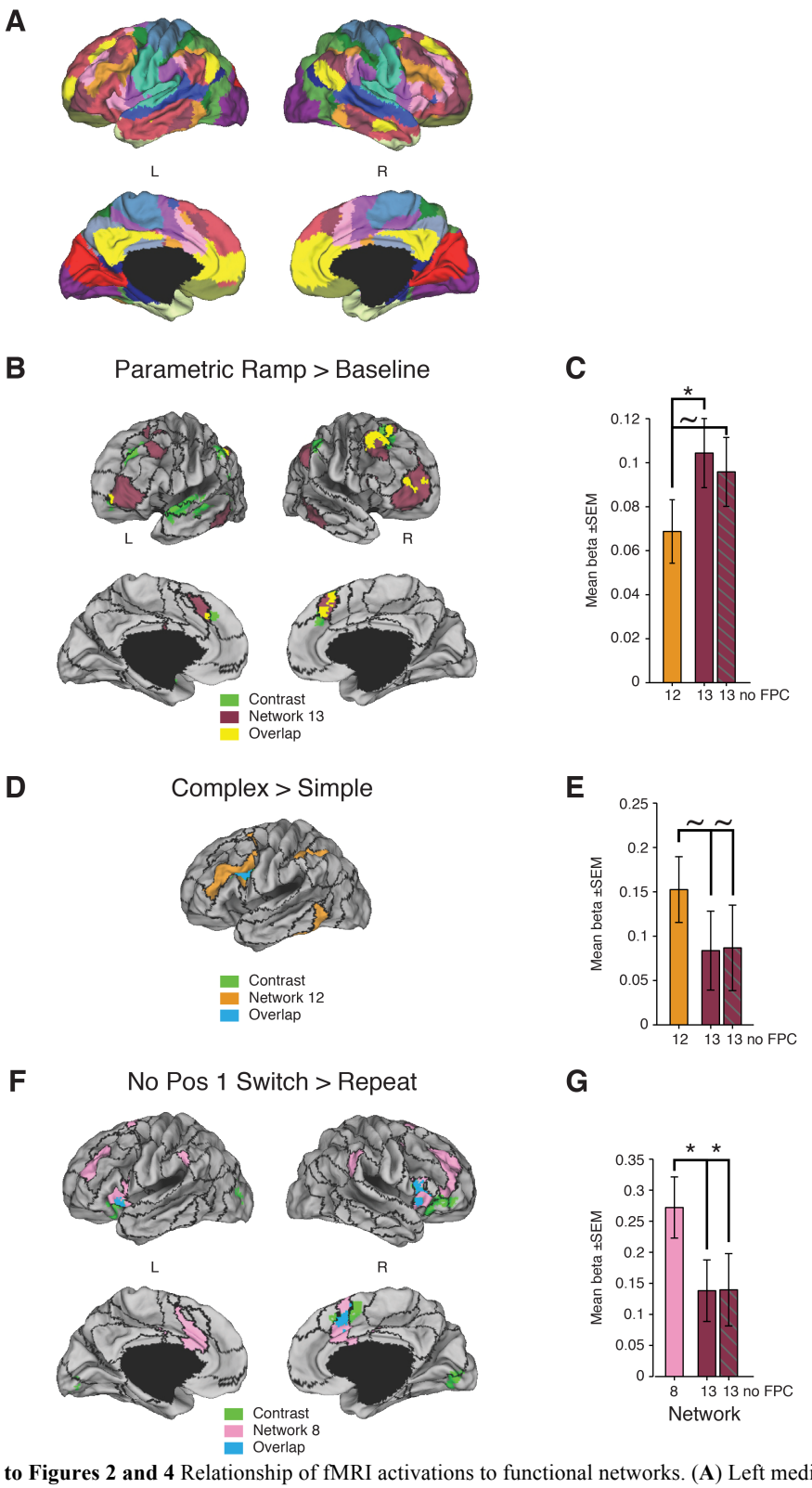

Figure S2, Related to Figures 2 and 4 Relationship of fMRI activations to functional networks. (**A**) Left medial and lateral views, right medial and lateral views of the 17-network parcellation with colors as in (Yeo et al., 2011). (**B**) Black outlines show borders of the 17-network parcellation. Parametric Ramp > Baseline contrast as in **Figure 2B**. (**C**) Mean beta values for Parametric Ramp > Baseline contrast from network 12 and 13 ROIs. A "*" indicates significance at the $P < 0.05$ level and "~" indicates marginal reliability (see SI text). (**D,E**) Same as in (B,C) but for the Complex > Simple contrast as in **Figure 2D**. (**F,G**) Same as in (B,C) but for the No Position 1 Switch > Repeat contrast shown in **Figure 4C**.

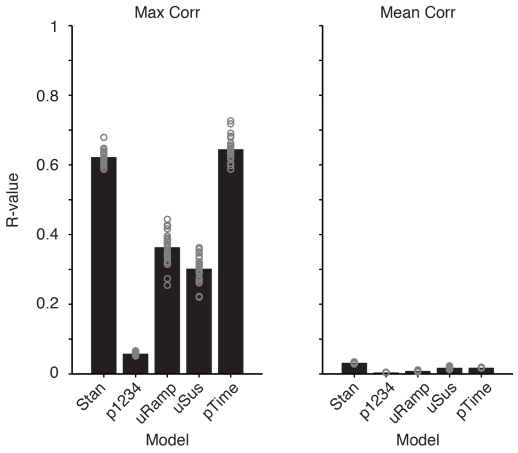

Figure S3, Related to Figure 4 Correlation values among regressors. Maximum (left) and mean (right) correlation values across subjects for each of the five models utilized. Bars indicate the mean across subjects and gray circles indicate values from individual subjects. Note that the values presented reflect the full matrix of correlations among all non-nuisance regressors, not solely those regressors that are entered into specific contrasts. The maximum values in the Standard Onsets model are due to correlations between regressors for adjacent positions in the sequence, e.g. Simple Position 1 and Simple Position 2. However, none of the contrasts presented here directly compare adjacent Positions.

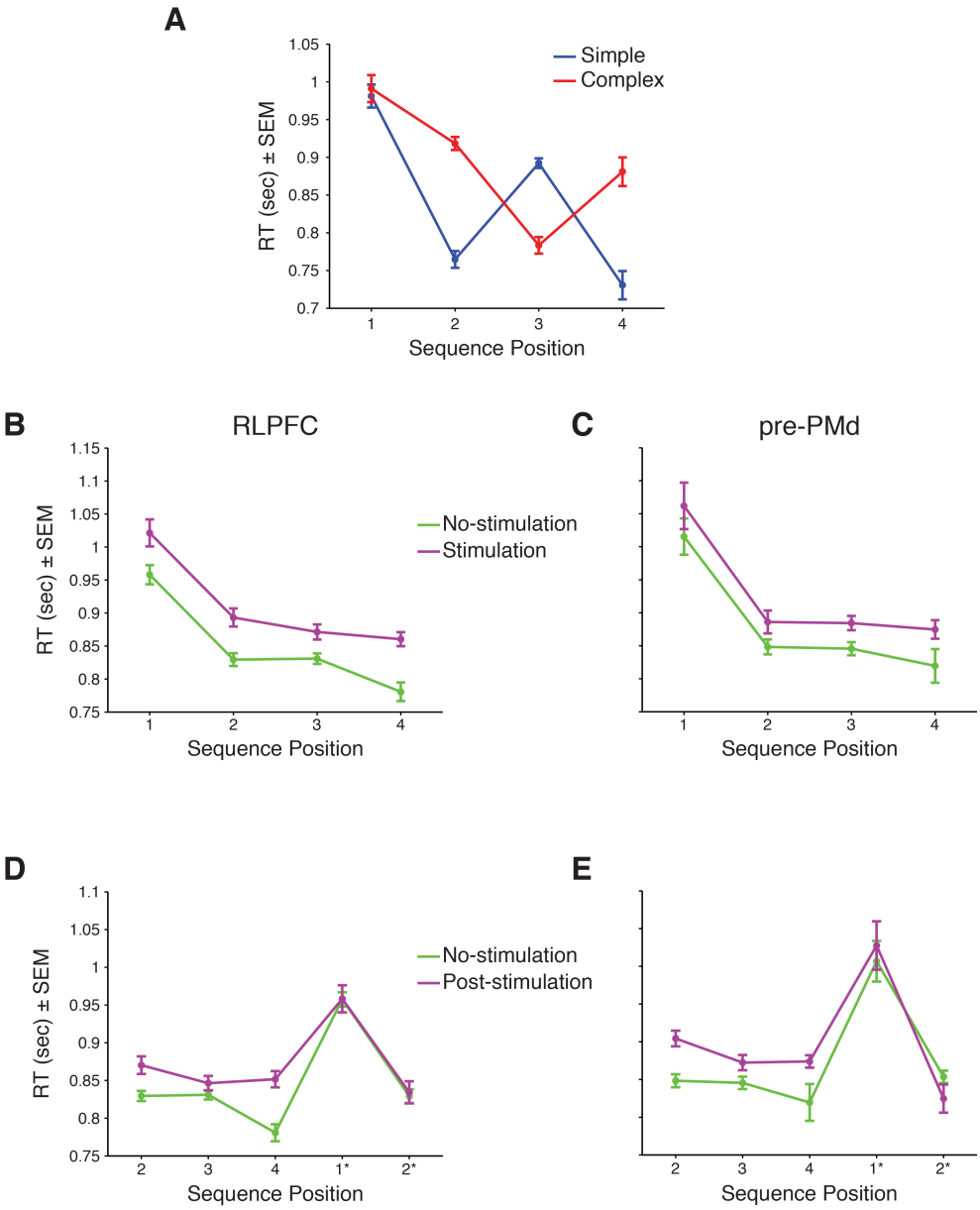

Figure S4, Related to Figure 5 RT results in TMS experiment. (**A**) RTs on no-stimulation trials combined across groups. (**B,C**) RTs including both Sequence Complexities in (**B**) RLPFC and (**C**) pre-PMd stimulation groups. (**D,E**) Post-stimulation RTs for (**D**) RLPFC and (**E**) pre-PMd stimulation groups. Positions 1* and 2* are in the subsequent trial. Post-stimulation trials had stimulation on one of the three previous trials, e.g. Position 3 post-stimulation contains trials where stimulation was delivered on either Position 1 or 2 of that sequence. Position 2* post-stimulation is made up of trials where stimulation was delivered at Position 3 or 4 of the *previous* sequence.

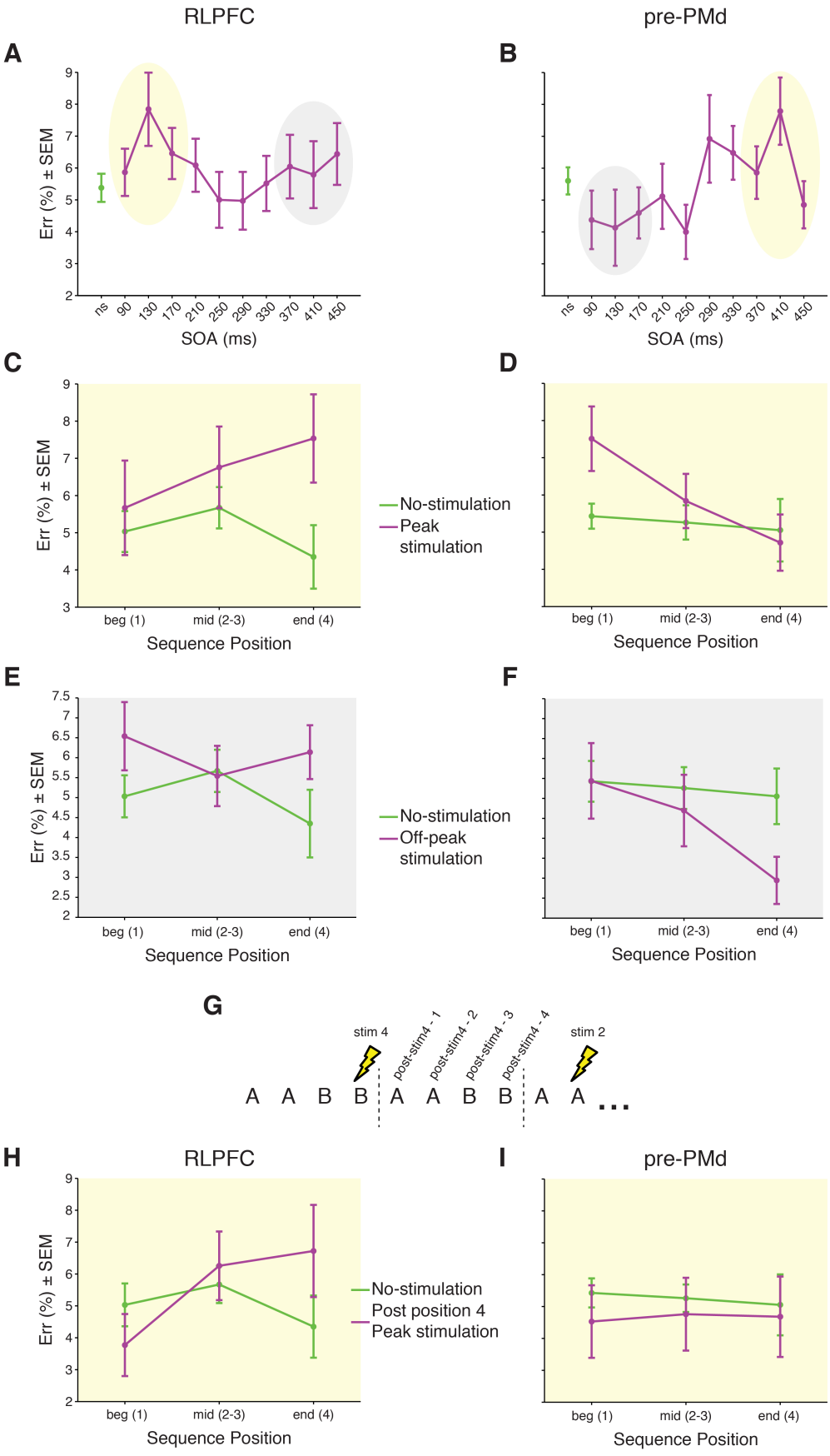
RLPFC
pre-PMd
A
B
C
D
E
F
G
H
I
Err (%) ± SEM
SOA (ms)
ns
90
130
170
210
250
290
330
370
410
450
Sequence Position
beg (1)
mid (2-3)
end (4)
No-stimulation
Peak stimulation
Off-peak stimulation
Post position 4
Peak stimulation
stim 4
stim 2
post-stim4 - 1
post-stim4 - 2
post-stim4 - 3
post-stim4 - 4
A A B B A A B B A A ...

Figure S5, Related to Figure 5 Mean ER in TMS experiment. (**A,B**) Same data as **Figure 5B** for (**A**) RLPFC **Figure 5C** for (**B**) pre-PMd. (**C,D**) Stimulation trials at the peak SOA ± 1 bin over the beginning (Position 1), middle (Positions 2 and 3), and end (Position 4) of each sequence for (C) RLPFC (peak SOA = 130 ms) and (**D**) pre-PMd (peak SOA = 410 ms). (**E,F**) Stimulation trials at the *off*-peak SOA ± 1 bin for (**E**) RLPFC (*off*-peak SOA = 410 ms) and (**F**) pre-PMd (*off*-peak SOA = 130 ms). (**G**) Schematic of trials where stimulation was delivered in the *preceding* sequence at Position 4. There was no stimulation delivered on post-position 4-stimulation trials themselves. (**H,I**) Post-position 4-stimulation trials at the peak SOA for (**H**) RLPFC and (**I**) pre-PMd.

Supplemental Tables

Table S1, Related to Figure 2, Activation coordinates

Contrast Location	BA	Extent (voxels)	*x*	*y*	*z*	Peak *t*-val.
Parametric Ramp > Baseline						
Left RLPFC	10	179	-28	58	-4	4.82
Right RLPFC	10/46	306	28	52	10	4.15
Bilateral Medial Frontal Cortex	6/32	321	-4	28	36	4.78
Left PFC	6/9	427	-24	10	40	5.23
Right Frontal		1,434				
pre-PMd/PMd	6		24	4	70	5.73
SMA/pre-SMA	6/8		10	22	52	4.23
Left STG & Temporal Pole		828				
Temporal Pole	22		-52	6	-10	4.77
STG	28/47		-36	-2	-24	5.3
Mid Temporal	21		-52	-24	-4	5.01
Right Temporal	22	191	58	-42	12	5.06
Left Posterior Mid Temporal	22/39	255	-60	-56	14	5.44
Right Parietal	40	545	54	-38	48	5.05
Left Inferior Parietal	40	183	-56	-52	44	4.49
Complex > Simple						
Mid-lateral PFC		391				
IFS	44		-58	12	30	4.35
pre-PMd	44		-48	4	28	4.35
PM	6		-48	-6	26	4.47
No 1 Switch > Repeat						
Right IFG	45/47	1,008	38	24	0	6.80
Left IFG	45/47	239	-30	20	-16	5.10
Right SMA/pre-SMA/Cingulate	6	417	10	12	28	4.81
Right Occipital	18	321	14	-76	-14	5.16
Left Occipital	18	336	-26	-86	8	5.21
Position 2 & 3 Switch > Repeat						
Right IFG	45/47	448	38	24	0	5.25
Right Occipital	19	189	26	-68	-16	4.88

Clusters reliable at $P < 0.05$ corrected. Coordinates are the center of mass in MNI.

Table S2, RT model regressors, Related to Experimental Procedures

Regressor	Mean Beta	t_{27}	P
Intercept	-0.078	-0.97	0.34
Run*	-0.097	-2.1	< 0.05
Block trial num.*	0.01	8.3	< 0.0001
Overall trial num.	0.00047	0.97	0.34
Sequence Complexity: Complex*	0.13	5.2	< 0.0001
Task Switch*	0.33	10	< 0.0001
Position 2*	-0.48	-14	< 0.0001
Position 3*	-0.43	-11	< 0.0001
Position 4*	-0.48	-15	< 0.0001
Previous ITI duration*	0.064	6.9	< 0.0001
Previous trial RT [AR(1)]*	0.11	6.5	< 0.0001
Response Incongruent*	0.21	7.8	< 0.0001
Response Switch*	-0.06	-2.6	< 0.05
Sequence Dwell Time (stepwise)*	0.021	3.1	< 0.01

* Indicates $P < 0.05$.

Supplemental Experimental Procedures

Behavioral analysis

The first sequence of each sequence block was excluded from analysis (fMRI: 80 trials, TMS: 128 trials total per participant). In addition, trials with reaction times < 100 ms were excluded (mean < 0.2% of trials). Error rate calculations were performed on the remaining trials. Reaction time (RT) analysis excluded error trials. Moreover, beyond the trial-based trimming, we also wanted to ensure that RT and imaging analysis focused on trials during which participants maintained the correct position within each sequence. We set a threshold of 2 error trials within a 4-trial moving window to trigger the start of a period where the participant potentially "lost" the position within sequence, and ended the "lost" period when the participant performed 4 trials in a row correctly. A mean of 2% of trials were removed because they were "lost". We collapsed across the two variants of each sequence type (e.g. CCSS and SSCC of AABB) because there were no differences in the patterns of data. Differences were assessed using repeated measures analysis of variance (RM-ANOVA), and, where applicable, paired t-tests.

Linear model of reaction times. RTs from the fMRI experiment were modeled using a linear model in Matlab using a least-squares fit that included the following predictors: run, trial number within block, overall trial number, Sequence Complexity (Simple/Complex), Local Task Switch (Repeat/Switch), Position in sequence (1-4), duration of previous ITI, previous trial RT (akin to an AR(1) correction), response Congruency (Congruent/Incongruent), local Response Switch (Repeat/Switch), and Sequence Dwell Time. Sequence Dwell Time was calculated as the time elapsed since the onset of the stimulus at position one in that sequence and then z-scored. All RT variables (response variable and previous trial RT) were log normalized and z-scored.

fMRI procedure

A Siemens 3T Trio Tim MRI system with a 32-channel head coil was used for whole-brain imaging. First, a high-resolution T1 weighted multiecho MPRAGE anatomical image was collected for visualization [repetition time (TR), 2200 ms; echo time (TE), 1.54, 3.36, 5.18, 7.01 ms; flip angle, 7°; 144 sagittal slices; 1.2 × 1.2 × 1.2 mm]. Functional images were acquired using a fat-saturated gradient-echo echo-planar sequence (TR, 2 s; TE, 28 ms; flip angle, 90°; 38 interleaved axial slices; 3 × 3 × 3 mm). There were five functional runs with a mean of 229 images collected per run (trials proceeded after each response, so the total length of the run was variable due to individual differences in RT). After the functional runs, a T1 in-plane anatomical image was collected as a possible aid in preprocessing (TR, 350 ms; TE 2.5 ms; flip angle, 70°; 38 interleaved transversal slices; 1.5 × 1.5 × 3 mm). Head motion was restricted using padding that surrounded the head. Visual stimuli were projected onto a screen and viewed through a mirror attached to the head coil. Participants responded using an MR compatible four-button response pad (Mag Design and Engineering).

Jittered ITI times were optimized for the rapid, event-related design by randomly choosing (with replacement) approximately half the times as "Short" (0.25, 0.5, or 0.75 s) and half the times as "Long" (1-8 s in 0.25 s increments) such that the mean time for the block was ~2 s. The Short trials were to ensure a sufficient number of trials to evaluate the behavioral effects. For each block, 10,000 sequences of times were then generated and then the design was estimated using SPM8. Each design estimate was evaluated for the correlation between regressors in the Onsets Model (see below). The 6 estimated designs with the lowest maximum correlations were used as a pool of block time sequences that would be drawn from for each participant and counter balanced across participants. This process was repeated for the four different block lengths (24, 25, 26, and 27 total trials).

Experiment control scripts were programed using the Psychophysics Toolbox in Matlab (Mathworks) and were displayed using an Apple computer running Mac OSX.

fMRI data analysis

No participants were excluded from analyses. Preprocessing and analysis of fMRI data were performed using SPM8 (http://www.fil.ion.ucl.ac.uk/spm). First, data were examined for artifacts, outliers and excessive variance in the global signal (functions: tsdiffana, art_global, art_movie). All data collected were used for subsequent analysis. The data were then corrected for differences in slice acquisition timing by resampling slices in time to match the first slice. Images were corrected for motion using a rigid transformation (using B-spline interpolation). Images were then normalized to Montreal Neurological Institute (MNI) stereotaxic space using affine regularization, smoothed using an 8 mm full-width at half-maximum isotropic Gaussian kernel, and resampled into 2 x 2 x 2 mm voxels using trilinear interpolation.

Across all models, regressors were generated by convolving onsets and parametric functions with the canonical hemodynamic response function (HRF). Onset regressors (not ramp or sustain regressors) were also convolved with the temporal derivative of the HRF. Nuisance regressors were included to remove variance due to the first sequence in a block, error trials, trials that were removed for behavioral analyses (see Supplemental Behavioral analysis), the six motion parameters (translation and rotation), linear drift over the course of each run, block instructions, and sequence position questions.

Betas related to each regressor were estimated using a subject-specific fixed-effects model. Whole brain, linear contrasts estimated subject-specific effects. We entered these estimates into a second-level analysis that treated subject as a random effect. Statistical reliability was assessed using a one-sample t-test against a contrast value of zero at each voxel ($P < 0.001$).

Whole brain group voxel-wise effects were considered significant at $P < 0.05$ corrected for multiple comparisons using extent thresholds at the cluster level to yield family-wise error correction. We rendered group contrasts on an inflated MNI canonical brain using Caret (Van Essen et al., 2001).

We assessed the strength of activation in each network of the 17-network parcellations (Yeo et al., 2011) by extracting the mean beta values across all five runs within a large ROI of each network (Choi and Badre, 2012), yielding up to 17 total values for each participant. These data were then subjected to RM-ANOVA. The values for **Fig. S2C** were extracted from the Parametric sequence position model and values for **Fig. S2E** and **S2G** were extracted from the standard Onsets model.

The overlap between activations in the present study and in the previous literature was explored using Sleuth (Fox and Lancaster, 2002).

Supplemental Results

A. FMRI behavioral results

Participants performed well at the fMRI task overall (error rate = 4.2%), and the mean reaction time (RT) across all task types was 1.1 s. At the sequence level, RT for Complex sequences was slower across Positions (1,161 ms) than for Simple sequences (1,117 ms). There was a significant overall effect of local task switching, (Positions 2-4; $F_{1,27} = 106$, $P < 0.0001$), Sequence Position ($F_{3,81} = 50.1$, $P < 0.0001$), and Sequence Complexity ($F_{1,27} = 22.5$, $P < 0.0001$) with an interaction between Position and Sequence Complexity ($F_{3,81} = 53.5$, $P < 0.0001$). The significant Complexity × Position interaction was driven by the alternating local switching effects at Positions 2-4 within the sequence rather than a difference at Position 1 ($t_{27} =$

0.34, $P > 0.7$). Local switching effects in RT did not interact with Position or Sequence Complexity (F's$_{1,27} < 2.2$).

To assess the sequence initiation effects, we compared the RT at Position 1 to the RT at Position 3. This was to equate the local task switching, e.g. in the Complex sequence, Position 1 is a task repeat, and Position 3 is also a task repeat (**A**B**B**A). We found the RT at Position 1 was significantly greater than the RT at Position 3 for both the Simple ($t_{27} = 6.3$, $P < 0.0001$) and Complex ($t_{27} = 7.6$, $P < 0.0001$) sequences individually, and when combined ($t_{55} = 8.7$, $P < 0.0001$).

Error rates showed reliable effects of task switching ($F_{1,27} = 14.5$, $P < 0.001$), Sequence Complexity ($F_{1,27} = 8.39$, $P < 0.01$), and an interaction of Position and Sequence Complexity ($F_{3,81} = 8.07$, $P < 0.001$; **Figure 1F**), but no effects of Position ($F_{3,81} = 8.39$, $P > 0.1$) or sequence initiation (Position 1 > 3, $t_{55} = -0.38$, $P = 0.7$), consistent with prior work (Schneider and Logan, 2006).

For the event-related fMRI design, it was necessary to jitter the ITI timing (0.25s - 8s, mean 2s). As preparation time can affect task switching, we examined the effect of the jittered ITI on reaction times (RT). We divided trials into two approximately equal groups (see Methods): trials where the preceding ITI was 0.75s or less ("Short"), and trials where the preceding ITI was 1s or more ("Long").

There was an effect of previous ITI such that RTs were faster following Short ITIs (1,106 ms) than Long ITIs (1,158 ms; $F_{1,27} = 13.1$, $P < 0.001$). There was also an interaction of ITI with Sequence Position ($F_{3,81} = 13.7$, $P < 0.0001$). Complex sequences showed a longer RT at Position 1 than Simple sequences on Short trials (one-tailed $t_{27} = 1.78$, $P < 0.05$), again replicating prior work (Schneider and Logan, 2006). The elevated RT at Position 1 and the

enhancement of this effect for Complex sequences provides evidence that participants relied on a sequence level representation.

The other basic behavioral effects were obtainable at both Short and Long ITIs. The effects of local task switching and sequence-level control observed in all trials were found to be reliable in both the Short and Long ITI conditions individually (Pos 1 > 3 all conditions t_{27} > 3.62, $P < 0.001$. Short: Seq $F_{1,27} = 17.3$, $P < 0.001$; Pos $F_{3,81} = 65.6$, $P < 0.0001$; Seq × Pos $F_{3,81}$ = 27.8, $P < 0.0001$. Long: Seq $F_{1,27} = 8.4$, $P < 0.01$; Pos $F_{3,81} = 23.7$, $P < 0.0001$; Seq × Pos $F_{3,81}$ = 26.5, $P < 0.0001$. Switch > Repeat Positions 2-4 all conditions $F_{1,27} = 65$, $P < 0.0001$). This indicates that despite having an influence on behavior, the jittered ITI did not abolish the critical effects, even after longer preparation time. Nevertheless, preparation time did interact with sequence level effects, such that the sequence initiation effect was larger following a short ITI ($t_{27} = 2.4$, $P < 0.05$).

All effects on error rates were again obtainable when analysis was restricted to trials preceded by a Short or Long ITI (Short: Switching $F_{1,27} = 6.9$, $P < 0.01$; Seq $F_{1,27} = 3.8$, $P <$ 0.06; Pos $F_{3,81} = 1.3$, $P > 0.3$; Seq × Pos $F_{3,81} = 3.8$, $P < 0.05$. Long: Switching $F_{1,27} = 15$, $P <$ 0.0001; Seq $F_{1,27} = 7.2$, $P < 0.01$; Pos $F_{3,81} = 1.8$, $P > 0.2$; Seq × Pos $F_{3,81} = 5.3$, $P < 0.01$.).

B. RLPFC ramp controls

To distinguish the ramping pattern of activation across sequence Positions 1-4 (**Figure 3D**) in RLPFC from a uniform, sustained signal throughout the sequence (**Figure 3C**) and compensate for any shared variance between the regressors (**Figure S3**) we used stepwise regressions to identify the unique variance accounted for by each function. The model containing the Unique Ramp regressor produced significant activation in the RLPFC (**Figure 4A**), but the Unique Sustain regressor model did not (**Figure 4B**). Similarly, beta values in the RLPFC ROI

were not significantly greater than zero for the Sustain regressor ($t_{55} = 1.8$, $P > 0.05$), though they were for Ramp regressor ($t_{55} = 3.3$, $P < 0.01$). Thus, even when the shared variance between Sustain and Ramp regressors was removed, the activity in the RLPFC was explained only by ramping, and not significantly by sustained, activation.

Unfortunately, our design did not permit us to unambiguously distinguish an event-related parametric ramp (**Figure 3B**) from a smooth ramp (**Figure 3D**) signal in the BOLD response. Including these regressors in the same model is analogous to a "hybrid model" (Donaldson et al., 2001) wherein both transient and epoched signals are modeled on the same BOLD response in order to distinguish them. However, as was demonstrated by Visscher et al. (2003), when sustained and transient regressors compete for variance in the same model, error in the transiently modeled BOLD can be systematically misallocated to the epoch model, more than vice versa. Unfortunately, we did not build in the necessary features to avoid this bias in the present fMRI design, as we did not anticipate discovering the ramping function when we designed the experiment.

To examine the relationship among the regressors, we performed pairwise correlations of the following regressors across all participants: onsets at all positions, onset-based parametric ramp (**Figure 3B**), smooth ramp (**Figure 3D**), and sustain (**Figure 3C**). The onsets at all positions (All Onsets) are not a ramp, but an instantaneous regressor at each stimulus onset that is the same across all positions. The table below shows the mean R-value of the correlations for each pair of convolved regressors (± standard deviation).

	Parametric Ramp	Smooth Ramp	Sustain
All Onsets	0.67 (± 0.03)	0.51 (± 0.03)	0.57 (± 0.03)
Parametric Ramp	1	0.87 (± 0.01)	0.81 (± 0.02)
Smooth Ramp		1	0.88 (± 0.02)

C. FMRI correlates of local task-level control

To more directly examine the effects of local task level control, we contrasted trials where the task Switched within a sequence (e.g. A to B in A**AB**B) to those where the task Repeated (e.g. B to B in A**BB**A). So that sequence initiation would not contaminate task switching, we first contrasted task Switches with Repeats at Positions 2 through 4 (**Figure 4C**, **Table S1**). This contrast yielded activation in bilateral IFG, right SMA/pre-SMA, and bilateral occipital cortex. Next, to equate the number of switches and repeats contributed by each Sequence Complexity, we performed the contrast of Switch versus Repeat trials using only Positions 2 and 3. Here, we found similar activation in right IFG, and right occipital cortex (**Table S1**).

We confirmed the effect of local task control in right SMA/pre-SMA in an unbiased ROI (see Methods). In particular, SMA/pre-SMA showed a significant effect of Task Switching ($F_{1,27}$ = 11.5, $P < 0.01$) and Position ($F_{3,81}$ = 2.7; $P < 0.05$ **Figure 4D**), and no effect of Sequence Complexity ($F_{1,27}$ = 2.4; $P > 0.1$).

SMA/pre-SMA was distinguished from pre-PMd and RLPFC based on their sensitivity to local task switching. SMA/pre-SMA showed a significant ROI × Switch/Repeat interaction with pre-PMd ($F_{1,27}$ = 15, $P < 0.001$), as well as RLPFC (ROI × Switch/Repeat: $F_{1,27}$ = 7.15, $P <$ 0.01).

D. Relationship to functional networks

The whole brain analysis revealed activation outside of RLPFC and pre-PMd related to sequential control (e.g. **Figure 2B**). Though this activation was widespread and distributed, recent functional connectivity studies using resting state fMRI data has provided evidence of

widely distributed association networks that share correlated activation (Yeo et al., 2011). This correlated activity may be reflective of intrinsic network connectivity among these distributed regions. Thus, to assess whether the distributed regions activated here were members of the same functional network, we tested the overlap of our functional activation with neocortical networks defined from a recently published 17-network parcellation of resting state fMRI data (N=1,000) (Yeo et al., 2011). For each functional contrast, we computed the mean beta estimates from key frontal network definitions across participants.

The most anterior association network (**Figure S2A**, maroon network) contained the greatest values for the Parametric Ramp > Baseline contrast (**Figure S2B**). This network had reliably different beta values than a nearby network (**Figure S2A,** orange network) that contained the pre-PMd ROI representative of the sequence complexity effects ($t_{27} = 2.4$, $P < 0.05$, **Figure S2C**). Importantly, the network overlap was not only due to the activation of RLPFC, but also overlap within medial frontal cortex and parietal cortex (**Figure S2B**, yellow). Indeed, the maroon network had marginally greater mean beta values than the orange network even when the RLPFC subregion was removed from the network definition ($t_{27} = 2$, $P < 0.06$, **Figure S2C** hatched bar). In contrast to medial frontal and parietal cortex, the superior temporal cortex activation did not overlap with the temporal lobe component of the maroon network definition.

The greatest values for the Complex > Simple conditions were in the mid-lateral association network (**Figure S2D**, orange network). While the difference between the orange and maroon (with and without RLPFC) networks was marginal (t_{27}'s > 1.5, P's < 0.1), in the whole brain, the majority of activation in Complex > Simple fell in the mid-lateral fronto-parietal network (orange) in the frontal cortex (**Figure S2E**). Contrasted against the ramping effect, this yielded a reliable network by effect interaction ($F_{1,27} = 4.5$, $P < 0.05$).

The Switch > Repeat contrast showed a third pattern of activations across the networks, with the greatest beta values in network 8 (pink; **Figure S2F**). Again the significant activations for this contrast were mostly confined to this association network. There were reliable differences between the betas in the maroon network, both with and without the RLPFC, and the pink network (t_{27}'s > 2.3, P's < 0.05, **Figure S2G**). Contrasting the ramping effect preference in the maroon network and the switch preference for the pink network yielded reliable network by effect interactions ($F_{1,27}$'s > 5, P's < 0.05).

E. TMS behavioral results

The behavioral results in the TMS experiment on the no stimulation trials replicated all the basic behavioral effects found in the fMRI experiment (**Figure S4A**). There was no difference in the RTs on trials that did not receive stimulation in between the RLPFC and pre-PMd stimulation groups (Region: $F_{1,32}$ = .02, P > 0.8; Position × Region: $F_{3,96}$ = 0.06, P > 0.9).

Combining data from no-stim trials in the two regions, we replicated the behavioral results from the fMRI experiment in the TMS experiment. There was a significant main effect of Sequence Complexity ($F_{1,34}$ = 41, P = 0.0), Position ($F_{3,102}$ = 46, P = 0.0), and a significant Sequence Complexity × Position interaction ($F_{3,102}$ = 37, P = 0.0; **Figure S4A**). The TMS experiment also replicated the sequence initiation effects observed in the fMRI experiment when contrasting Position 1 with Position 3 (matched for switching and repeating in individual sequences). Position 1 showed significantly greater RTs than Position 3, regardless of Sequence Complexity (Simple: t_{34} = 8, P = 0.0; Complex: t_{34} = 6, P = 0.0; Combined: Simple: t_{69} = 9, P = 0.0). To preserve statistical power in this large design, for the remainder of the analyses we will combine Simple and Complex sequences.

F. TMS stimulation results

In both RLPFC and pre-PMd, RTs were significantly impacted by stimulation on both the trial that stimulation was delivered ($F_{1,17}$'s > 12, P's < 0.01, **Figure S4B,C**) and on trials that followed stimulation ($F_{1,17}$'s > 17, P's < 0.001), but were still within the same sequence. Importantly, this post-stimulation effect did not cross over into the subsequent sequence (Position × Post-stimulation interaction $F_{4,68}$'s > 3.5, P's < 0.01; **Figure S4D,E**). Though these RT results suggest the involvement of RLPFC and pre-PMd sequential control, we must be cautious about the interpretation. Given that both regions showed comparable effects on RT and a third control region, RMPFC, marginally showed these effects (Position × Post-stimulation $F_{4,56} > 2.1$, $P < 0.09$), we cannot be sure if the effect of TMS on RT is a function the stimulation of these specific areas (and so their functional role) or rather a generalized effect of the TMS stimulation (e.g., general irritation or distraction).

Although the RTs are numerically different, there were no reliable differences in the effects of stimulation on RT between RLPFC and pre-PMd in TMS1 (Region: $F_{1,32} = 0.06$, $P > 0.8$; Stimulation × Region: $F_{1,32} = 0.4$, $P > 0.5$; Position × Region: $F_{3,96} = 0.08$, $P > 0.9$; Stimulation × Position × Region $F_{3,96} = 0.4$, $P > 0.7$). We also found no reliable differences in the effects of stimulation on RT between RLPFC and pre-PMd in TMS2 (Region: $F_{1,28} = 1.5$, $P > 0.2$; Stimulation × Region: $F_{1,28} = 0.01$, $P > 0.9$; Position × Region: $F_{3,84} = 0.09$, $P > 0.9$; Stimulation × Position × Region $F_{3,84} = 1.1$, $P > 0.3$) or between RLPFC and RMPFC in TMS2 (Region: $F_{1,14} = 0.11$, $P > 0.7$; Stimulation × Region: $F_{1,14} = 2.4$, $P > 0.1$; Position × Region: $F_{3,42} = 2.4$, $P > 0.08$; Stimulation × Position × Region $F_{3,42} = 0.7$, $P > 0.5$). Therefore, with no reliable differences in the effect of stimulation on RT between RLPFC and pre-PMd in either TMS experiment, and also no reliable differences in the effect of stimulation between RLPFC

and RMPFC, we conclude that an account of the effects found in ER being due only to the amount of distraction participants experienced as a result of each pulse is not supported.

Unlike the RTs, in ER we found a dissociation between TMS to the RLPFC and pre-PMd (**Figure S5C,D**). This dissociation is further supported by a significant difference between the effects of stimulation at the last position in the sequence between RLPFC and pre-PMd ($F_{1,32}$ = 6.7, $P < 0.01$; **Figure 5D**). The effects of TMS were specific to the peak SOA, as the Stimulation (no-stim, 130 ms peak, 410 ms peak) × Region interaction was reliable ($F_{2,64} = 3.4$, $P < 0.05$) and the Position × Stim × Region interaction was not significant at non-peak SOA ($F_{2,64} = 0.7$, $P >$ 0.4; **Figure S5E,F**). Furthermore, the effects of stimulation on ER did not carry over into the following sequence. When we examined the effect on ER in the sequence immediately following stimulation at Position 4 (**Figure S5G-I**), the Position × Stimulation × Region interaction was not reliable ($F_{2,64} = 2.8$, $P > 0.05$).

G. TMS post-test questionnaire results

In TMS1 the mean amount of discomfort experienced by participants in the RLPFC group was 1.8, and in the pre-PMd group was 1.2. The two groups did not have significantly different ratings (Wilcoxon rank sum test, $z = 1.7$). In TMS2 neither the amount of distraction ($z = 0.97$, P > 0.3) nor the amount of discomfort ($z = -0.1$, $P > 0.9$) differed between RLPFC and pre-PMd, or between RLPFC and RMPFC (distraction: $z = -0.6$, $P > 0.5$; discomfort: $z = 0.74$, $P > 0.4$). As illustrated in **Figure 5F**, the area with the greatest distraction ratings, RMPFC, showed the smallest effects of stimulation. Similarly, pre-PMd showed numerically greater discomfort rating than RLPFC, and RLPFC exhibits greater effects later in the sequence.

To further support that the amount of distraction or discomfort had no effect on the observed effects of stimulation on error rates, when the amount of distraction and discomfort

were entered as covariates either independently or together, the crossover interactions between RLPFC and pre-PMd were unchanged (Position x Stimulation x Region: $F_{2,56}$'s > 3.4, P's < 0.04). The crossover interaction between RLPFC and RMPFC also remained unchanged with the same covariates (Position x Stimulation x Region: $F_{2,28}$'s > 3.4, P's < 0.05). These results again support the conclusion that there is a significant difference between stimulation delivered to RLPFC and pre-PMd and between RLPFC and RMPFC that is not due to differences in the amount of distraction or discomfort.

Supplemental Discussion

G. Relationship to prior stimulation studies

Only two previous studies have examined TMS stimulation to the RLPFC (Hanlon et al., 2013; De Pisapia et al., 2012), and of them, only one examined the behavioral consequences of such stimulation (De Pisapia et al., 2012). Participants were asked to maintain a numeral in working memory, and then perform five arithmetic operations. During one of the five operations, the remembered numeral would be required and thus be "integrated" from working memory. Repetitive TMS pulses were delivered at the third step. They found an increase in RT when TMS was delivered to the RLPFC after an integration step, and no effects on error rates. The authors speculate that the RLPFC increases activation up to the integration step, though they did not observe such a ramping pattern in their data. It is striking that we did observe this hypothesized dynamic in fMRI and in TMS. Though, we interpret this effect as relating to uncertainty rather than integration. But these ideas are not incommensurate, for example, it is possible that an integration step is required to resolve uncertainty about task and position in our task. Likewise, because the occurrence of the integration step was unpredictable in this prior study, the RLPFC's

involvement may have reflected the need to resolve uncertainty about when the integration would take place across the multiple steps of their task.

H. Prior studies of RMPFC in sequential control

In Yoshida and Ishii (2006), participants navigate through a virtual maze towards a goal. The paper uses an elegant hidden Markov model to predict the participants' sampling and updating of their spatial position in the virtual maze. They draw from this model an estimate of the participants' uncertainty about their position, the hidden current position (HCP) entropy. They found activity in the superior frontal gyrus and medial anterior frontal cortex that correlated with HCP entropy.

In one sense, the results of Yoshida and Ishii (2006) might complement the broadest conclusions of the present study, in as much as they focus on uncertainty resolution during a sequential task. However, there are two major distinctions that make the results of the current study a fundamental and novel addition that go well beyond Yoshida and Ishii (2006). First, we discovered new brain dynamics not previously observed by Yoshida and Ishii (2006). Specifically, we identify a ramping pattern of activation over the course of a task sequence in RLPFC and associated areas. This pattern is the primary novel discovery from the fMRI experiment that motivates our two subsequent TMS studies. Yoshida and Ishii (2006) did not observe this ramping dynamic in RLPFC or elsewhere. And, further, they do not even observe changes in RLPFC. Their effects are in medial anterior PFC, which is not activated by position in the present experiment and which our follow up TMS experiment now shows to be unnecessary for our task. Thus, strictly speaking we fail to replicate Yoshida and Ishii (2006), if uncertainty resolution is the whole story.

Of course, a more likely reason for these differences is that the task paradigms

themselves are different in fundamental ways. We asked participants to conduct a repeating sequence of tasks. Yoshida and Ishii (2006) studied spatial navigation. Even if one interprets their task in terms of sequential decision making, the spatial navigation task amounts to a sequence of spatio-motor responses rather than a sequence of tasks. As we describe in the Introduction of this paper, among the novel contributions of this experiment is that we challenge the assumption that motor-based and task-based sequences are the same. There have been very few studies that explicitly test task sequences. Most, like Yoshida and Ishii (2006), use motor-based sequences. Our study provides such a test, and indeed, we observe novel dynamics not seen previously in Yoshida and Ishii (2006) or elsewhere.

The Koechlin et al. (2002) study shares elements of the task paradigm with the current study; however, the results and conclusions are different between these studies. Koechlin et al. (2002) focused on sequence learning, whereas the present study was explicitly instructed, hence not requiring learning at all, and was focused instead on sequential control. In Koechlin et al. (2002), participants learned to respond to sequences of letter stimuli in either a motor sequence (color indicated the finger to move to respond), or a cognitive sequence (color indicated one of three tasks to perform on the current letter). These motor and cognitive sequences were learned during early scanning blocks and performed without cues (letters without colors) in later blocks. They found that the anterior medial frontal cortex was selectively activated in during cognitive sequence cue learning, whereas the ventral striatum was selectively activated during motor sequence cue learning. Further, they found lateral anterior prefrontal cortex was more engaged early in learning, when control processes were theoretically more necessary because the behaviors were less automatic.

Though again perhaps providing a complement to our results, the present study clearly reports new discoveries that lead to fundamental new conclusions about RLPFC and sequential

control beyond what was described in Koechlin et al. (2002). First, the within-sequence ramping pattern that is the focus of the present study was not reported by Koechlin et al. (2002). They only reported block-level activation changes, such as over the course of learning, and do not investigate dynamics within a sequence. Further, Koechlin et al. (2002) focused on effects within medial frontal cortex. We again note that we located our effects laterally in RLPFC, not in medial frontal cortex. And further, the follow up TMS study demonstrated that medial anterior PFC is not necessary for our task. Though again, this might be interpreted as failing to replicate the Koechlin et al. (2002) result, we reiterate that Koechlin et al. (2002) was focused on changes over the course of learning and we investigated changes within a task sequence.

The above discussion of Yoshida and Ishii (2006) and Koechlin et al. (2002) provides evidence that the present fMRI findings go well beyond these prior reports. The observation of within-sequence ramping pattern in RLPFC has not only not been observed before, this temporal dynamic provides new insights into the mechanisms of hierarchical control that go well beyond the general prior association of anterior frontal cortex with higher order tasks. Rather, we provide important new clues regarding the dynamics by which RLPFC region comes online to influence sequential control.

Supplemental References

Choi, E.-Y., and Badre, D. (2012). Resting-state functional connectivity evidence for asymmetric rostro-to-caudal prefronto-striatal connectivity. In Annual Meeting of the Society for Neuroscience,.

Donaldson, D.I., Petersen, S.E., Ollinger, J.M., and Buckner, R.L. (2001). Dissociating state and item components of recognition memory using fMRI. Neuroimage *13*, 129–142.

Van Essen, D.C., Drury, H.A., Dickson, J., Harwell, J., Hanlon, D., and Anderson, C.H. (2001). An integrated software suite for surface-based analyses of cerebral cortex. J. Am. Med. Inform. Assoc. *8*, 443–459.

Fox, P.T., and Lancaster, J.L. (2002). Opinion: Mapping context and content: the BrainMap model. Nat. Rev. Neurosci. *3*, 319–321.

Hanlon, C.A., Canterberry, M., Taylor, J.J., Devries, W., Li, X., Brown, T.R., and George, M.S. (2013). Probing the Frontostriatal Loops Involved in Executive and Limbic Processing via Interleaved TMS and Functional MRI at Two Prefrontal Locations: A Pilot Study. PLoS One *8*, e67917.

Koechlin, E., Danek, A., Burnod, Y., and Grafman, J. (2002). Medial prefrontal and subcortical mechanisms underlying the acquisition of motor and cognitive action sequences in humans. Neuron *35*, 371–381.

De Pisapia, N., Sandrini, M., Braver, T.S., and Cattaneo, L. (2012). Integration in working memory: a magnetic stimulation study on the role of left anterior prefrontal cortex. PLoS One *7*, e43731.

Schneider, D.W., and Logan, G.D. (2006). Hierarchical control of cognitive processes: switching tasks in sequences. J. Exp. Psychol. Gen. *135*, 623–640.

Visscher, K.M., Miezin, F.M., Kelly, J.E., Buckner, R.L., Donaldson, D.I., McAvoy, M.P., Bhalodia, V.M., and Petersen, S.E. (2003). Mixed blocked/event-related designs separate transient and sustained activity in fMRI. Neuroimage *19*, 1694–1708.

Yeo, B.T.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J. Neurophysiol. *106*, 1125–1165.

Yoshida, W., and Ishii, S. (2006). Resolution of Uncertainty in Prefrontal Cortex. Neuron *50*, 781–789.