Neural Substrates of Working Memory Updating

Gal Nir-Cohen[1], Yoav Kessler[1], and Tobias Egner[2]

Abstract

■ Working memory (WM) needs to protect current content from interference and simultaneously be amenable to rapid updating with newly relevant information. An influential model suggests these opposing requirements are met via a BG–thalamus gating mechanism that allows for selective updating of PFC WM representations. A large neuroimaging literature supports the general involvement of PFC, BG, and thalamus, as well as posterior parietal cortex, in WM. However, the specific functional contributions of these regions to key subprocesses of WM updating, namely, gate opening, content substitution, and gate closing, are still unknown, as common WM tasks conflate these processes. We therefore combined fMRI with the reference-back task, specifically designed to tease apart these subprocesses. Participants compared externally presented face stimuli to a reference face held in WM, while alternating between updating and maintaining this reference, resulting in opening versus closing the gate to WM. Gate opening and substitution processes were associated with strong BG, thalamic, and frontoparietal activation, but intriguingly, the same activity profile was observed for sensory cortex supporting task stimulus processing (i.e., the fusiform face area). In contrast, gate closing was not reliably associated with any of these regions. These findings provide new support for the involvement of the BG in gate opening, as suggested by the gating model, but qualify the model's assumptions by demonstrating that gate closing does not seem to depend on the BG and that gate opening also involves task-relevant sensory cortex. ■

INTRODUCTION

Working memory (WM) refers to our ability to temporarily hold information in mind, manipulate it, and update it in the service of goal-directed behavior (Oberauer et al., 2018; Cowan, 2017). Models of WM have long emphasized the tension between its maintenance and updating functions (Fallon, van der Schaaf, ter Huurne, & Cools, 2017; Badre, 2012; O'Reilly, 2006; Frank, Loughry, & O'Reilly, 2001; Miller & Cohen, 2001): Current WM content has to be shielded from interference by irrelevant information, while at the same time being amenable to updating when new goal-relevant information appears.

An influential neurocognitive theory addressing this dilemma is the PBWM (PFC, BG, and WM) network model, which postulates a selective input gate for WM (O'Reilly & Frank, 2006; Frank et al., 2001). Specifically, the model proposes a BG gating mechanism that separates perceptual input, represented in sensory cortex, from WM representations, maintained in (or via) dorsolateral PFC (dlPFC). By default, the gate is closed, thus enabling robust maintenance/shielding of WM content. However, in response to salient signals, like task-relevant stimuli or reward cues, the BG gate opens (based on phasic dopaminergic input from the midbrain), allowing for the inflow of new information into WM via a thalamus–PFC pathway.

There is a vast neuroimaging literature supporting the assumption that the dlPFC—usually in conjunction with the medial PFC (mPFC) and posterior parietal cortex (PPC)—contributes to the maintenance of WM content (e.g., D'Esposito & Postle, 2015; Nee et al., 2013; Feredoes, Heinen, Weiskopf, Ruff, & Driver, 2011; Roth, Serences, & Courtney, 2006). The same broad set of regions, often referred to as the frontoparietal network (FPN), has also been implicated in updating WM, as inferred from *n*-back (e.g., Owen, McMillan, Laird, & Bullmore, 2005), AX-CPT (e.g., Lopez-Garcia et al., 2016), and task-switching (e.g., Kim, Cilles, Johnson, & Gold, 2012) studies. These regions have therefore been referred as the core network of WM (Johnson et al., 2019; Harding, Yücel, Harrison, Pantelis, & Breakspear, 2015; Rottschy et al., 2012). A smaller set of studies has also provided evidence to support the involvement of the BG (Chatham & Badre, 2015; Murty et al., 2011; Cools, Sheridan, Jacobs, & D'Esposito, 2007) and/or the midbrain (D'Ardenne et al., 2012; Murty et al., 2011) in WM updating. More specifically, the involvement of the BG in gating goal-relevant information into WM—a key subcomponent of WM updating—has been implicated in several studies. For example, previous studies reported BG involvement during switching attention between objects (van Schouwenburg et al., 2014; Cools, Clark, & Robbins, 2004) and tasks (Leber, Turk-Browne, & Chun, 2008). Moreover, van Schouwenburg, den Ouden, and Cools (2010) found that the BG mediated the connectivity between PFC and visual cortex during attentional shifts, triggered by a bottom–up cue. Finally, McNab and Klingberg (2008) demonstrated that activity in both PFC and BG preceded the selection of relevant information for WM maintenance and that this activity was associated with individual

[1]Ben-Gurion University of the Negev, [2]Duke University

Journal of Cognitive Neuroscience 32:12, pp. 2285–2302
https://doi.org/10.1162/jocn_a_01625

differences in WM capacity, resonating with the notion that capacity is related to filtering (i.e., gating) ability (Vogel, McCollough, & Machizawa, 2005).

Although there is broad agreement on the FPN, BG, and thalamus being the key players in WM input gating, the mapping of these regions to the processes underlying WM gating and updating is presently unclear. These processes include opening the gate to allow information into WM, modifying the relevant items while removing outdated information, and returning to a closed-gate, perceptually shielded state when updating is complete. The reason for a lack of such process-specific brain mapping in the prior literature is mainly because of task impurity. For instance, standard *n*-back, AX-CPT, and task-switching protocols conflate item encoding, updating, substitution, and other processes and do not provide a means to differentiate gate opening, gate closing, substitution, and item removal processes (discussed in Lewis-Peacock, Kessler, & Oberauer, 2018; Kessler, 2017; Rac-Lubashevsky & Kessler, 2016a, 2016b; Ecker, Lewandowsky, Oberauer, & Chee, 2010).

The goal of this study was therefore to examine potential functional specialization in the WM network with respect to above-described subprocesses involved in WM updating. To this end, we paired fMRI with the recently developed "reference-back" task (Rac-Lubashevsky & Kessler, 2016a, 2016b, 2018), which has been shown to successfully disentangle processing costs associated with four key WM updating operations: (1) opening the gate to WM; (2) updating information in WM, which may either take the form of reinforcing current content or (3) substituting old with new information; and (4) closing the gate to enable robust maintenance of the newly updated information. In addition, we examined the neural correlates of being in an "updating mode" (see Kessler & Oberauer, 2014). Unlike the processes described above, the updating mode refers to the state of the gate to WM—whether it is open for new input or not.

By interrogating neural responses in the BG, FPN, and thalamus, as well as in visual regions with known sensitivity to our task stimuli (see below), we observed distinct patterns of neural substrates supporting the different WM updating processes. Whereas dlPFC, BG, and thalamus were preferentially involved in the gate opening process, parietal cortex also contributed to this process but additionally displayed a stronger contribution to substitution. In contrast, these regions were not involved in gate closing.

METHODS

Participants

To mitigate the dangers of false-positive and false-negative findings, we based our sample size on effect size estimation (Button et al., 2013). Specifically, a recent meta-analysis of a large fMRI data set indicated a moderate effect size for WM task contrasts (Poldrack et al., 2017). For a desired power of 0.8 to detect this size of effect in within-subject contrasts, under assumption of a conservative (low) level of correlation between paired observations ($r = .3$), we aimed for a minimal sample size of $n = 45$ (based on GPower; Erdfelder, Faul, & Buchner, 1996).

Sixty-one healthy students from Ben-Gurion University of the Negev participated in the experiment in exchange for monetary compensation. Thirteen participants were excluded from the analysis because of technical problems with the MRI during the scan (6), extensive head movements (2) or a low accuracy rate (5; <80%). The final sample included 48 participants (29 women; age: $M = 25.5$ years, $SD = 2$ years). All participants were right-handed and reported normal or corrected-to-normal vision. None of the participants had any history of neurological or psychiatric problems. The experiment was approved by the Helsinki committee of the Soroka Medical Center, Beer Sheva, Israel.

Stimuli

The reference-back task used eight face images of neutral facial expression (four male, four female; two faces per block) from the FEI faces database (fei.edu.br/~cet/facedatabase.html). The faces were displayed inside blue (RGB values: 0, 0, 255) and red (RGB values: 255, 0, 0) colored frames (see Figure 1A). The faces' diameter was approximately 180 pixels (4.76 cm, subtending a visual angle of 2.7° from a 100-cm viewing distance). The frame's dimensions were 380 × 380 pixels (10 × 10 cm), subtending a visual angle of 5.7°. The use of face stimuli (in combination with an independent "localizer" scan) enabled us to identify the fusiform face area (FFA; Kanwisher, McDermott, & Chun, 1997) to assess neural task stimulus processing in visual cortex as a function of WM updating operations. The localizer task employed grayscale images of famous familiar faces, unfamiliar faces, buildings, objects, and scrambled objects. Those images were presented within an elliptical shape (14.5 × 8 cm, 8.3° × 4.5°) against a black background. All stimuli were projected on a screen at the back of the scanner bore and viewed via a mirror affixed to the head coil.

Procedure

The scanning session took about an hour, in the following sequence: anatomical structural scan (10 min); 1-back task with face stimuli, serving as a face localizer task (10 min); and four 80-trial blocks of the reference-back task (30 min). The participants completed a behavioral practice session (two blocks) of the reference-back task a day or two before the experimental session in the scanner.

The Reference-Back Task

We employed the reference-back task (Rac-Lubashevsky & Kessler, 2016a, 2016b), which enabled us to disentangle WM updating subprocesses (i.e., gate opening, gate closing, and substitution). This task is based on the *n*-back task. In the standard *n*-back task, the participant is presented with

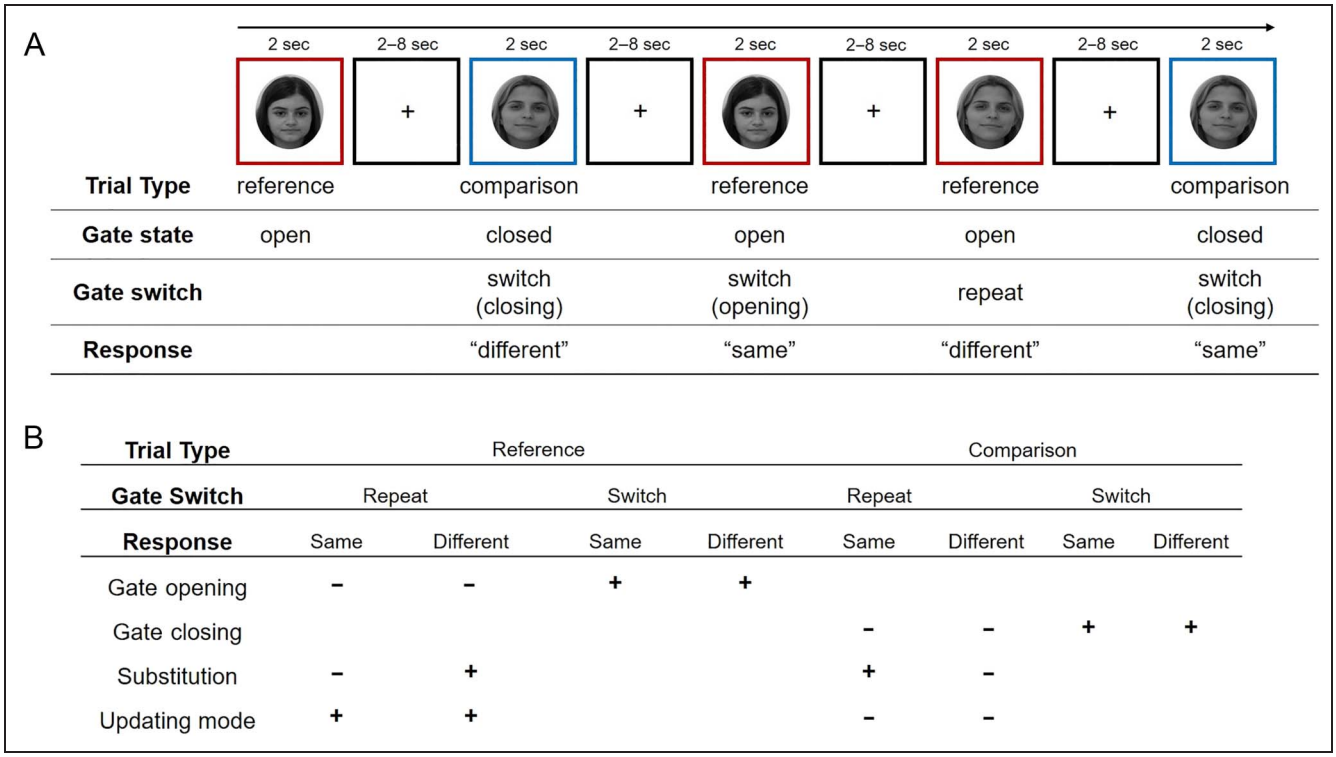

Figure 1. (A) An example sequence of reference (red frame) and comparison (blue frame) trials in the reference-back task (top), along with an illustration of the putative states of the WM gating process and responses required (bottom). (B) Contrast weights for defining distinct WM updating subprocesses.

a sequence of stimuli and is asked to decide whether the current stimulus is identical or not to the stimulus presented *N* trials before (Owen et al., 2005; Jonides et al., 1997). Because updating, including various subprocesses, takes places in each trial of the *N*-back task, it is difficult to isolate the subprocess and identify their distinctive neural markers. To overcome this limitation, the reference-back paradigm was developed.

This task is composed of two trial types: reference and comparison. Specifically, in each trial (see Figure 1A), a face stimulus was presented inside a red or blue frame, and the participant was required to indicate whether or not the stimulus was identical to the one presented in the most recent red frame. Trials involving a red frame are denoted "reference trials." In these trials, the participant must first compare the presented stimulus to the one held in WM, that is, the face that appeared in the previous red frame (making a same/different decision); the participant then has to update his or her WM with the stimulus that appears in the present trial, which serves as the reference for future trials. Trials involving a blue frame are denoted "comparison trials." Like in red frame trials, the participant is required to make a same/different decision between the currently presented face and the reference held in WM; however, unlike in reference trials, WM does not have to be updated, because blue-framed faces do not serve as a reference for future trials.

Accordingly, both reference and comparison trials involve a same/different decision against a WM referent, but only the former require WM updating. This means that the gate to WM should be open in reference trials but kept closed in comparison trials. By considering the state of the gate on the previous trial, this protocol further allows one to distinguish between trials where the gate needs to be opened and trials where the gate needs to be closed. Specifically, trials in which the previous trial type is repeated (e.g., two reference trials in a row) do not entail a change in the state of the gate: The gate remains open for successive reference trials and remains closed for successive comparison trials. However, switching from a comparison trial to a reference trial requires gate opening, whereas switching from a reference trial to a comparison trial requires gate closing (see Figure 1A for a trial-by-trial example).

Consequently, the reference-back task enables one to distinguish among WM updating subprocesses using three predefined orthogonal contrasts. These contrasts were utilized in previous studies, showing robust behavioral effects (Kessler, 2017; Rac-Lubashevsky & Kessler, 2016a, 2016b) as well as EEG correlates (Rac-Lubashevsky & Kessler, 2018) and association with spontaneous eye blink rates, an index of central phasic dopaminergic activity (Rac-Lubashevsky, Slagter, & Kessler, 2017). Moreover, RT costs for these contrasts demonstrate split-half reliabilities of .85–.86 and are correlated with performance in the standard *N*-back task (Rac-Lubashevsky & Kessler, 2016b).

The contrasts were defined as follows (also see Figure 1B). (1) "Gate opening" is the difference between reference-switch and reference-repeat trials. All reference trials require the

gate to WM to be open, but only the former, where participants are switching from a comparison trial to a reference trial, involves the process of gate opening. Using a similar logic, (2) "gate closing" was defined as the difference between comparison-switch and comparison-repeat trials—all comparison trials require a closed gate, but only on trials where participants switch from a reference to a comparison trial does the process of gate closing take place. Importantly, because each of the two face stimuli can appear in each of the conditions, the above contrasts are orthogonal to the correct response (being "same" or "different"). Finally, (3) "substitution" refers to replacing old with new information in WM, which occurs on those reference trials where the current stimulus does not match the previous reference (see also Ecker et al., 2010). To substitute old with new information, the irrelevant previously updated information should be removed (Kessler, 2018; Lewis-Peacock et al., 2018; Ecker, Lewandowsky, & Oberauer, 2014). In the reference-back paradigm, removal and substitution are coupled (i.e., each time an item is substituted, the previous one should be removed; but see Kessler, 2018, for a version of this paradigm that enables observing the aftereffects of removal by $n - 2$ repetition costs). Hence, in the current study, we are not attempting to isolate activity specifically related to removing items from memory, and part of the activations identified in the substitution contrast may reflect such a removal process.

It is important to de-confound substitution from the difference between making a "same" versus a "different" response, and this can be achieved by using the difference between "same" and "different" responses in comparison trials as a baseline. Accordingly, substitution is calculated as an interaction contrast, reflecting a larger difference between "same" and "different" responses in reference trials than in comparison trials: ("different"$_{\text{reference}}$ − "same"$_{\text{reference}}$) − ("different"$_{\text{comparison}}$ − "same"$_{\text{comparison}}$).

In addition, the reference-back design enabled us to examine the differential neural activity of being in an open-gate state ("updating mode"; see Kessler & Oberauer, 2014, 2015) compared to a closed-gate state. Accordingly, the "updating mode" is defined by the overall difference between reference trials, where the WM referent has to be updated, and comparison trials, where the referent does not have to be updated. The updating mode contrast only involved trial-type repetition trials, in order not to confound it with gate switching. Thus, whereas "substitution" refers to the situation where updating involves replacing of the old referent with a new one (and possibly also includes removing the now-irrelevant item; see Lewis-Peacock et al., 2018), the "updating mode" refers to the more general situation of being in an open-gate state, regardless of whether the referent has to be replaced ("different" trials) or not. These process designations and the specific contrasts isolating the different updating operations are shown in Figure 1B.

In mapping these reference-back gating costs onto the PBWM model, it should be noted that the latter assumes that the WM gate is closed by default and only opens transiently to relevant inputs, whereas the reference-back contrasts assume the gate to remain open after an updating event (a reference trial) until the next input is evaluated. The fact that robust gate-opening costs are reliably observed (Rac-Lubashevsky et al., 2017; Rac-Lubashevsky & Kessler, 2016a, 2016b; Kessler & Oberauer, 2014, 2015) argues against the PBWM model assumption, because such costs should not be obtained if the gate were closed automatically after each updating event (i.e., reference trial RTs should not differ as a function of the preceding trial being a reference or comparison trial). However, we suggest that the PBWM model can be reconciled with these data via the plausible assumption of context-sensitive gating policies (e.g., Bhandari & Badre, 2018). To wit, in situations where updating is rarely required and distracters are frequent, it would make sense to keep the gate closed by default. By contrast, when updating is required frequently, as in the present task (on 50% of the trials), it would be more efficient to maintain the gate state from the previous trial until the next input is observed, because this policy would minimize the number of gate state switches (cf. Kessler & Oberauer, 2014).

Each block of the reference-back task started with a reference trial, to which the participants did not respond. Then, in each subsequent trial, a framed face was presented for 2 sec, followed by a blank intertrial interval for 2, 4, 6, or 8 sec. Each of the eight conditions (Trial Type × Gate Switching × Response) was presented 10 times in each block, resulting in 320 trials (40 trials per conditions). The order of trials, as well as the duration of the intertrial interval jitter, was determined using Optseq (FreeSurfer analysis tools; Greve, 2002). Each of the four experimental blocks involved two face stimuli from the same gender. The stimuli were changed from one block to another and were counterbalanced between participants. We employed different faces in each block to avoid contributions of long-term memory to performance. Moreover, using only two faces within a given block ensures that there is a high potential for interference, which promotes the use of WM over familiarity-based strategies (e.g., Szmalec, Verbruggen, Vandierendonck, & Kemps, 2011). Note that the contrasts of interest (Figure 1B) are orthogonal with respect to whether a specific face is repeated from one trial to the next, thus preventing face stimulus repetition suppression effects from confounding our results.

FFA Localizer Task

As an FFA localizer task, we employed a block-design 1-back task (taken from Avidan et al., 2014). Different stimulus categories (familiar faces, unfamiliar faces, buildings, objects, and scrambled objects) were presented in 10-sec blocks, with 6-sec intervals between blocks. Within each block, 10 images were presented, each for 800 msec followed by a 200-msec intertrial interval. Within each block, nine images were unique, whereas one image was presented twice in a

row. The participants were asked to keep track of the stimuli and to press a key each time an image was presented twice in a row (1-back). There were seven repetitions of each block type.

fMRI Data Acquisition and Preprocessing

All fMRI data were collected at the Brain Imaging Research Center, Soroka Medical Center, Beer-Sheva, using a 3-T Philips Ingenia MRI scanner. The scanning of each participant started with a 3-D structural scan, acquired by a T1-weighted sequence that yielded high-resolution images of 1-mm^3 voxel size with a matrix of 256 × 256 for 170 slices. Functional data were collected by a T2*-weighted sequence (repetition time = 2000 msec, echo time = 35 msec, flip angle = 90°). Thirty-five slices were scanned in ascending order with a 96 × 96 matrix size and a 2.61 × 2.61 mm voxel resolution with 3-mm thickness. Two hundred fifteen volumes were acquired. Behavioral responses were recorded using a two-key box the participants held in their right hand and pressed with their index or middle finger for different and same responses, respectively. Imaging data were preprocessed and analyzed using SPM 12 (Welcome Trust Centre for Neuroimaging, London, UK; www.fil.ion.ucl.ac.uk/spm). Each participant's functional images were realigned, coregistered to the anatomical image, and slice-time corrected. Then, the images were normalized into Montreal Neurological Institute space (with a 2-mm^3 voxel size interpolation) and smoothed using a 6-mm Gaussian kernel to FWHM.

Statistical Data Analysis

In a first-level analysis, for each participant, a task model was constructed with event-based stick functions, convolved with a canonical hemodynamic response function, and high-pass filtered (128 sec) to remove low-frequency signal drift. The participant-level task matrices included one regressor for each of the eight conditions resulting from the 2 (Trial Type: reference vs. comparison) × 2 (Gate Switch: repeat vs. switch) × 2 (Response: same vs. different) factorial design shown in Figure 1B. The models also included a regressor accounting for error trials, null trials, the grand mean, and six head-movement regressors. Four linear contrasts were defined to estimate activation for gate opening, gate closing, substitution, and updating mode, respectively, as explained in the task description above (see Figure 1B). For the FFA localizer task, the individual task models were constructed in the same manner, but coding for blocks of face stimuli versus nonface stimuli, which were contrasted against each other.

The individual participants' contrast images were then submitted to a second-level one-sample *t* test group-level analysis, where participants were treated as random effects. We pursued two broad sets of fMRI analyses, the first being an exploratory whole-brain analysis and the second being an ROI analysis grounded in our a priori interest of closely interrogating the role of different nodes of the WM network (the constituent parts of the FPN, the BG nuclei, and the thalamus) and the FFA in distinct aspects of the WM updating process. Accordingly, we took a very conservative approach for guarding against false positives in the exploratory analysis, using a voxel-based FWE with a threshold of $p < .05$ and a minimum cluster size (K_E) of 10 significant voxels for the whole-brain analysis, and employed a less conservative approach for the a priori ROI analysis (which in turn is less likely to avoid false negatives) by using a voxel-based false discovery rate with a threshold of $p < .05$. Note that we used voxel-based rather than cluster-based thresholding throughout to bypass recent concerns about common cluster-based correction approaches (Cox, Chen, Glen, Reynolds, & Taylor, 2017; Eklund, Nichols, & Knutsson, 2016). On the basis of a large WM literature (D'Esposito & Postle, 2015; Harding et al., 2015; Nee et al., 2013; Rottschy et al., 2012; McNab & Klingberg, 2008; O'Reilly & Frank, 2006; Frank et al., 2001), the WM network ROIs were defined anatomically, using the WFU PickAtlas toolbox (Maldjian, Laurienti, Kraft, & Burdette, 2003) masks of the BG and thalamus (using automated anatomical labeling labels; Tzourio-Mazoyer et al., 2002), and assembling an FPN mask, using Brodmann's areas, by combining masks of the dlPFC (BA 8, BA 9, BA 46), mPFC/ACC (BA 24, BA 32), and the PPC (BA 7, BA 40). We specified a functional ROI of the FFA via the localizer scan. Finally, in addition to searching for significant clusters of activation within each ROI, we also ran the abovementioned contrasts on mean activity (beta estimates) extracted from each ROI. Bayes factors favoring the alternative (BF_{10}) and null (BF_{01}) hypotheses were calculated using JASP software (JASP Team, 2019) with a default prior. The latter is especially important to establish meaningful null effects. Finally, we also conducted some exploratory brain–behavior correlation analyses between mean ROI betas and behavioral contrast defining gate-opening and gate-closing costs.

RESULTS

Behavioral Results

All the conditions, along with the four a priori contrasts of interest, were tested on both RT and accuracy; mean RTs for the key conditions are shown in Figure 2, and descriptive and inferential statistics are presented in Tables 1 and 2, respectively. The results fully replicated those of the original studies on the reference-back paradigm (Rac-Lubashevsky & Kessler, 2016a, 2016b): As can be seen in Figure 2, mean RT for reference trials was substantially slower than that for comparison trials, reflecting the cost of the WM updating mode (54 msec, $p < .001$). As shown in Figure 2A, switch trials were found to be significantly slower than repeat trials, both in reference trials and in comparison trials, reflecting the costs of gate opening (72 msec, $p < .001$) and gate closing (52 msec, $p < .001$),

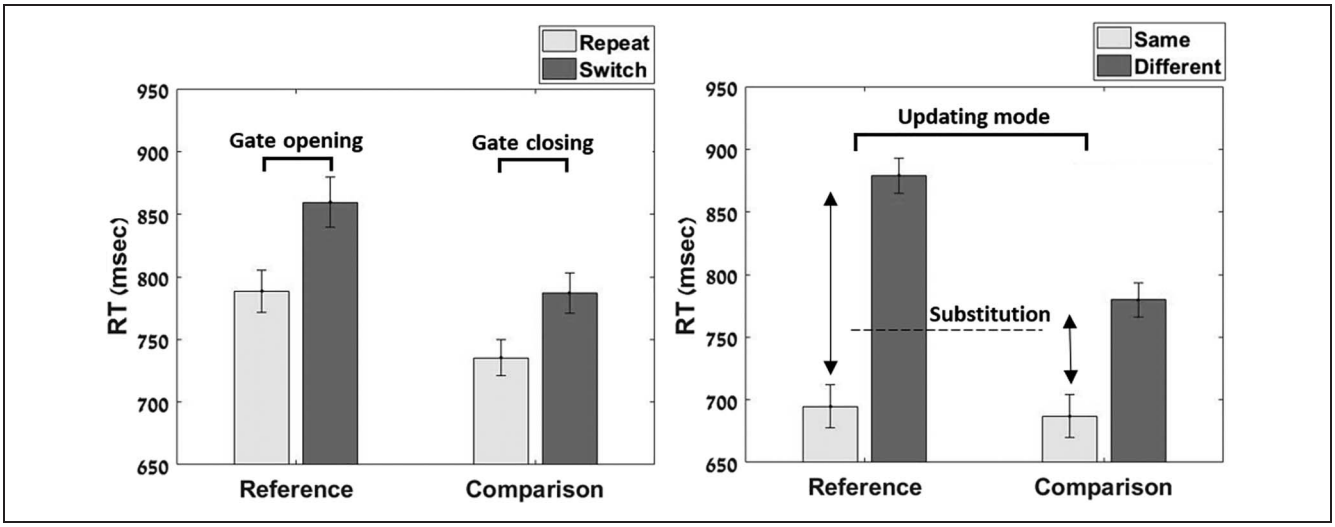

Figure 2. Mean RTs (and standard error of the mean) for reference and comparison trials are displayed as a function (A) of whether the condition was repeated or switched to and (B) of whether the stimulus/response was the same or different to the reference stimulus held in WM. The figure also highlights the four key contrasts defining gate opening, gate closing, substitution, and updating mode.

respectively. Finally, the interaction contrast comparing "same" and "different" response conditions between reference and comparison trials revealed a robust substitution cost (92 msec, $p < .001$). In summary, the behavioral results showed that our adaptation of the reference-back protocol was successful in revealing the behavioral signatures of updating, gate opening and closing, and substitution processes in WM, thus providing a solid basis for interrogating the fMRI data for neural substrates of these processes.

Imaging Results

Exploratory Whole-Brain Analysis

We began with an exploratory whole-brain analysis of each contrast of interest using a conservative correction threshold (voxel-wise FWE $p < .05$, $K_E > 10$). Dorsal views of cortical activations revealed by each contrast are presented in Figure 3 (for a full list of activated clusters, see Table 3). The process of gate opening (required in reference trials that follow a comparison trial) was associated with increased activation in dorsal and dorsomedial frontal and parietal regions, with particularly large clusters of activity observed in the posterior and medial aspects of the PPC, including the precuneus. In addition, gate opening was associated with activity in the thalamus, as well as an extensive posterior cluster stretching from the cuneus into parts of visual cortex, including the fusiform gyrus (for a full list of activated clusters, see Table 3). Given that this contrast controls for basic visual input (which is equated between reference and comparison trials), the latter data suggest that the process of gating visual information into WM may be directly reflected

Table 1. Descriptive Statistics of the RT and Accuracy Data

	Conditions		*RT*		*Accuracy*	
Trial Type	*Gate Switch*	*Response*	*Mean (msec)*	*SD*	*Mean (%)*	*SD*
Reference	Repeat	Same	694	113	98	3.02
		Different	882	158	97	3.28
	Switch	Same	773	160	98	2.62
		Different	945	199	96	4.48
Comparison	Repeat	Same	688	116	98	2.47
		Different	782	145	98	3.15
	Switch	Same	716	114	98	2.93
		Different	857	165	97	2.67

Table 2. Summary of the Inferential Statistics of the RT and Accuracy Data Analysis

	RT				*Accuracy*			
Contrast	*Mean Difference (msec)*	*F(1, 47)*	*p*	η_p^2	*Mean Difference (%)*	*F(1, 47)*	*p*	η_p^2
Gate opening	72	71.12	<.001	.60	−0.4	1.02	.31	.02
Gate closing	52	78.33	<.001	.62	−0.6	3.04	.08	.06
Substitution	92	162.45	<.001	.77	−0.1	0.71	.40	.01
Updating mode	54	57.88	<.001	.47	−0.1	2.36	.13	.05

in enhanced activity in relevant visual regions (see also ROI Analysis section below).

The analysis of the gate-closing process (required on comparison trials that follow reference trials) did not yield significant activations with the conservative FWE whole-brain correction. To probe further for potential neural substrates of gate closing, we applied a more lenient form of whole-brain correction (voxel-wise FDR $p < .05$, $K_E > 10$) to this contrast, which revealed primarily activity in bilateral PPC, specifically in the superior parietal lobule/intraparietal sulcus, along with smaller clusters of dorsal frontal activation (for full listing of active clusters, see Table 3). When WM did not only have to be updated but information in WM had to be replaced ("substitution," required on reference trials where the current stimulus mismatched the WM referent), activity was enhanced in left dlPFC (middle frontal gyrus) and inferior parietal lobule (for full listing of active clusters, see Table 3). Finally, being in an "updating mode" was also associated with increased activation of dorsal frontal and parietal regions, including most prominently the left PPC.

In summary, in line with expectations, the exploratory whole-brain analysis identified core components of the FPN as supporting the regulation of WM updating/protection processes. However, these contrasts also suggest some regional differences, suggesting a relatively greater involvement of medial posterior parietal (and visual) cortex in gate opening, of more lateral posterior parietal regions in gate closing operations and relatively stronger prefrontal involvement in the substitution process.

ROI Analysis

Activity related to the different WM updating operations in a priori ROIs was examined with ROI-wide FDR correction, using a voxel-wise threshold of FDR $p < .05$. Dorsal views/axial slices illustrating key findings are shown in Figure 4. A list of peak coordinates is shown in Table 4. In addition, we extracted and analyzed mean activity estimates from each of the ROIs. Although this analysis is necessarily less sensitive, as it averages activity over entire anatomical regions, it allowed us to further quantify the potential regional functional specializations with respect to WM updating subprocesses and to ensure that the inferences derived from the ROI-based search do not simply reflect (quantitative) thresholding effects, but genuinely (qualitatively) different activity patterns. Figures 5 and 6 present mean beta values extracted for each ROI (broken down into nuclei, in the case of the BG). A summary of the entire statistical analysis, including Bayes factors, is presented in Table 5.

BG. In the ROI-based search, we observed enhanced activity in BG nuclei for gate opening and substitution, but not for gate closing or being in an updating mode (see Figures 4 and 5). More specifically, opening of the gate to WM was associated with the most widespread increase in activity, involving all of the BG nuclei bilaterally (Figure 5). In contrast, no activation increase was detected in any part of the BG during WM gate closing or updating mode. Finally, the substitution of old WM content with new information was associated with increased activation in the caudate, left putamen, and left pallidum.

This pattern of results was largely replicated in the analysis of mean activation estimates for individual nuclei (Table 5; Figure 5). Mean activation of the BG was significantly higher during switch than repeat trials in reference trials, but not in comparison trials. Bayes factors corroborated the implication of the BG in gate-opening processes,

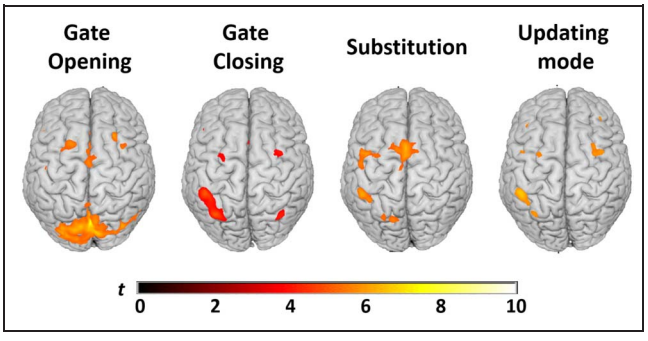

Figure 3. Whole-brain group search results for neural substrates of the different WM updating subprocesses/states, displayed as rendered dorsal 3-D views (voxel-wise FWE $p < .05$, $K_E > 10$). The gray clusters presented in "gate-closing" contrast refer to the clusters found using voxel-wise FDR threshold correction ($p < .05$).

Table 3. Whole-Brain Analysis' List of Peak Activation in MNI Coordinates

			MNI Coordinates			
Region	*HM*	K_E	*x*	*y*	*z*	*Z*
Gate opening (reference switch > reference repeat)						
Precuneus		4647	0	−70	34	7.33
Fusiform gyrus	R	133	30	−74	−10	7.05
SMA	L	43	−2	−8	52	5.56
MFG	L	20	−22	14	58	5.41
IFG	L	37	−50	24	22	5.39
Thalamus	L	18	−8	−18	10	5.34
Insula	L	17	−28	28	0	5.25
Gate closing (comparison switch > comparison repeat)						
Parietal lobule	L	943	−36	−66	50	5.05
Parietal lobule	R	276	40	−52	−28	5.18
BA 46	L	42	−46	28	24	4.23
Middle frontal gyrus	R	50	34	2	58	4.10
Middle frontal gyrus	L	46	−30	−2	66	3.97
SMA	L	57	−2	14	48	4.10
BA 10	L	30	−42	44	0	4.10
Substitution (different–same reference > different–same comparison)						
SMA	L	259	−2	8	56	6.06
Middle frontal gyrus	L	83	−50	−44	50	5.65
IPL	L	37	−44	2	50	6.53
Updating (repeat reference > repeat comparison)						
IPL	L	302	−48	−42	50	6.17
IPL	R	82	34	−44	38	5.61
MFG	R	49	38	4	58	5.54
SPL	L	22	−34	−64	52	5.42
MFG	L	10	−28	0	66	5.21

Z refers to *z* score at peak activated voxel. MNI = Montreal Neurological Institute; SPL = superior parietal lobule.

especially in the caudate and pallidum (gate opening: caudate $BF_{10} = 25.17$, putamen $BF_{10} = 6.38$, pallidum $BF_{10} = 100.46$; gate closing: caudate $BF_{10} = 0.20$, putamen $BF_{10} = 0.40$, pallidum $BF_{10} = 0.19$), and they also provided support against an involvement of these two nuclei in gate closing (caudate: $BF_{01} = 4.91$; pallidum: $BF_{01} = 5.14$). Updating mode-related activity was not significant for any contrast, and the noninvolvement of the BG in that mode was supported by the Bayes factor results, which favored the null hypothesis (putamen: $BF_{01} = 5.96$; pallidum: $BF_{01} = 4.87$). Finally, only mean activity in the caudate also exhibited some evidence for an involvement in the substitution process ($BF_{10} = 3.48$). In summary, the present data support the long-standing proposal of BG involvement in input gating of WM content, by showing that the BG are robustly associated with the process of opening the gate to WM. We also observed some evidence for substitution-related activity, but most importantly, supported by the

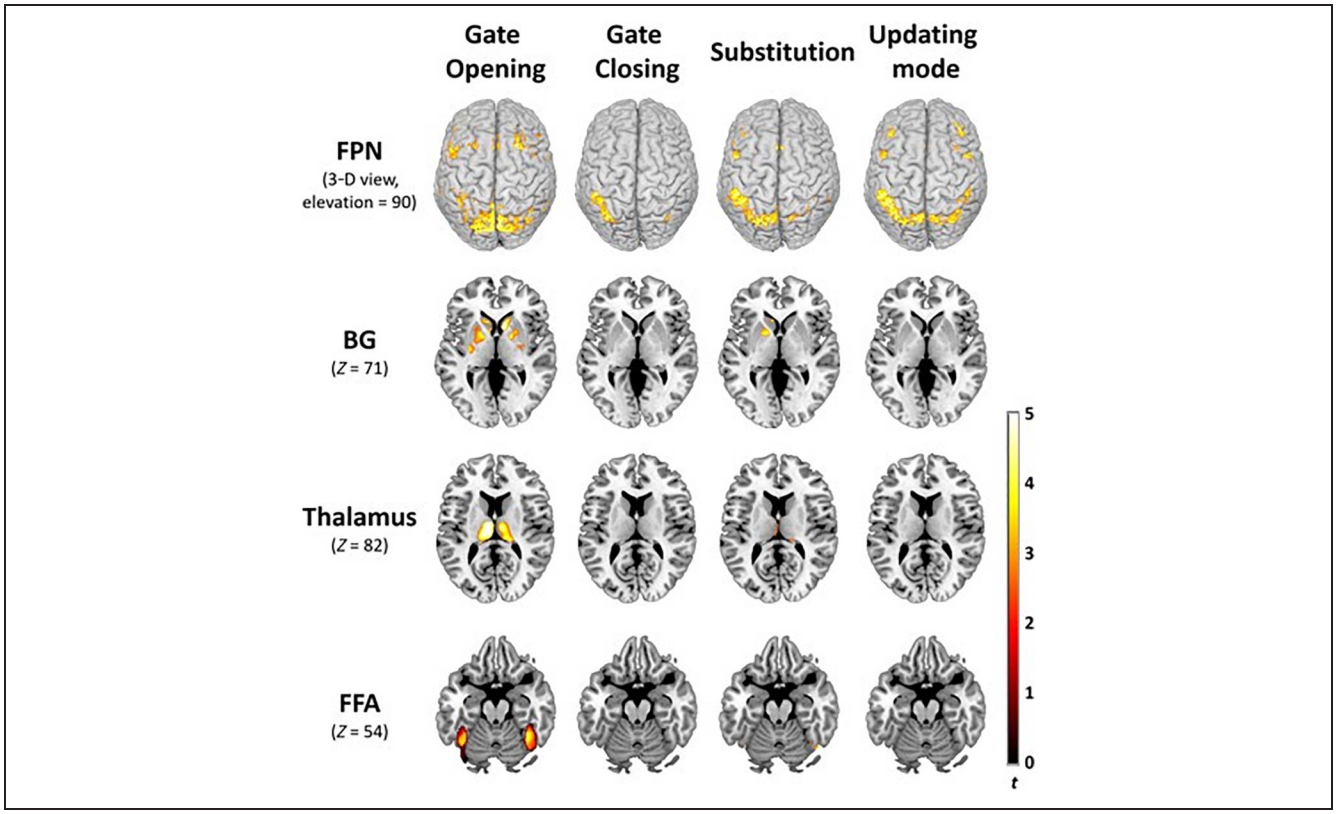

Figure 4. Results of the ROI analyses are displayed as a function of WM updating subprocess (columns) and ROI (rows) on dorsal-view rendered 3-D brains (top row) and axial slices (other rows).

Bayesian analysis results, the BG nuclei seem to play no active role in closing the gate to WM.

Thalamus. Similar to the BG, the thalamus was found to display activity increases during WM gate opening but displayed no detectable increase in activity during the gate-closing operation or with respect to being in an updating mode. This was born out both by the search for significant clusters within the thalamus ROI (Figure 4) and by the analyses run on mean thalamus activation (Figure 5): Switch trials evoked higher mean activation than repeat trials in reference trials but not in comparison trials, and this gate-opening effect was strongly supported by Bayes factor analysis (BF_{10} = 3801). By contrast, we observed some evidence against the thalamus' involvement in gate closing (BF_{01} = 3.63). Moreover, neither the substitution cost nor updating mode contrasts were significant, with the Bayes factor analysis speaking against thalamus involvement in the updating mode (BF_{01} = 6.34). In summary, as in the BG, we observed strong evidence for activity increase in the thalamus when the gate to WM had to be opened, whereas we observed some evidence against an involvement in gate closing.

FPN. As already suggested by the whole-brain analysis above, we found that components of the FPN were activated by all WM updating subprocesses, but the pattern of activation was suggestive of a functional fractionation of the different FPN nodes (see Figures 4 and 6). Specifically, the search for significant activated clusters within the ROIs showed that, whereas the entire FPN was robustly and bilaterally activated during the gate-opening operation and by being engaged in an updating mode, the process of WM content substitution produced much more lateralized activity in the left parietal and left lateral frontal cortex, and gate closing was associated almost exclusively with enhanced left parietal activation along the IPS (Figure 4).

The analysis of mean ROI activity and the Bayes factors confirmed this picture: As shown in Figure 6 (and in Table 5), the gate-opening contrast was significant for mean activation in all FPN components. Moreover, the Bayes factors favoring the hypothesis of these regions' involvement in gate opening supported this pattern (dlPFC: BF_{10} = 1375.50; mPFC\ACC: BF_{10} = 40.34; PPC: BF_{10} = 378.96). In contrast, the gate-closing contrast at the level of mean ROI activity was not found to be significant in any of the FPN components. Bayes factors favored the null hypothesis of no involvement in gate closing for the mPFC (BF_{01} = 6.24) but did not support either hypotheses for the other components (see Table 5). The substitution cost contrast was significant for the PPC, with strong support from the Bayes factor analysis (BF_{10} = 14.47), but not the frontal FPN components, although Bayes factors indicated some

Table 4. ROI Analysis Peak Activation in MNI Coordinates

				MNI Coordinates			
Region		*HM*	K_E	*x*	*y*	*z*	*Z*
Gate opening (reference switch > reference repeat)							
BG	Caudate	L	477	−8	22	2	4.66
		R	450	10	14	−2	4.90
	Putamen	L	608	−28	12	−4	3.76
		R	254	36	−8	−2	3.48
	Pallidum	L	206	−14	0	6	4.98
		R	108	16	4	2	4.09
Thalamus		L	1031	−8	−18	10	5.34
		R	865	8	−14	8	4.71
FPN	dlPFC	R	567	26	20	44	5.58
		L	428	−46	20	26	4.97
	mPFC\ACC	R	578	4	−8	50	4.67
		L	661	−2	−8	50	5.54
	PPC	L	2286	−2	−72	34	6.83
		R	1697	4	−68	32	7.10
FFA		L	208	−38	−62	−18	4.93
		R	276	36	−52	−16	5.88
Gate closing (comparison switch > comparison repeat)							
FPN	dlPFC	L	10	−46	28	24	4.23
	mPFC\ACC	L	15	−4	14	46	3.99
	PPC	L	425	−46	−40	50	4.96
		R	25	34	−68	50	4.05
Substitution (different–same reference > different–same comparison)							
BG	Caudate	L	24	−8	22	2	3.60
		R	68	16	4	20	3.46
	Putamen	L	42	−16	6	6	3.61
	Pallidum	L	35	−14	6	2	3.89
Thalamus		L	13	−2	−10	12	3.21
		R	15	14	−30	10	3.26
FPN	dlPFC	L	129	−50	6	40	5.03
		R	24	4	14	56	4.56
	mPFC\ACC	L	162	−2	8	52	5.43
		R	51	4	16	46	3.98
	PPC	L	1289	−52	−40	50	5.57
		R	161	14	−66	60	4.42

Table 4. (*continued*)

Region		*HM*	K_E	*MNI Coordinates*			*Z*
				x	*y*	*z*	
FFA		R	14	44	−64	−20	3.90
		L	14	−40	−62	−20	3.52
Updating (repeat reference > repeat comparison)							
FPN	dlPFC	L	122	−50	8	42	4.26
		R	81	38	40	36	5.42
	mPFC\ACC	L	33	−4	12	46	4.25
	PPC	L	1218	−50	−40	48	6.11
		R	830	40	−42	40	5.07

Z refers to *z* score at peak activated voxel.

evidence for the dlPFC's involvement in substitution (BF_{10} = 3.88). Finally, the updating mode contrast was significant for mean activity in the PPC and marginally significant for the frontal regions. The Bayes factor analysis did not provide support for or against this involvement, however (see Table 5).

In summary, in line with the results of the whole-brain analysis, the ROI-based analysis provided additional evidence for a functional dissociation of FPN components' roles in WM updating, with the frontal and parietal nodes being concerned with gate opening and being in an updating mode, but PPC additionally contributing to substitution of information in WM. Whereas the ROI-based search revealed some activity in the left PPC during gate closing, Bayesian analysis on the mean beta values of PPC provided neither support for (BF_{10} = 0.62) nor against

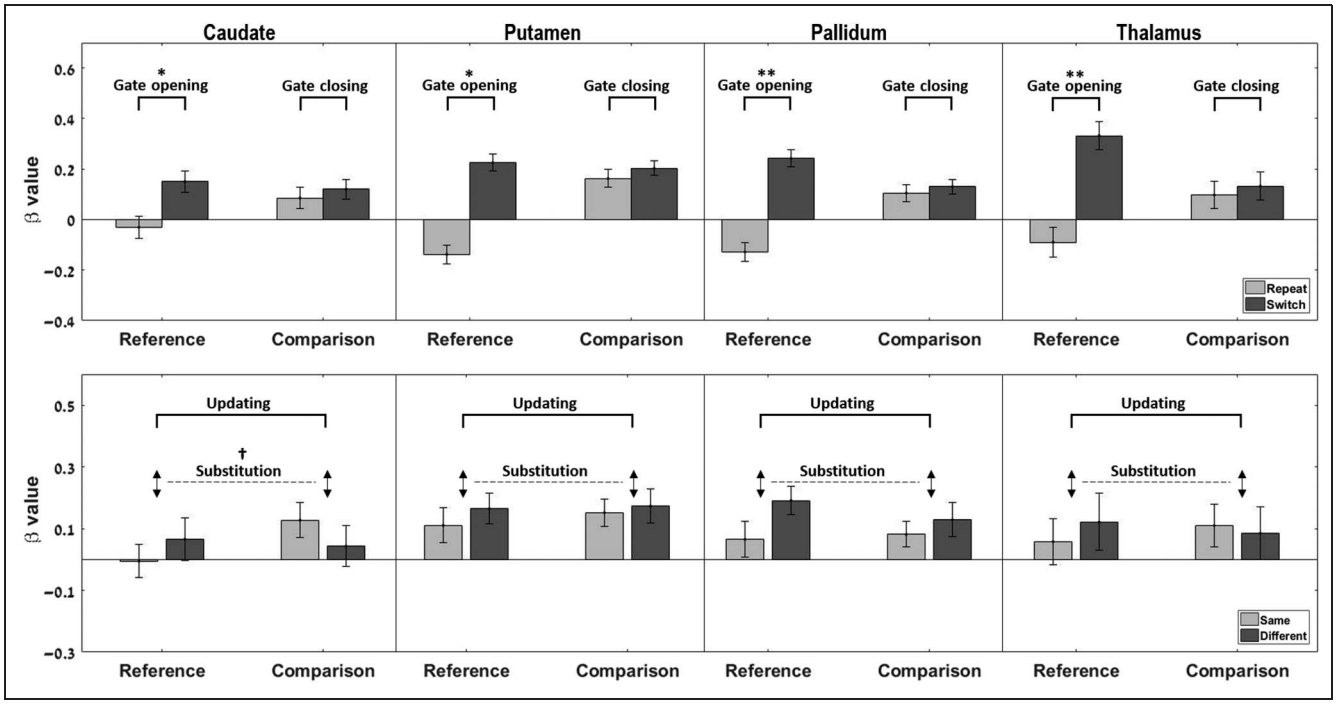

Figure 5. Mean activity estimates (and standard error of the mean) for the BG nuclei and thalamus are shown for reference and comparison trials as a function of whether the condition was repeated or switched to (top panel) and of whether the stimulus/response was the same or different to the reference stimulus held in WM (bottom panel). The figure also highlights the four key contrasts defining gate opening, gate closing, substitution, and updating mode (†p < .01, *p < .05, **p < .001).

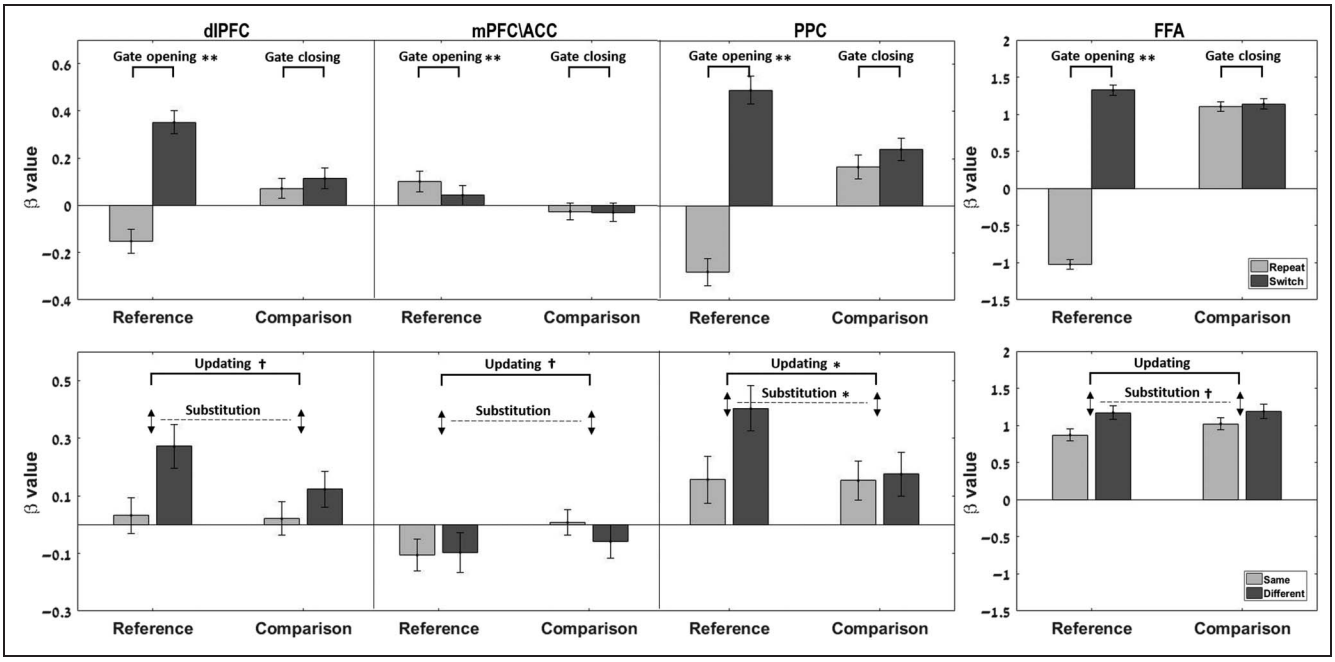

Figure 6. Mean activity estimates (and standard error of the mean) for the FPN components and the FFA are shown for reference and comparison trials as a function of whether the condition was repeated or switched to (top panel) and of whether the stimulus/response was the same or different to the reference stimulus held in WM (bottom panel). The figure also highlights the four key contrasts defining gate opening, gate closing, substitution, and updating mode (†$p < .01$, *$p < .05$, **$p < .001$).

($BF_{01} = 1.59$) this region's involvement in gate closing. Together, we view these data as merely suggestive of a possible involvement of PPC in gate closing.

FFA. The (functionally defined) FFA ROI displayed a similar activity pattern to that observed in the BG, thalamus, and frontal (but not parietal) FPN components. In the search for active clusters, we observed loci in the FFA that showed significant activation increases during substitution, but the most pronounced activity was observed during gate opening (Figure 4). In contrast, no activity increase was detected during gate closing or in relation to the updating mode. The same pattern was confirmed in the analysis of mean FFA activation (Figure 6) and in the Bayesian analysis. Specifically, there was greater mean activation during switch trials than repeat trials in reference trials but not in comparison trials, and an involvement of the FFA in gate opening was strongly supported by the Bayes factor analysis ($BF_{10} = 50.30$). Moreover, we found some evidence against the FFA's involvement in gate closing, with the Bayes factor favoring the null hypothesis ($BF_{01} = 3.75$). Substitution cost and updating mode contrasts were not significant at the mean ROI activity level; however, the Bayes factor indicated evidence in favor of the FFA's involvement in substitution ($BF_{10} = 8.13$). Thus, intriguingly, processes related to allowing sensory content to enter WM (gate opening) and/or to replace WM representations (substitution) seem to have a direct impact on activity in the sensory regions that are involved in processing and/or representing the relevant stimulus material.

Brain–Behavior Correlations

Finally, we conducted exploratory, post hoc correlation analyses across participants between ROI activity (contrast beta weights) and behavioral (RT) gate-opening and gate-closing costs. On the basis of the above results, we expected to observe positive correlations between gate-opening costs and the corresponding ROI contrasts as well as null correlations for gate-closing costs. It should be noted, however, that even with our larger-than-average sample ($n = 48$), brain–behavior correlation analyses of this type are likely underpowered (Cremers, Wager, & Yarkoni, 2017; Poldrack et al., 2017), and the below results should therefore be interpreted with caution. The correlations between gate-opening cost and the corresponding ROI contrast were all positive in direction and small ($r = .15$–$.32$), with only the thalamus data passing the conventional statistical significance criterion (dlPFC: $r = .15$, $p = .316$; mPFC: $r = .15$, $p = .321$; PPC: $r = .13$, $p = .384$; caudate: $r = .24$, $p = .104$; putamen: $r = .16$, $p = .271$; pallidum: $r = .16$, $p = .267$; thalamus: $r = .32$, $p = .025$; FFA: $r = .17$, $p = .236$). The correlations between gate-closing cost and the corresponding ROI contrast hovered around zero or trended negative in direction, and none of them reached significance (dlPFC: $r = -.14$, $p = .335$; mPFC: $r = -.09$, $p = .560$; PPC: $r = -.17$, $p = .253$; caudate: $r = -.09$, $p = .460$; putamen: $r = -.04$, $p = .762$; pallidum: $r = -.09$, $p = .522$; thalamus: $r = -.26$, $p = .080$; FFA: $r = -.02$, $p = .867$). In summary, although the direction of correlations observed was in line with expectations (and significant for the thalamus and gate-opening cost), the results were largely inconclusive.

Table 5. Summary of the Statistical Analysis on the Mean Activity Beta Values, Divided by ROIs and by the Subprocesses

	Gate Opening				*Gate Closing*				*Substitution*				*Updating Mode*			
	F(1, 47)	*p*	BF_{10}	BF_{01}	*F(1, 47)*	*p*	BF_{10}	BF_{01}	*F(1, 47)*	*p*	BF_{10}	BF_{01}	*F(1, 47)*	*p*	BF_{10}	BF_{01}
Caudate	9.66	.003**	25.17	0.04	0.68	.41	0.20	4.91	2.80	.10*	3.48	0.28	2.09	.15	0.41	2.39
Putamen	5.28	.02**	6.38	0.15	1.25	.26	0.40	2.47	0.12	.72	0.33	3.03	0.38	.53	0.16	5.96
Pallidum	13.41	<.001**	100.46	0.01	0.57	.45	0.19	5.14	0.86	.35	0.78	1.27	0.38	.53	0.20	4.87
Thalamus	20.99	<.001**	3801.41	2.63e−4	0.41	.52	0.27	3.63	0.49	.48	0.57	1.74	0.01	.89	0.15	6.34
dlPFC	22.93	<.001**	1375.50	7.27e−4	1.24	.26	0.36	2.70	2.38	.12	3.88	0.25	3.75	.06*	0.77	1.28
mPFC\ACC	12.41	<.001**	40.34	0.02	0.01	.91	0.16	6.24	0.71	.40	0.38	2.58	3.81	.06*	0.81	1.22
PPC	19.24	<.001**	378.96	0.003	2.34	.13	0.62	1.59	5.83	.02**	14.47	0.06	4.26	.04**	2.39	0.41
FFA	24.95	<.001**	50.30	0.02	0.35	.55	0.26	3.75	2.90	.09*	8.13	0.12	2.56	.11	0.39	2.52

For each Region × Subprocess, Bayes factors favoring the hypothesis (BF_{10}) and the null hypothesis (BF_{01}) were calculated as well.

*$p < .1$.

**$p < .05$.

DISCUSSION

There is copious evidence for the involvement of the FPN, BG, and thalamus in maintaining and updating WM content, but how the regions may differentially contribute to different subprocesses of WM updating is not well understood. To address this important question, we combined fMRI with the recently developed reference-back task, which enabled us—for the first time—to tease apart neural substrates of gate-opening, gate-closing, and substitution processes, as well as an updating mode of operation. Moreover, we used face stimuli to examine updating-related activity in sensory cortex specialized for processing the WM items, namely, the FFA.

Our behavioral results fully replicated previous studies using this protocol (Kessler, 2017; Rac-Lubashevsky & Kessler, 2016a, 2016b). Moreover, the imaging data indicate that the FPN, BG, and thalamus all contribute to gate-opening and substitution processes. Being in an updating mode associated with FPN activity, mainly the PPC, but not the subcortical components. In addition, we found that FFA activity displayed robust effects of WM gate-opening and substitution processes.

Gate Opening, Substitution, and Updating Mode in the WM Network

The use of the reference-back paradigm allowed us to isolate different subprocesses involved in WM updating processes, thus enabling more precise process-to-brain region attribution and, accordingly, a stronger test of the PBWM model (O'Reilly & Frank, 2006; Frank et al., 2001). In general support of that model's proposal of BG–thalamus–PFC circuits supporting WM, we observed involvement of all components of this network (as well as PPC) in the processes of opening the gate to WM and in subsequently replacing the current content of WM with new perceptual information (substitution). The PBWM model posits that it is specifically the BG (and not the FPN) that are implementing the gating operation. However, under the assumptions of the model, it is nevertheless plausible that one would also observe gate opening and substitution-related activity more broadly throughout the WM network, as the opening of the gate, and in particular the substitution process, would be expected to have knock-on effects in the thalamus and FPN. Moreover, in support of the model's differentiation between the gating mechanism in the BG and the representation of WM content in FPN, we found that an "updating mode," which refers to an open-gate state that does not involve any change in gating status, was associated exclusively with FPN and not with BG or thalamic activity. Hence, it appears that opening the gate to WM relies on activating the fronto-thalamic-striatal loop, but only frontoparietal cortex is involved in keeping the gate in an open state to allow for continuous updating.

Gate Opening versus Gate Closing

Intriguingly, whereas we observed strong evidence for the BG, thalamus, and FPN regions' involvement in gate opening, we did not observe strong evidence for any of these region's involvement in gate closing but substantial evidence against such an involvement for the BG and thalamus. Previous behavioral studies demonstrated that both opening and closing the gate to WM involve an RT cost. This was observed both using the reference-bask task (Rac-Lubashevsky & Kessler, 2016a, 2016b, 2018; Rac-Lubashevsky et al., 2017) and in a sequence updating task (Kessler & Oberauer, 2014, 2015). Here, we find that, under roughly equivalent behavioral costs of opening and closing the gate, the neural mechanisms involved in the two operations appear to be distinct. Notably, supported by Bayesian analysis, the BG and the thalamus showed clear single dissociations—with strong evidence for gate-opening-related activity and strong evidence against gate-closing-related activity. However, we did not observe the obverse pattern in any other region. In fact, the only region where we detected any sign of potential involvement in gate closing was the PPC, but the evidence was not conclusive.

We offer three possible interpretations of these results that raise interesting questions for follow-up studies. First, the lack of a BG activation increase in relation to gate closing per se may plausibly reflect the fact that the closed gate (i.e., tonic inhibition of the thalamus) represents a BG default state (e.g., Chevalier & Deniau, 1990), returning to which might not impose additional local metabolic demands on the BG. In other words, the observation of robust behavioral gate-closing costs (in both the current and prior studies), combined with the lack of strong evidence for any one region's contribution to closing the gate to WM, suggests that gate closing may be a time-consuming but not an active process. This possibility fits with the PBWM model suggesting that the closed gate represents a default state. Second, on the basis of our observation of some (although not strong) evidence for an involvement of the PPC in gate closing, it is also possible that gate closing does require active cortical engagement and that this process originates in the PPC. The PPC has been frequently implicated in WM maintenance (e.g., Quentin et al., 2019; Majerus et al., 2016), and it has ample direct anatomical projections to the BG (e.g., Jarbo & Verstynen, 2015; Cavada & Goldman-Rakic, 1991), thus making it a plausible contributor to gating processes. However, given the inconclusive nature of its involvement in gate closing in this study, future studies, perhaps involving targeted neurostimulation of the PPC, are needed to rigorously test this possibility.

Third, another possible account for the lack of active BG (and thalamic and FPN) involvement in the gate-closing process could be that closing the gate to WM relies on direct projections from the ventral tegmental area (VTA) to frontal regions and does not activate the BG (cf. Chatham & Badre, 2015; D'Ardenne et al., 2012; Braver &

Cohen, 2000). Specifically, whereas gate opening should be selective by nature, and involve activating specific BG–PFC loops, gate closing may be nonselective and could operate through the (nonspecific) VTA–PFC pathway, and our imaging protocol may not have been sensitive enough to detect metabolic changes in the VTA, being a very small midbrain region. Future studies could potentially probe whether a different neural signature emerges under conditions where gate closing has to be selective, too, such as when items from multiple categories have to be updated independently and/or employ fMRI protocols optimized for gauging VTA activity.

Finally, it should be noted that, although in our design, gate switching always involved a change in the color of the frame (from red to blue or vice versa), the dissociation between gate opening and closing, and the brain regions involved in the two, makes it extremely unlikely that the gating effects merely reflect the perceptual effects of switching between frame colors. In such a case, symmetrical effects of gate opening and gate closing were expected, unlike our clear indications for a dissociation between the two. Moreover, although it is tempting to interpret gate-switch costs (opening and closing) as task-switch costs, several empirical and theoretical considerations argue against this interpretation. On a theoretical level, such an interpretation suggests that being in an open-gate state represents a different task set than being in a closed-gate state. However, what this view would consider to be switching between two "tasks" is identical to what we refer to as switching between the gate states. Hence, appealing to the notion of task switching would not represent an alternative account but a relabeling of the gate-switching operations and one that in our view is theoretically less coherent. Empirically, the interpretation of gate switching as task switching implies similar neural correlates for gate-opening and gate-closing contrasts (as both compare "task switches" to "task repetitions"). In contrast to this prediction, we observed distinct activity for the two. Finally, previous behavioral work found evidence for task-switching and gate-opening operations to proceed in parallel, thus implying that they are distinct processes (Kessler, 2017).

The Role of Posterior Cortex in WM Updating

Many prior studies have established that PFC is a crucial component of STM (Fuster & Alexander, 1971), cognitive control (Miller & Cohen, 2001), and WM updating and maintenance (e.g., Narayanan et al., 2005). However, the specific role it plays in supporting these abilities is still under debate. One traditional view, which is also implemented in PBWM, holds that PFC plays a key role in encoding and storing goal-directed representation within WM (O'Reilly & Frank, 2006; Courtney, Petit, Haxby, & Ungerleider, 1998; Goldman-Rakic, 1996). In contrast, a more recent view (the sensory recruitment hypotheses) does not identify PFC with temporary storage per se but rather characterizes the lateral PFC as responsible for directing selective attention toward memory representations, with the actual representations being held in the posterior cortex, that is, in areas that also serve the perception and long-term storage of the memoranda (Serences, 2016; Lara & Wallis, 2015; Feredoes et al., 2011; Postle, 2006; D'Esposito, Postle, & Rypma, 2000).

Relevant to this debate, our findings demonstrate the involvement of the FFA in WM updating subprocesses. In particular, gate-opening and substitution processes, but not gate closing, were associated with elevated neural activity in the FFA. This finding supports the notion that posterior, "perceptual" regions play a major role not only in WM maintenance (as already posited by the sensory recruitment account) but also in WM updating. We offer two tentative accounts for this finding. The first is that the posterior cortex is more heavily recruited for WM maintenance in situations that call for substitution of the information or for switching from "passive maintenance" (as reflected in comparison trials) to updating. A second, not mutually exclusive interpretation is that the same signaling cascade that leads to gating information into WM also serves to concurrently boost attention to the to-be-encoded items represented in posterior perceptual regions. Although input gating involves a feed-forward flow of information from perception to WM, representations held in WM may in turn lead to directing attention toward perceptual regions, in a concurrent feed-backward fashion. These possibilities could plausibly be evaluated by combining the reference-back task with more time-sensitive measures of neural activity in future studies.

Conclusions

To conclude, this study revealed for the first time distinct neural activity related to gate opening, gate closing, substitution, and updating mode, by combining the reference-back protocol with fMRI. Whereas gate opening was associated with activation of the BG–thalamus–PFC loop, in accordance with the PBWM model, the same was true for FFA activity. Moreover, we observed strong evidence against an active involvement of the BG and thalamus in the gate-closing process. These results supply important novel data to inform our evolving theories of the neuroanatomical mechanisms supporting WM.

Acknowledgments

This study was funded by U.S.-Israel Binational Science Foundation (BSF) grant awarded to Y. K. and T. E.

Reprint requests should be sent to Gal Nir-Cohen, Department of Cognitive and Brain Sciences, Zlotowski Center for Neuroscience, Ben-Gurion University of the Negev, POB 653, Beer-Sheva, Israel 84105, or via e-mail: nga@post.bgu.ac.il.

REFERENCES

Avidan, G., Tanzer, M., Hadj-Bouziane, F., Liu, N., Ungerleider, L. G., & Behrmann, M. (2014). Selective dissociation between core and extended regions of the face processing network in congenital prosopagnosia. *Cerebral Cortex*, *24*, 1565–1578.

DOI: https://doi.org/10.1093/cercor/bht007, **PMID:** 23377287, **PMCID:** PMC4064011

Badre, D. (2012). Opening the gate to working memory. *Proceedings of the National Academy of Sciences, U.S.A.*, *109*, 19878–19879. **DOI:** https://doi.org/10.1073/pnas.1216902109, **PMID:** 23144220, **PMCID:** PMC3523858

Bhandari, A., & Badre, D. (2018). Learning and transfer of working memory gating policies. *Cognition*, *172*, 89–100. **DOI:** https://doi.org/10.1016/j.cognition.2017.12.001, **PMID:** 29245108, **PMCID:** PMC6016086

Braver, T. S., & Cohen, J. D. (2000). On the control of control: The role of dopamine in regulating prefrontal function and working memory. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII: Control of cognitive processes* (pp. 713–737). Cambridge, MA: MIT Press.

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., et al. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*, 365–376. **DOI:** https://doi.org/10.1038/nrn3475, **PMID:** 23571845

Cavada, C., & Goldman-Rakic, P. S. (1991). Topographic segregation of corticostriatal projections from posterior parietal subdivisions in the macaque monkey. *Neuroscience*, *42*, 683–696. **DOI:** https://doi.org/10.1016/0306-4522(91)90037-O

Chatham, C. H., & Badre, D. (2015). Multiple gates on working memory. *Current Opinion in Behavioral Sciences*, *1*, 23–31. **DOI:** https://doi.org/10.1016/j.cobeha.2014.08.001, **PMID:** 26719851, **PMCID:** PMC4692183

Chevalier, G., & Deniau, J. M. (1990). Disinhibition as a basic process in the expression of striatal functions. *Trends in Neurosciences*, *13*, 277–280. **DOI:** https://doi.org/10.1016/0166-2236(90)90109-N

Cools, R., Clark, L., & Robbins, T. W. (2004). Differential responses in human striatum and prefrontal cortex to changes in object and rule relevance. *Journal of Neuroscience*, *24*, 1129–1135. **DOI:** https://doi.org/10.1523/JNEUROSCI.4312-03.2004, **PMID:** 14762131, **PMCID:** PMC6793591

Cools, R., Sheridan, M., Jacobs, E., & D'Esposito, M. (2007). Impulsive personality predicts dopamine-dependent changes in frontostriatal activity during component processes of working memory. *Journal of Neuroscience*, *27*, 5506–5514. **DOI:** https://doi.org/10.1523/JNEUROSCI.0601-07.2007, **PMID:** 17507572, **PMCID:** PMC6672352

Courtney, S. M., Petit, L., Haxby, J. V., & Ungerleider, L. G. (1998). The role of prefrontal cortex in working memory: Examining the contents of consciousness. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *353*, 1819–1828. **DOI:** https://doi.org/10.1098/rstb.1998.0334, **PMID:** 9854254, **PMCID:** PMC1692423

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, *24*, 1158–1170. **DOI:** https://doi.org/10.3758/s13423-016-1191-6, **PMID:** 27896630

Cox, R. W., Chen, G., Glen, D. R., Reynolds, R. C., & Taylor, P. A. (2017). fMRI clustering in AFNI: False-positive rates redux. *Brain Connectivity*, *7*, 152–171. **DOI:** https://doi.org/10.1089/brain.2016.0475, **PMID:** 28398812, **PMCID:** PMC5399747

Cremers, H. R., Wager, T. D., & Yarkoni, T. (2017). The relation between statistical power and inference in fMRI. *PLoS One*, *12*, e0184923. **DOI:** https://doi.org/10.1371/journal.pone.0184923, **PMID:** 29155843, **PMCID:** PMC5695788

D'Ardenne, K., Eshel, N., Luka, J., Lenartowicz, A., Nystrom, L. E., & Cohen, J. D. (2012). Role of prefrontal cortex and the midbrain dopamine system in working memory updating. *Proceedings of the National Academy of Sciences, U.S.A.*, *109*, 19900–19909. **DOI:** https://doi.org/10.1073/pnas.1116727109, **PMID:** 23086162, **PMCID:** PMC3523834

D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. *Annual Review of Psychology*, *66*, 115–142. **DOI:** https://doi.org/10.1146/annurev-psych-010814-015031, **PMID:** 25251486, **PMCID:** PMC4374359

D'Esposito, M., Postle, B. R., & Rypma, B. (2000). Prefrontal cortical contributions to working memory: Evidence from event-related fMRI studies. In W. X. Schneider, A. M. Owen, & J. Duncan (Eds.), *Executive control and the frontal lobe: Current issues* (pp. 3–11). Berlin, Heidelberg: Springer. **DOI:** https://doi.org/10.1007/978-3-642-59794-7_2

Ecker, U. K. H., Lewandowsky, S., & Oberauer, K. (2014). Removal of information from working memory: A specific updating process. *Journal of Memory and Language*, *74*, 77–90. **DOI:** https://doi.org/10.1016/j.jml.2013.09.003

Ecker, U. K. H., Lewandowsky, S., Oberauer, K., & Chee, A. E. H. (2010). The components of working memory updating: An experimental decomposition and individual differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 170–189. **DOI:** https://doi.org/10.1037/a0017891, **PMID:** 20053053

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences, U.S.A.*, *113*, 7900–7905. **DOI:** https://doi.org/10.1073/pnas.1602413113, **PMID:** 27357684, **PMCID:** PMC4948312

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers*, *28*, 1–11. **DOI:** https://doi.org/10.3758/BF03203630

Fallon, S. J., van der Schaaf, M. E., ter Huurne, N., & Cools, R. (2017). The neurocognitive cost of enhancing cognition with methylphenidate: Improved distractor resistance but impaired updating. *Journal of Cognitive Neuroscience*, *29*, 652–663. **DOI:** https://doi.org/10.1162/jocn_a_01065, **PMID:** 27779907

Feredoes, E., Heinen, K., Weiskopf, N., Ruff, C., & Driver, J. (2011). Causal evidence for frontal involvement in memory target maintenance by posterior brain areas during distracter interference of visual working memory. *Proceedings of the National Academy of Sciences, U.S.A.*, *108*, 17510–17515. **DOI:** https://doi.org/10.1073/pnas.1106439108, **PMID:** 21987824, **PMCID:** PMC3198359

Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory: A computational model. *Cognitive, Affective & Behavioral Neuroscience*, *1*, 137–160. **DOI:** https://doi.org/10.3758/CABN.1.2.137, **PMID:** 12467110

Fuster, J. M., & Alexander, G. E. (1971). Neuron activity related to short-term memory. *Science*, *173*, 652–654. **DOI:** https://doi.org/10.1126/science.173.3997.652, **PMID:** 4998337

Goldman-Rakic, P. S. (1996). The prefrontal landscape: Implications of functional architecture for understanding human mentation and the central executive. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *351*, 1445–1453. **DOI:** https://doi.org/10.1098/rstb.1996.0129, **PMID:** 8941956

Greve, D. N. (2002). Optseq home page. http://surfer.nmr.mgh.harvard.edu/optseq/.

Harding, I. H., Yücel, M., Harrison, B. J., Pantelis, C., & Breakspear, M. (2015). Effective connectivity within the frontoparietal control network differentiates cognitive control and working memory. *Neuroimage*, *106*, 144–153. **DOI:** https://doi.org/10.1016/j.neuroimage.2014.11.039, **PMID:** 25463464

Jarbo, K., & Verstynen, T. D. (2015). Converging structural and functional connectivity of orbitofrontal, dorsolateral prefrontal, and posterior parietal cortex in the human striatum. *Journal of Neuroscience*, *35*, 3865–3878. **DOI:** https://doi.org/10.1523/JNEUROSCI.2636-14.2015, **PMID:** 25740516, **PMCID:** PMC4461697

JASP Team. (2019). JASP (Version 0.11.1) [Computer software]. Available from https://jasp-stats.org/.

Johnson, E. L., King-Stephens, D., Weber, P. B., Laxer, K. D., Lin, J. J., & Knight, R. T. (2019). Spectral imprints of working memory for everyday associations in the frontoparietal network. *Frontiers in Systems Neuroscience*, *12*, 65. **DOI:** https://doi.org/10.3389/fnsys.2018.00065, **PMID:** 30670953, **PMCID:** PMC6333050

Jonides, J., Schumacher, E. H., Smith, E. E., Lauber, E. J., Awh, E., Minoshima, S., et al. (1997). Verbal working memory load affects regional brain activation as measured by PET. *Journal of Cognitive Neuroscience*, *9*, 462–475. **DOI:** https://doi.org/10.1162/jocn.1997.9.4.462, **PMID:** 23968211

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*, 4302–4311. **DOI:** https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997, **PMID:** 9151747, **PMCID:** PMC6573547

Kessler, Y. (2017). The role of working memory gating in task switching: A procedural version of the reference-back paradigm. *Frontiers in Psychology*, *8*, 2260. **DOI:** https://doi.org/10.3389/fpsyg.2017.02260, **PMID:** 29312095, **PMCID:** PMC5742995

Kessler, Y. (2018). *N*–2 repetition leads to a cost within working memory and a benefit outside it. *Annals of the New York Academy of Sciences*, *1424*, 268–277. **DOI:** https://doi.org/10.1111/nyas.13607, **PMID:** 29542817

Kessler, Y., & Oberauer, K. (2014). Working memory updating latency reflects the cost of switching between maintenance and updating modes of operation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 738–754. **DOI:** https://doi.org/10.1037/a0035545, **PMID:** 24446752

Kessler, Y., & Oberauer, K. (2015). Forward scanning in verbal working memory updating. *Psychonomic Bulletin & Review*, *22*, 1770–1776. **DOI:** https://doi.org/10.3758/s13423-015-0853-0, **PMID:** 25962687

Kim, C., Cilles, S. E., Johnson, N. F., & Gold, B. T. (2012). Domain general and domain preferential brain regions associated with different types of task switching: A meta-analysis. *Human Brain Mapping*, *33*, 130–142. **DOI:** https://doi.org/10.1002/hbm.21199, **PMID:** 21391260, **PMCID:** PMC3421461

Lara, A. H., & Wallis, J. D. (2015). The role of prefrontal cortex in working memory: A mini review. *Frontiers in Systems Neuroscience*, *9*, 173. **DOI:** https://doi.org/10.3389/fnsys.2015.00173, **PMID:** 26733825, **PMCID:** PMC4683174

Leber, A. B., Turk-Browne, N. B., & Chun, M. M. (2008). Neural predictors of moment-to-moment fluctuations in cognitive flexibility. *Proceedings of the National Academy of Sciences, U.S.A.*, *105*, 13592–13597. **DOI:** https://doi.org/10.1073/pnas.0805423105, **PMID:** 18757744, **PMCID:** PMC2527350

Lewis-Peacock, J. A., Kessler, Y., & Oberauer, K. (2018). The removal of information from working memory. *Annals of the New York Academy of Sciences*, *1424*, 33–44. **DOI:** https://doi.org/10.1111/nyas.13714, **PMID:** 29741212

Lopez-Garcia, P., Lesh, T. A., Salo, T., Barch, D. M., MacDonald, A. W., III, Gold, J. M., et al. (2016). The neural circuitry supporting goal maintenance during cognitive control: A comparison of expectancy AX-CPT and dot probe expectancy paradigms. *Cognitive, Affective & Behavioral Neuroscience*, *16*, 164–175. **DOI:** https://doi.org/10.3758/s13415-015-0384-1, **PMID:** 26494483, **PMCID:** PMC4819423

Majerus, S., Cowan, N., Péters, F., Van Calster, L., Phillips, C., & Schrouff, J. (2016). Cross-modal decoding of neural patterns associated with working memory: Evidence for attention-based accounts of working memory. *Cerebral Cortex*, *26*, 166–179. **DOI:** https://doi.org/10.1093/cercor/bhu189, **PMID:** 25146374, **PMCID:** PMC4717284

Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage*, *19*, 1233–1239. **DOI:** https://doi.org/10.1016/S1053-8119(03)00169-1

McNab, F., & Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience*, *11*, 103–107. **DOI:** https://doi.org/10.1038/nn2024, **PMID:** 18066057

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202. **DOI:** https://doi.org/10.1146/annurev.neuro.24.1.167, **PMID:** 11283309

Murty, V. P., Sambataro, F., Radulescu, E., Altamura, M., Iudicello, J., Zoltick, B., et al. (2011). Selective updating of working memory content modulates meso-cortico-striatal activity. *Neuroimage*, *57*, 1264–1272. **DOI:** https://doi.org/10.1016/j.neuroimage.2011.05.006, **PMID:** 21596142, **PMCID:** PMC3908780

Narayanan, N. S., Prabhakaran, V., Bunge, S. A., Christoff, K., Fine, E. M., & Gabrieli, J. D. E. (2005). The role of the prefrontal cortex in the maintenance of verbal working memory: An event-related fMRI analysis. *Neuropsychology*, *19*, 223–232. **DOI:** https://doi.org/10.1037/0894-4105.19.2.223, **PMID:** 15769206

Nee, D. E., Brown, J. W., Askren, M. K., Berman, M. G., Demiralp, E., Krawitz, A., et al. (2013). A meta-analysis of executive components of working memory. *Cerebral Cortex*, *23*, 264–282. **DOI:** https://doi.org/10.1093/cercor/bhs007, **PMID:** 22314046, **PMCID:** PMC3584956

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D. A., Conway, A., Cowan, N., et al. (2018). Benchmarks for models of short-term and working memory. *Psychological Bulletin*, *144*, 885–958. **DOI:** https://doi.org/10.1037/bul0000153, **PMID:** 30148379

O'Reilly, R. C. (2006). Biologically based computational models of high-level cognition. *Science*, *314*, 91–94. **DOI:** https://doi.org/10.1126/science.1127242, **PMID:** 17023651

O'Reilly, R. C., & Frank, M. J. (2006). Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation*, *18*, 283–328. **DOI:** https://doi.org/10.1162/089976606775093909, **PMID:** 16378516

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, *25*, 46–59. **DOI:** https://doi.org/10.1002/hbm.20131, **PMID:** 15846822, **PMCID:** PMC6871745

Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafò, M. R., et al. (2017). Scanning the horizon: Towards transparent and reproducible neuroimaging research. *Nature Reviews Neuroscience*, *18*, 115–126. **DOI:** https://doi.org/10.1038/nrn.2016.167, **PMID:** 28053326, **PMCID:** PMC6910649

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, *139*, 23–38. **DOI:** https://doi.org/10.1016/j.neuroscience.2005.06.005, **PMID:** 16324795, **PMCID:** PMC1428794

Quentin, R., King, J.-R., Sallard, E., Fishman, N., Thompson, R., Buch, E. R., et al. (2019). Differential brain mechanisms of selection and maintenance of information during working memory. *Journal of Neuroscience*, *39*, 3728–3740. **DOI:** https://doi.org/10.1523/JNEUROSCI.2764-18.2019, **PMID:** 30833510, **PMCID:** PMC6510345

Rac-Lubashevsky, R., & Kessler, Y. (2016a). Dissociating working memory updating and automatic updating: The reference-back paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 951–969. **DOI:** https://doi.org/10.1037/xlm0000219, **PMID:** 26618910

Rac-Lubashevsky, R., & Kessler, Y. (2016b). Decomposing the n-back task: An individual differences study using the

reference-back paradigm. *Neuropsychologia*, *90*, 190–199. **DOI:** https://doi.org/10.1016/j.neuropsychologia.2016.07.013, **PMID:** 27425422

Rac-Lubashevsky, R., & Kessler, Y. (2018). Oscillatory correlates of control over working memory gating and updating: An EEG study using the reference-back paradigm. *Journal of Cognitive Neuroscience*, *30*, 1870–1882. **DOI:** https://doi.org/10.1162/jocn_a_01326, **PMID:** 30125218

Rac-Lubashevsky, R., Slagter, H. A., & Kessler, Y. (2017). Tracking real-time changes in working memory updating and gating with the event-based eye-blink rate. *Scientific Reports*, *7*, 2547. **DOI:** https://doi.org/10.1038/s41598-017-02942-3, **PMID:** 28566762, **PMCID:** PMC5451427

Roth, J. K., Serences, J. T., & Courtney, S. M. (2006). Neural system for controlling the contents of object working memory in humans. *Cerebral Cortex*, *16*, 1595–1603. **DOI:** https://doi.org/10.1093/cercor/bhj096, **PMID:** 16357333

Rottschy, C., Langner, R., Dogan, I., Reetz, K., Laird, A. R., Schulz, J. B., et al. (2012). Modelling neural correlates of working memory: A coordinate-based meta-analysis. *Neuroimage*, *60*, 830–846. **DOI:** https://doi.org/10.1016/j.neuroimage.2011.11.050, **PMID:** 22178808, **PMCID:** PMC3288533

Serences, J. T. (2016). Neural mechanisms of information storage in visual short-term memory. *Vision Research*, *128*, 53–67. **DOI:** https://doi.org/10.1016/j.visres.2016.09.010, **PMID:** 27668990, **PMCID:** PMC5079778

Szmalec, A., Verbruggen, F., Vandierendonck, A., & Kemps, E. (2011). Control of interference during working memory updating. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 137–151. **DOI:** https://doi.org/10.1037/a0020365, **PMID:** 20731517

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, *15*, 273–289. **DOI:** https://doi.org/10.1006/nimg.2001.0978, **PMID:** 11771995

van Schouwenburg, M. R., den Ouden, H. E. M., & Cools, R. (2010). The human basal ganglia modulate frontal–posterior connectivity during attention shifting. *Journal of Neuroscience*, *30*, 9910–9918. **DOI:** https://doi.org/10.1523/JNEUROSCI.1111-10.2010, **PMID:** 20660273, **PMCID:** PMC6632831

van Schouwenburg, M. R., Onnink, A. M. H., ter Huurne, N., Kan, C. C., Zwiers, M. P., Hoogman, M., et al. (2014). Cognitive flexibility depends on white matter microstructure of the basal ganglia. *Neuropsychologia*, *53*, 171–177. **DOI:** https://doi.org/10.1016/j.neuropsychologia.2013.11.015, **PMID:** 24291339

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, *438*, 500–503. **DOI:** https://doi.org/10.1038/nature04171, **PMID:** 16306992