

Published in final edited form as:

J Cogn Neurosci. 2009 May ; 21(5): 922–937. doi:10.1162/jocn.2009.21060.

Rule-dependent prefrontal cortex activity across episodic and perceptual decisions: An fMRI investigation of the criterial classification account

Sanghoon Han¹, Scott A. Huettel², and Ian G. Dobbins³

¹ Dept. of Psychology & Neuroscience, Duke University, Durham, North Carolina, USA

² Brain Imaging and Analysis Center, Duke University Medical Center, Durham, North Carolina, USA

³ Dept. of Psychology, Washington University, Saint Louis, Missouri, USA

Abstract

Although lateral prefrontal cortex (LPFC) is clearly involved in decision-making, competing functional characterizations exist. One characterization posits that activation reflects the need to select among competing representations. In contrast, recent fMRI research suggests that activation is driven by the criterial classification of representations, even with minimal competition. To adjudicate between these hypotheses, we used event-related fMRI and contrasted tasks that required different numbers of criterial classifications prior to response in both perceptual and memory domains. Additionally, we manipulated the level of inter-stimulus competition by increasing the number of probes. Experiment 1 demonstrated that LPFC activation tracked the number of intermediate classifications during trials, yet was insensitive to the number of competing probes and the behavioral decline accompanying competition. Furthermore, Experiment 2 demonstrated equivalent increases in LPFC activation for a task requiring two overt criterial classifications (independent-classification) and one requiring two covert criterial classifications prior to the single overt response (same-different judgment). As found in Experiment 1, both tasks showed greater activation than a judgment requiring only one classification act (forced-choice). These data indicate that LPFC responses reflect the number of executed criterial classifications or judgments, independent of the number of competing stimuli and the overt response demands of the decision task.

Keywords

lateral prefrontal cortex; fMRI; decision rule; classification judgment

Considerable evidence suggests that damage to lateral prefrontal cortex (LPFC), especially dorsolateral PFC, impairs judgments across a wide variety of cognitive domains, including judgments about episodic memory (Budson et al., 2005; Levine, Stuss, & Milberg, 1997; Owen et al., 1996; Shallice & Burgess, 1991). For example, one putative role for LPFC's contribution to episodic remembering has been termed "retrieval monitoring", and assumed to assist in the evaluation of recovered memories with respect to specific retrieval goals or plans (Dobbins, Foley, Schacter, & Wagner, 2002; Henson, Rugg, Shallice, & Dolan, 2000). Under this characterization, as memories become impoverished or ambiguous with respect to the retrieval goal, the monitoring demands are held to increase (Henson et al., 2000; Mitchell, Johnson, Raye, & Greene, 2004). This account has been extended to suggest that LPFC is engaged

whenever one monitors the sufficiency of stimulus representations across differing task types (Cabeza et al., 2003; Fleck, Daselaar, Dobbins, & Cabeza, 2006), even when these representations are perceptual instead of episodic.

The monitoring hypothesis shares a conceptual similarity with another hypothesis holding that LPFC is critical for the selection of context-appropriate representations (Liu, Banich, Jacobson, & Tanabe, 2006; Rowe & Passingham, 2001). As competition among mnemonic or perceptual representations increases, either through manipulating the number or similarity of stimuli, an increasing demand for LPFC function is predicted in order to select among currently active representations (Bunge et al., 2002; Incisa della Rocchetta & Milner, 1993; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997), perhaps through the foregrounding or attentional biasing of goal-relevant feature representations (Hester, D'Esposito, Cole, & Garavan, 2007; Raye et al., 2007). Thus both selection and monitoring accounts assume that competition, either through increasingly impoverished and/or numerous representations, underpins the need for

LPFC function

A potentially different characterization of LPFC function has arisen out of working memory research examining the mental “manipulation” of representations. (D'Esposito, Postle, Ballard, & Lease, 1999; Smith & Jonides, 1999). Under the manipulation framework, LPFC is critical for actively working with representations with respect to a standard or criterion. More specifically, it is not the need to select a particular stimulus representation as task appropriate, but the need to re-order or perform other abstracted assignment operations on the representations (i.e., ranking, sorting, alphabetization) that is deemed critical. In keeping with the selection and monitoring accounts however, the manipulation characterization of LPFC arguably also predicts that recruitment would necessarily increase as the number of items to be manipulated increased since this would amplify competition among the item representations (i.e., Braver, Cohen, Nystrom and others, 1997).

Despite the central role afforded competition among representations in these current frameworks there is evidence that suggests insensitivity of LPFC to situations that should lead to increased competition or the need for heightened selection. For example, researchers have found that LPFC activation was not closely tied to performance during classification judgments in the AX-CPT task (Continuous Performance Test) (Barch et al., 1997). This task requires conditional “target” or “non-target” classification of a standard depending upon which cue it follows. The authors examined performance on this task crossing a manipulation of delay between items (long or short) with stimulus degradation (perceptually degraded or intact). Although behavior was unaffected by delay manipulation, a tonic elevation of LPFC was observed in response to the long delay condition. In contrast, the stimulus degradation significantly reduced performance, yet it did not elevate LPFC activation. Since the latter manipulation should clearly increase competition or the need for selection, these findings suggest that competition does not directly influence the region's response. Likewise, Carter and colleagues also found that LPFC was not activated in AX-CPT task conditions expected to elicit competing representations (Carter et al., 1998). Finally, another example of insensitivity of the region to competition was observed in a recent study that separated decision rule-based activation from evidence-based PFC activation during episodic remembering and lexical decisions (Dobbins & Han, 2006). These authors argued that rule complexity and the quality of available evidence represented independent factors governing observed success rates. Consistent with the distinction they found that LPFC activation did not track the large differences in the level of mnemonic evidence supporting judgments. In contrast, the region appeared modulated by the requirement to employ different decision rules, and the authors suggested that LPFC activation varied as a function of the required number of criterial classifications prior to responding.

In the current study, two separate event-related fMRI experiments were designed to test the hypothesis that LPFC activation specifically reflects the number of criterial classifications rendered during a decision problem. If correct, activation should be sensitive to the complexity of different decision rules. Additionally, activation should **not** directly track the level of competition among the representations to which the rules are applied. Furthermore, task difficulty, as indexed by accuracy and reaction time, should also fail to account for LPFC activation. Under the criterial classification hypothesis, it is the action of judging a representation with respect to a criterion that is assumed to drive LPFC activity. We examined combinations of three types of decision rules that theoretically differ in the number of criterial classifications required prior to responding (forced-choice, same-different (Experiment 1), and additionally independent-classification (Experiment 2); see Methods and Figure 1), to see if LPFC activation tracked the number of criterial judgments assumed across the different rule types. The design also allowed us to test whether activation was linked to the number of competing stimuli or response representations, as the selection-based accounts instead predict, by increasing the number of competing stimulus representations present during the forced-choice decision type. For example, selecting a target among three candidates should clearly involve more competition than selecting a target among two. Additionally, we imposed a delay between probe presentations in Experiment 2 to see if the regional activity was sensitive to the need to hold intermediate covert classifications in working memory. Decisions were examined across verbal memory and perceptual judgment domains in order to verify that lateral PFC activity was recruited in a highly domain general manner (Duncan, 2001).

METHODS

Participants

Twelve native English-speaking volunteers were included in Experiment 1. Participants were 4 males and 8 females, with a mean age of 21.6 years (age range 19–27 years). An independent sample of fifteen native English speakers (7 female; ages 19–28 years, $M = 22.5$) was enrolled in Experiment 2. Informed consent was obtained in a manner approved by the Institutional Review Board of Duke University Medical Center. The participants were paid \$20 for each hour of participation.

Materials and Tasks

Experiment 1 examined memory judgments for words intermixed with gender judgments for faces. For the words, 568 nouns were drawn randomly from a pool of 1216 words for each subject (on average 7.09 letters, and 2.34 syllables, with a Kucera-Francis corpus frequency of 8.85). From this list, two lists of 284 items (114 old, 170 new items for each cycle) were constructed for use in two memory study/test cycles. For the gender judgment, a total of 496 face pictures (192 females, 304 males) were selected and used in two face runs (Endl et al., 1998). The stimuli were cropped to exclude hair and lower chin areas in order to make the judgments more challenging. Experiment 2 consisted of only the gender judgment task, and a total of 424 face pictures (212 female) were selected from the same stimulus set and pseudo-randomly assigned to the conditions for each subject.

Experiment 1 examined forced-choice and same-different judgments, whereas Experiment 2 added an independent classification judgment and a delay manipulation. During **forced-choice** decisions subjects were presented with an array of probes containing one target and a number of lures, and required to select the target satisfying a single criterion, for example, selecting which of three faces was a female. Regardless of the number of lures present during forced-choice, only one of the presented items satisfies the decision criterion, with the judgment postponed until the relative evidence among competitors has been compared. Thus criterial classification demands are assumed relatively constant across levels of forced-choice even

though increasing the number of lures should generally reduce performance and increase competition among stimulus representations (see Figure 1). During **same-different** judgments subjects are asked to determine whether each of the members arises from the same or different categories. For example, if a pair of faces contained a male face and a female face, the correct response would be “different”. Although same-different and forced-choice tasks can be fully matched in terms of probe construction and motor requirements, same-different judgments require additional criterial classifications prior to response. More specifically, observers must independently classify each probe item with respect to a criterion, hold these intermediate classifications in mind, and finally compare the classifications in order to reach the terminal decision, “same” or “different”. Although the two covert classification operations should heavily tax a region critical for criterial classification, they are not clearly related to a hypothetical mechanism for selecting among competing stimuli or response representations.

Finally, we also investigated **Independent-classification** decisions (Reynolds, McDermott, & Braver, 2006). During these judgments, subjects are asked to overtly and separately classify each presented item. For example, if a pair of faces contained a male and a female face, the observer should separately respond “male” to the left item and then “female” to the right. Here, as with same-different judgments, both items are judged with respect to a decision criterion although the judgments are overt during independent-classification versus covert during same-different trials. Additionally, there is no requirement to hold intermediate judgments in mind during the independent-classification task since the judgment to the first item need not be directly contrasted with the judgment to the second item. Although the independent classification task can be matched to the forced-choice and same-different tasks in terms of the nature and number of probe stimuli, it necessarily involves more overt responses.

fMRI data acquisition

In Experiment 1, scanning was performed on a 4T, General Electric (Waukesha, Wisconsin) scanner using a standard head coil. Functional data were acquired using a spiral-in pulse sequence (acquisition matrix 64×64 , TR = 2000ms, TE = 31 ms, 34 axial slices parallel to AC-PC plane with near-isotropic voxels of $3.75 \times 3.75 \times 3.8$ mm, no gap) designed to minimize susceptibility artifact. Before functional data collection, four dummy volumes were discarded to allow for T1 equilibration. Participants' head motion was minimized using foam padding and a forehead strap. High-resolution T1-weighted anatomical images (3-dimensional SPGR) were acquired for visualization. In Experiment 2, we used a 3T General Electric scanner using a standard head coil and functional data were acquired using a standard EPI pulse sequence. Other scanning parameters/protocols were identical to those in Experiment 1.

fMRI Data Analyses

Analyses focused on memory retrieval and gender discrimination runs, the memory encoding phases of Experiment 1 were not included. Data were preprocessed using SPM99 (Wellcome Dept. of Neurology, UK; <http://www.fil.ion.ucl.ac.uk/spm/>). Slice acquisition timing was corrected by resampling all slices in time relative to the middle slice collected, followed by rigid body motion correction across all runs. Functional data were spatially normalized to a canonical EPI template using a 12-parameter affine and nonlinear cosine transformation and then spatially smoothed with an 8-mm full-width at half-maximum (FWHM) isotropic Gaussian kernel. Each scanning session was rescaled such that the mean global signal was 100 across the volumes.

For the analyses, volumes were treated as a temporally correlated time series and modeled by convolving a synthetic hemodynamic response function and its temporal and dispersion derivative using the onset times for the events. The motion parameters for each run were entered into the model as user-specified regressors. The resulting functions were used as covariates in

a general linear model, along with a basis set of cosine functions that were used to high-pass filter the data, and a covariate representing session effects. The least squares parameter estimates of the best fitting synthetic HRF for each condition of interest (averaged across runs) were used in pair-wise contrasts and stored as a separate image for each subject. These difference images were then tested against the null hypothesis of no difference between contrast conditions using one-tailed *t* tests. The data were statistically analyzed treating subjects as a random effect.

To control for family-wise error resulting from multiple comparisons, a Monte Carlo simulation procedure was conducted to determine the height and extent threshold sufficient for enforcing a maximum Type-1 error of .05 for reported clusters (Slotnick, Moo, Segal, & Hart, 2003). Based on the results of the simulation, clusters were considered significant and further scrutinized if they consisted of 24 or more contiguous voxels (3mm isotropic) exceeding a minimum probability threshold of .003. Functional regions of interest (ROIs) were extracted by MarsBar Toolbox using peristimulus time averaging for the event-related fMRI data surviving the contrasts (Brett, Anton, Valabregue, & Poline, 2002). Percent signal averages were obtained for the above threshold voxels within an 8 mm radius of each of the SPM-identified maxima. For the delayed runs in Experiment 2, selective averaging was performed on the data after filtering and adjustment for all other effects such as low-frequency signal components or effects of other trial conditions, using a finite impulse response model. Mean signal comparisons within the predefined ROIs were considered reliable at an alpha level of .05.

EXPERIMENT 1

Procedures

During Experiment 1 subjects alternated between runs of memory and perceptual judgment tasks using the same decision rules (described below). During the study phase directly preceding memory tests, subjects made a concrete/abstract judgment for each presented word (2 seconds). The study phase had 142 trials including 28 interspersed fixation baseline trials. During the memory test phase, subjects were presented with a single word, a word pair, a triplet, or a quartet in each retrieval trial for 6 s. Prompts indicating the appropriate decision rule were simultaneously presented underneath the words. During the two-alternative forced-choice (2AFC) trials, subjects were presented with a word pair (one target, one lure), above the prompt “Which old?”, which indicated participants should select the location of the studied word (left or right). There were also three-alternative forced-choice (3AFC) (one target, two lures) and four-alternative forced-choice (4AFC) (one target, three lures) tasks, yielding a total of three levels of forced-choice in which the number of task relevant representations increased, but only one item satisfied the criterion. In contrast to these forced-choice trials, there were also two-alternative same-different (2ASD) trials where subjects were required to determine whether the pair contained one old and one new item (different), or were constructed from the same item types (both studied or both lure - same). Immediately below the probes the prompt “Different?” indicated they should press the first button “yes” if they believed that the probes were different or the second button “no” button if instead they believed them to be from the same memory category. There were three-alternative same-different (3ASD) trials where the subjects determined if the triplet contained an odd item type (viz., one old and two new items). The order of 152 retrieval trials¹ including 30 fixation trials (~20% of total trials) was

determined using an optimization algorithm for fMRI experimental design (Wager & Nichols, 2005). This optimal sequencing program was used to order all conditions in all runs of the current experiments. In two thirds of the 2AFC trials the probes were constructed identically as those used during 2AFC trials, direct comparison. To guard against response biases, one third of the remaining 2ASD pairs were constructed of the same item types (either both old or both new), however, these are not analyzed since they differ in construction from the comparison forced-choice trials. Similarly, a third of the 3ASD trials included triplets of the same item types (either all old or all new) that were not analyzed. Finally, there were also simple singleton trials during which the perceptual gender discrimination, subjects were presented with male or female face of probes could not be matched with the forced-choice and same-different trials, all singleton trials are omitted from the analyses and discussion.

(Different?) decisions analogous in form to those used during the memory judgments. Thus the decision rules in the gender discrimination were fully matched with the decision rules in the memory tests and the timing and response deadlines were also matched. In total, Experiment 1 contained two runs of the verbal memory task interspersed with two runs of the gender discrimination task. All subjects started with the gender discrimination task.

Results – Task Performance

During verbal recognition, linear trend analyses across the levels of forced-choice decisions demonstrated both decreasing accuracy ($F(1,11) = 205.97$, $MSe = .004$, $p < .001$) and increasing reaction time ($F(1,11) = 24.03$, $MSe = 138120$, $p < .001$) as the number of lures increased (Table 1). During gender discrimination, accuracy did not significantly decline with increasing numbers of lures ($F(1,11) = 1.14$, $MSe = .002$, $p > .30$) however, reaction time did significantly increase ($F(1,11) = 102.91$, $MSe = 27260$, $p < .001$). Overall, the data demonstrate that performance declined as the number of lures increased during forced-choice judgment, although the decline was more prominent for verbal memory judgments than perceptual gender judgments. Table 1 also lists correct response rates for same-different performance. It is clear that subjects generally found same-different judgments more demanding than forced-choice when the number of stimuli in the arrays was the same. This was anticipated given current decision models of these tasks (Macmillan & Creelman, 1991).

Of key interest was whether there was a condition under which forced-choice and same-different responding was similarly difficult. If there were a forced-choice task with more stimuli, that exhibited the same level of performance as a same-different task with fewer stimuli, this would serve as a strong test of the criterial classification versus competition accounts in the fMRI data. Given this goal, we focused exclusively on contrasting 2ASD versus 4AFC. For verbal memory, there was no reliable difference in accuracy between 4AFC and 2ASD (.49 vs. .57; $t(11) = 1.71$, $p = .12$), or in reaction times (3133ms vs. 3396ms; $t(11) = 1.57$, $p = .15$) during correct responding, indicating that these trials were similar in difficulty for verbal memory. For gender discrimination same-different performance was lower than forced-choice (.85 vs. .64; $t(11) = 6.85$, $p < .001$). Reaction times however did not differ between the 4AFC and 2ASD (2754ms vs. 2916ms; $t(11) = 1.36$, $p = .20$). Thus unlike the verbal memory task, the behavioral data suggested that same-different responding remained more difficult than forced-choice, even when the number of alternatives for the latter was increased to four.

Results – fMRI Data

Rule-Based Modulation in LPFC—The initial analysis focused on contrasting same-different and forced-choice decisions under conditions that were fully matched for the number of alternatives (viz., two and three alternatives) during both memory and perceptual judgments. Furthermore, to rule out domain specific effects, voxels were removed from this conjunction map if they demonstrated sensitivity to the domain of judgment (verbal memory or gender discrimination) at a liberal threshold (.3 no voxel extent). This ensured that any voxel that illustrated even a minor tendency to differentially respond across words or faces regardless of rule type was removed from the resulting map (Figure 2, Table 2).

¹In two thirds of the 2ASD trials the probes were constructed identically as those used during 2AFC enabling direct comparison. To guard against response biases, one third of the remaining 2ASD pairs were constructed of the same item types (either both old or both new), however, these are not analyzed since they differ in construction from the comparison forced-choice trials. Similarly, a third of the 3ASD trials included triplets of the same item types (either all old or all new) that were not analyzed. Finally, there were also simple singleton trials during which subjects received a single item and made memory judgments (old?). For brevity and because the numbers of probes could not be matched with the forced-choice and same-different trials, all singleton trials are omitted from the analyses and discussion.

The results revealed greater activation for the same-different compared to forced-choice rule in bilateral prefrontal cortex (approximate Brodmann's area 46/9) and medial superior frontal (~BA 6) regions, as well as in parietal (~BA 7) and extrastriate areas (~BA 17/18/19). (Table 2; No relative increases were observed during the reverse contrast). Example time courses reconstructed from the right LPFC region (Figure 2A and 2B) confirm that the masking was successful and clearly demonstrate that the response was driven predominantly by the rule and not by the nature of the task domain (verbal memory or gender discrimination). Mean response activation (6 to 8s. post stimulus onset) was further analyzed to see if the number of stimuli itself modulated activation during forced-choice responding in the LPFC regions (4 versus 3 versus 2). In the right LPFC, trend analysis suggested no reliable relationship between the number of stimuli and the level of activation during forced-choice responding for either perceptual gender ($F(1,11) = .001, p > .97$) or verbal memory judgments ($F(1,11) = 1.29, p > .27$). This null pattern was repeated on the left LPFC, which again demonstrated no increase in response for the perceptual ($F(1,11) = .26, p > .61$) or memory ($F(1,11) = .81, p > .38$) forced-choice judgments.

This insensitivity of LPFC activation to the number of competing stimuli during forced-choice was also demonstrated in a parametric whole brain analyses. For both verbal memory and gender discrimination, neither the LPFC regions of interest nor other lateral or dorsal PFC regions were implicated in the parametric whole-brain analyses (4- versus 3- versus 2AFC) even when the threshold was relaxed to .05, 5 voxels. The parametric analyses only revealed bilateral extrastriate areas (~BA 17/18/19) showing a tendency to increase response as a function of the number of task relevant stimuli during the forced-choice decision type.

Thus although the above behavioral analyses clearly indicated forced-choice appropriately became more difficult as more lures were added, the LPFC regions identified in the map (Figure 2) were insensitive to this difficulty effect and to the number of stimuli competing for selection. Furthermore, although the behavioral analysis indicated that 4AFC and 2ASD trials were similar in difficulty for verbal memory, Figure 2 and direct contrasts clearly showed a much higher response for the 2ASD versus the 4AFC verbal memory conditions in both left and right LPFC regions (.75 vs. .48, $t(11) = 3.88, p < .01$; .36 vs. .21, $t(11) = 2.65, p < .05$). That is, even when performance was similar, activation in the LPFC region was significantly higher for the task requiring more intermediate criterial classifications, namely, the same-different judgment. Overall, the data demonstrate that LPFC regions responded as a function of the demand for criterial classification.

Domain Specific Activation

We identified regions sensitive to the material domain –(verbal memory versus perceptual gender discrimination) by directly contrasting the domains² collapsed across rules and number of alternatives. To restrict the resulting map to domain specific differences, regions demonstrating even slight sensitivity to the nature of the decision rule were excluded using a liberal threshold (.3, no extent). Thus the resulting regions would largely be driven by material domain, but would be insensitive to differences in the specific task rule. This domain specific pattern was shown in left PFC (~BA 45/46/9), left parietal (~BA 39/19/7), left temporal lobe (~BA 20/37), and bilateral extrastriate areas (~BA 17/18/19) favoring the verbal memory trials, compared to right PFC (~BA 44/6/9/10) regions, right parietal (~BA 40), right temporal lobe (~BA 20/37), and bilateral extrastriate areas (~BA 17/18/19) favoring the perceptual discrimination of faces (Table 3).

²It is important to note that any regions identified could reflect either the cognitive domain (memory versus perception) or the material domain (faces versus words) or their interaction. Since a key aim of the current study was to emphasize the domain generality of the LPFC response, this was an anticipated limitation of the design and we simply refer to this contrast as one of domain specificity.

Double Dissociation within Each Hemisphere: Rule by Domain Interaction across PFC regions

The data suggest that within prefrontal cortex, proximal dorsal and ventral LPFC regions may display qualitatively different response patterns with more ventral areas showing specificity to the domain and more dorsal regions specificity to the type of decision rule (Figure 3). To directly test this, we conducted separate ANOVAs on the mean response of the left and right PFC lateral regions implicated above, using factors of region (more dorsal vs. more ventral), rule (same-different vs. forced-choice), and material domain (verbal memory vs. face discrimination). For left PFC, as can be seen in Figure 3A and 3B, a double dissociation between more dorsal and more ventral regions was observed. Whereas more dorsal LPFC was sensitive to the decision rule ($F(1,11) = 23.41, p < .001$) showing greater activation of same-different than forced-choice decisions, it did not differentially respond across domain ($F < 1$). In contrast, the more ventral region demonstrated a domain specific response (verbal memory greater than face discrimination ($F(1,11) = 30.08, p < .01$), yet was relatively insensitive to the decision rule manipulation ($F(1,11) = 3.38, p = 0.09$) although there was a minor, possible trend towards greater activity during forced-choice compared to same-different tasks. For right PFC (Figure 3C & 3D), dorsal LPFC was again sensitive to the different decision rules ($F(1,11) = 33.69, p < .001$) yet insensitive to the task domain ($F < 1$). In contrast, a posterior and slightly more ventral region was sensitive to the task domain showing greater activation for the perceptual discrimination task than verbal memory task ($F(1,11) = 19.91, p < .001$). However, this region did not differentially respond across the decision rule types ($F < 1$).

Overall, these data demonstrate a separate double dissociation within each hemisphere for very proximal prefrontal regions. Consistent with prior literature, the more dorsal LPFC areas appear to respond independently of large differences in stimulus types or task domains, but are instead critically tied to the nature of the decision operations recruited by the task rule (Dobbins & Han, 2006; Fleck et al., 2006; Hayama et al., in press). In contrast, more ventral LPFC regions are sensitive to the domain of judgment or the nature of the stimulus materials, yet insensitive to the particular rule-related decision operations. While one or the other dissociation has been reported in the literature (see review by Badre & Wagner, 2007), this appears to be the first time that a separate double dissociation has been observed in each hemisphere within the same sample of subjects.

EXPERIMENT 2

Although the criterial classification account suggests that separately rendered classifications for each probe during same-different judgments governed LPFC activation, it could be also potentially argued that the same-different responding also requires that these intermediate judgments be temporarily held in mind so that the observer can ultimately determine whether the probes are in fact the same or different. To address this, Experiment 2 included a task that simply required the observer to overtly and independently classify each member of an array (*independent classification*). The independent classification of the items has no requirement for either maintenance or comparison of rendered decisions. Direct comparison of same-different tasks and independent classification tasks should reveal increased LPFC activation for the former if the need for working memory maintenance is the critical process, whereas the criterial classifications account would predict an identical response across same-different and independent-classification tasks. For efficiency, we used only the gender-judgment task.

Procedures

As in Experiment 1, for half of the runs, probes and decision questions were presented simultaneously. During these *simultaneous* runs we contrasted forced-choice, same-different, and independent classification decisions (2AIC– two-alternative independent classification)

for pairs of items. The latter decision rule simply required the subject to first classify the left item and then classify the right item and therefore differs from the other two in requiring two overt motor responses. Only two alternatives were used throughout these *simultaneous* display runs (2AFC, 2ASD, and 2AIC). There were 16 trials for each decision condition along with 16 resting trials for each run resulting in a total of 32 trials per condition of interest. In contrast to the *simultaneous* runs, separate runs were also conducted in which the onset of the probes of each pair were temporally offset. This latter manipulation was designed to further investigate any working memory maintenance requirement inherent in same-different decisions (e.g., Barch et al., 1997). In these *delayed* runs, we examined 2ASD and 2AIC conditions. On each trial, the first face, along with the task cue above, was presented for 3s, and then disappeared while the task cue remained for the following 6s. After this delay, the second face was presented for 3s and disappeared while the task cue remained for another 6s (18s SOA). In the 2ASD condition, the decision cue was “Different?” and subjects were asked to make a same-different gender judgment when they encountered the second face. For 2AIC condition, the task cue read “Female?” and subjects made a yes/no response for the first face gender classification (before the second face appeared), and then made a yes/no response for the second face (before the following trial started). Thus, subjects made two independent classifications for this condition, but there was no requirement to maintain or compare these classifications. The *delayed* and *simultaneous* runs alternated and there were four runs in total for each subject. All subjects began with the *delayed* run.

Results – Task Performance

As with Experiment 1, during simultaneous classification, the behavioral data clearly show that same-different responding is more difficult than forced-choice when the number of stimuli are matched across the two (See Table 4).

Results – fMRI Data

Simultaneous classification—To identify similar responses across the tasks (same-different, independent classification) we looked for regions demonstrating greater activation during 2ASD responding compared to 2AFC *and* greater responding during 2AIC compared to 2AFC, (each thresholded at the corrected level) (Figure 4). This map implicated many of the regions implicated in Experiment 1, including left LPFC (~BA 46/10)³, medial premotor (~BA 6/32) regions, as well as lateral premotor and posterior extrastriate areas (~BA 18/19) (Table 5). In both the anterior and more posterior cluster of the left LPFC response, the extracted time courses demonstrated virtually identical responses for the 2ASD and 2AIC tasks despite the fact that inclusion in the map required no such similarity of response (Figure 4A & 4B). Direct comparison of the mean response (6s to 8s) for the two conditions yielded t-values well below 1 for the anterior ($t(1,14) = 0.16, p > .87$) and posterior ($t(1,14) = 0.10, p > .91$) LPFC clusters. Furthermore, the LPFC activations are proximal and overlapping with those observed in Experiment 1. Figure 5A demonstrates this by examining the overlap of three contrasts, namely, same-different greater than forced-choice in Experiment 1 (both domains), same-different greater than forced-choice in Experiment 2, and independent-classification greater than forced-choice in Experiment 2 at a conservative threshold for each contrast (.003*.003*.003) (Figure 5A, top) and also at a more liberal threshold (.01*.01*.01) (Figure 5A, bottom). Overall, these results suggest that the greater left LPFC recruitment during same-different

³Anterior PFC (BA10) regions, especially frontopolar areas have been implicated in a wide variety of tasks suggesting a role in sub-goal processing, information integration, attention toward internally derived information, and prospective remembering, among others (e.g., Fletcher & Henson, 2001; Reynold, McDermott, & Braver, 2006; Braver & Bongiolatti, 2002; Smith, Keramian, & Christoff, 2007). Although the Same-Different decision type may require an integration process before arriving at a final response, there were no integration requirements in the Independent-classification task. Given that the Independent-classification and Same-Different responses were virtually identical, the activation pattern is more in accord with the classification account predictions. Although the anterior region in Figure 4 straddles BA10 and 46 areas, both regions are treated as lateral PFC given their highly similar response profiles.

judgment compared to forced-choice responding reflects the execution of two separate classification operations, and that maintenance of these operations in working memory is not a requirement for the elevated response.

Next, same-different and independent classification trials were directly compared to identify putative substrates for additional operations during same-different trials. This comparison yielded additional greater activation during same-different judgment in left LPFC, midline PFC areas, and right extrastriate cortex (Table 6; Figure 5B). Direct comparison of mean response (6s to 8s) across proximal LPFC regions (slice; $z = 29$) demonstrated region (middle frontal vs. inferior frontal) by condition (same-different vs. independent classification) interactions in LPFC regions ($F(1,14) = 7.95, p < .05$). However, this increased regional response is unlikely to reflect operations wholly unique to same-different classification because there was also prominent above baseline activation for forced-choice and independent classification tasks, albeit lower than during same-different judgments. Given this, the response may reflect more general working memory maintenance of stimuli, response options, or the mapping between the two. Alternatively this increased response may indicate the greater range of potential outcomes across the task. For example, with two items there are four potential outcomes of a same-different task (both old, both new, left old and right new, right old and left new). However, for a forced-choice task either left item is old and the right is new, or the reverse. If subjects keep such information in working memory then the same-different tasks would entail a greater maintenance demand.

Delayed classification

The primary analysis focused on whether regions demonstrating identical responses during independent-classification and same-different judgments compared to forced-choice in the simultaneous runs also showed identical response patterns even with the imposition of a delay period. We extracted time courses from ROIs defined during simultaneous presentation (see Figure 4) to determine whether these regions were also sensitive to an external delay manipulation. Figure 6 shows the extracted response for key regions, and suggests several general patterns. First, extrastriate visual areas (BA 18), while showing distinct responses to each stimulus onset, were not differentially affected by imposing a delay across 2ASD and 2AIC judgments (Figure 6A). In contrast, left LPFC regions that demonstrated identical responses during 2ASD and 2AIC during simultaneous presentation diverged considerably when a delay demand was imposed (Figure 6B). This tonic elevation of response was not shown in all regions of PFC as demonstrated by a more ventral and posterior PFC region, which similar to visual extrastriate regions, did not show a differential response across the tasks with the imposition of a delay (Figure 6C). Finally, Figure 6D shows a motor cortex ROI that demonstrates the expected pattern given the different motor requirements of the two trial types. Namely, this region shows clear responses during both stimulus onsets in the 2AIC, but only a single late response during the 2ASD judgment. Overall, the delayed runs demonstrated that LPFC is tonically engaged when subjects must hold decisions in mind across time. This requirement does not exist on the 2AIC trials and activation is generally depressed. Thus this region demonstrates identical responses during simultaneous trials for these two judgment types, yet responses diverge when a delay is imposed and so the region appears sensitive to both the number of criterial classifications required, and to the need to maintain these when it arises.

DISCUSSION

The current findings indicate that LPFC activation was closely tied to the number of classifications required by the decision rule, supporting the hypothesis that the activation reflects the execution of criterial classifications during the course of a trial (Dobbins & Han,

2006). In contrast, the activation neither tracked the number of relevant stimulus representations present, nor the behavioral difficulty of the trials (Lau, Rogers, Ramnani, & Passingham, 2004). Specifically, during Experiment 1 this region was insensitive to the level of stimulus competition and task difficulty during forced-choice responding. Furthermore, the region responded in a domain-general fashion across verbal memory and perceptual gender discrimination task reflecting a common amodal decision mechanism (Duncan, 2001). Experiment 2 further demonstrated equivalent LPFC responses for a task requiring two overt classifications (independent classification) and one requiring two covert classifications prior to the single overt response (same-different judgment). Consistent with Experiment 1, both tasks showed greater activity than judgments requiring only one classification act (forced-choice), further suggesting criterial classifications per se, whether overt or covert, were driving regional activation.

Despite the current findings, it is important to emphasize that our account is not meant to suggest that competition induced selection operations are not important. Indeed extant literature suggests more ventrolateral (Badre et al., 2005; Thompson-Schill et al., 1997) or midline prefrontal (e.g., pre-SMA)/anterior cingulate regions may play a key role in appropriate responding in the face of increased competition among representations or actions (Badre & Wagner, 2004; Kerns et al., 2004) although these regions were not affected by the number of probes in the current design. This null effect may reflect the fact that stimuli in the current design were not associated with pre-potent responses. Furthermore, frontal activation near premotor areas (~BA 6/8) has also been reported in several working memory studies manipulating the level of task difficulty (Chein & Fiez, 2001; Fockert, Rees, Frith, & Lavie, 2001). Critically, the current report identifies one specific mechanism that may render one task more difficult than another, namely, the number of intermediate criterial judgments that must be correctly rendered prior to task completion. Tasks may also differ in perceived or actual difficulty for other reasons, such as the number of stimuli competing for attentional or response selection, or the quality of those stimulus representations. The current study and previous research thus suggests that treating difficulty as a unitary explanatory construct is inappropriate and that performance can decline on a task for a host of separable reasons.

The criterial classification account is similar to the notion of “manipulation” used in working memory paradigms (D’Esposito, Postle, Ballard, & Lease, 1999). However, it potentially differs in that manipulation, like selection, is also often characterized in a stimulus representation dependent fashion, with greater manipulation assumed as stimulus numbers increase, or as stimuli become increasingly ambiguous. In contrast, the criterial classification account does not directly assume this and as shown in Experiment 1, performance can markedly decline and stimulus competition increase without greater LPFC recruitment. One way to explain previous findings under this criterial classification account is to assume that the tasks often studied under the rubric of manipulation are those which require the subject to execute numerous covert criterial judgments prior to responding. For example, alphabetic sorting of stimuli is often used during working memory studies (i.e., D’Esposito et al., 1999; Han & Kim, 2004) and such a task clearly requires a series of criterial judgments before arriving at the desired order of stimuli. Critically, although the number of criterial judgments and stimuli are confounded during sorting tasks, the current study demonstrates that they can be effectively disentangled.

Another domain in which LPFC activity has been implicated is individual differences in problem solving ability (Unterrainer et al., 2004). A general finding is that subjects with higher scores on tests that purportedly measure general intelligence (g) also tend to exhibit greater LPFC activity during demanding cognitive judgments (Gray, Chabris, & Braver, 2003). Similarly, during factor analysis, tasks that load highly on the construct of generalized intelligence evoke more LPFC activity than those that load lowly (Duncan et al., 2000). From

a selection-based perspective, these data would suggest that high “g” individuals excel in the ability to overcome competition, perhaps through top-down biasing of task relevant representations (Miller & Cohen, 2001) or that high “g” tasks are those that induce considerable competition among representations. Instead, under the criterial classification account these differences would be explained by the ability to render multiple, perhaps interdependent, intermediate classifications during complex decision-making. Under the criterial classification account, it is the tendency to exhaustively judge the relevant dimensions prior to responding that would separate the groups.

How might the criterial classification account also accommodate the data where no decision rule manipulation was in place? For example, Henson and colleagues (2000) suggested that as memory evidence approached the decision criterion, right LPFC would be increasingly engaged, reflecting the greater monitoring demands required under conditions of impoverished memory evidence. This is in essence a selection account in which there is increased competition or ambiguity between the evidence and the criterion representations as memory declines. The data confirmed the monitoring prediction demonstrating increasing LPFC activation with less confident judgments. To accommodate such findings, it is important to stress that because the criterial classification account assumes a direct link between activation and executing criterial judgments, it also predicts increased activation whenever subjects are likely to change the classification judgment about the criterial status of an item. In the case of difficult recognition judgment, one would expect that reclassification of items (albeit covert) would occur more often for low versus high confidence responses and thus the criterial classification account in fact predicts greater LPFC activity with decreasing confidence. This prediction would be easy to test by allowing subjects to overtly reverse their classifications during testing without penalty. The primary prediction would be that regardless of accuracy, trials with reversed responses would yield greater activity than single responses because the latter reflects two classification operations (See also Barber and Carter, 2004).

The above logic also applies to the frequently studied oddball task in which observers either count or overtly classify rare targets enmeshed in a train of non-target “standards” (Casey et al., 2001; Huettel & McCarthy, 2004; Kirino, Belger, Goldman-Rakic, & McCarthy, 2000). A typical finding is that target responses yield greater LPFC activation than non-target responses, and this difference is amplified as the targets become increasingly rare (i.e., Casey et al., 2001). Again however, this is precisely what would be predicted if reclassifications yielded more activation than single classifications, because rendering targets increasingly rare is likely to promote more erroneous initial classifications. Although researchers examining the oddball task exclude error trials, it is important to note that subjects can covertly reclassify an item prior to overtly responding and thus similar to the same-different judgment studied here, the critical classification operations can be covert.

Although the criterial classification account better explains the current data than accounts that rely on competition-induced selection, not all aspects of the data are necessarily accommodated. For example, there were no clear differences between same-different judgments rendered for arrays of two- versus three probes. If three probe arrays require more classification operations than two probe arrays one might expect increased activation under the model (Figure 2). However, this fails to take into account that it is only on a minority of trials that subjects should need to classify all three items in order to correctly determine that an array is “different”. Indeed, with random singleton placement subjects will only need to classify the last item on a third of the trials in order to correctly conclude that the response should be “different”. Unfortunately, given the numerous other factors in the present design, there were insufficient trials to break down the three probe arrays in this fashion and contrast those trials in which the singleton was in the final position versus those trials in which it was

in the first or second position. However, this discussion point serves as yet another illustration of the greater specificity of predictions possible under the criterial classification account.

Another potential confusion surrounding the criterial classification account might arise from the finding of the delayed condition in Experiment 2, where activation in LPFC region was elevated in the 2ASD versus 2AIC trials. In contrast, during the simultaneous condition the activation of 2ASD and 2AIC trials was equivalent in LPFC and this equivalence was used to argue that it was the number of criterial judgments that drove the region's response. If so, then why is activation elevated during the same-different conditions when a delay is interspersed between probes? Such concerns however reflect the tacit assumption that the region can only support one cognitive demand. In contrast, we assume that the region underlies the execution of classifications, and that the secondary need to maintain these classifications for substantial durations also yields a tonic increase in activation (along with other regions). Similar to others (Raye et al., 2007) we assume that the delay-related response reflects the elevated activation in a larger scale interactive network and that this up-regulation is the mechanism by which information about stimuli or rendered classification decisions is kept active. In the case of LPFC, this activity may allow observers to rapidly access what they have already concluded even in the complete absence of the representation that elicited the judgment.

In conclusion, the current data strongly weigh against competition induced selection accounts of LPFC during decision-making. Instead, the pattern of results in LPFC was predicted by assuming that LPFC activation would track the number of criterial classifications rendered during the course of a trial. Unlike selection accounts, which sometimes emphasize the need for cognitive control prior to executing a decision, the current account is more in line with LPFC accounts that stress the role of this region during decision actions (e.g., Lau et al., 2004), although under our account these decision actions can be covert with no motor requirement.

Acknowledgments







National Institute of Mental Health (MH073982 to I.G.D.)

References

- Badre D, Poldrack RA, Pare-Blagoev E, Insler RZ, Wagner AD. Dissociable Controlled Retrieval and Generalized Selection Mechanisms in Ventrolateral Prefrontal Cortex. *Neuron* 2005;47(6):907–918. [PubMed: 16157284]
- Badre D, Wagner AD. Selection, integration, and conflict monitoring; assessing the nature and generality of prefrontal cognitive control mechanisms. *Neuron* 2004;41(3):473–487. [PubMed: 14766185]
- Badre D, Wagner AD. Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia* 2007;45(13):2883–2901. [PubMed: 17675110]
- Barber AD, Carter CS. Cognitive Control Involved in Overcoming Prepotent Response Tendencies and Switching Between Tasks. *Cerebral Cortex* 2004;15(7):899–912. [PubMed: 15459082]
- Barch DM, Braver TS, Nystrom LE, Forman SD, Noll DC, Cohen JD. Dissociating working memory from task difficulty in human prefrontal cortex. *Neuropsychologia* 1997;35(10):1373–1380. [PubMed: 9347483]
- Braver TS, Bongiolatti SR. The role of frontopolar prefrontal cortex in subgoal processing during working memory. *Neuroimage* 2002;15:523–536. [PubMed: 11848695]
- Braver TS, Cohen JD, Nystrom LE, Jonides J, Smith EE, Noll DC. A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage* 1997;5:49–62. [PubMed: 9038284]
- Brett, M.; Anton, J.L.; Valabregue, R.; Poline, J.B. Region of interest analysis using and SPM toolbox. Paper presented at the the 8th International Conference on Functional Mapping of the Human Brain; Sendai, Japan. 2002.

- Budson AE, Dodson CS, Vatner JM, Daffner KR, Black PM, Schacter DL. Metacognition and false recognition in patients with frontal lobe lesions: the distinctiveness heuristic. *Neuropsychologia* 2005;43(6):860–871. [PubMed: 15716158]
- Bunge SA, Hazeltine E, Scanlon MD, Rosen AC, Gabrieli J. Dissociable Contributions of Prefrontal and Parietal Cortices to Response Selection. *Neuroimage* 2002;17(3):1562–1571. [PubMed: 12414294]
- Cabeza R, Dolcos F, Prince SE, Rice HJ, Weissman DH, Nyberg L. Attention-related activity during episodic memory retrieval: A cross-function fMRI study. *Neuropsychologia* 2003;41(3):390–399. [PubMed: 12457763]
- Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, Cohen JD. Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science* 1998;280(5364):747–749. [PubMed: 9563953]
- Casey B, Forman SD, Franzen P, Berkowitz A, Braver TS, Nystrom LE, et al. Sensitivity of prefrontal cortex to changes in target probability: A functional MRI study. *Human Brain Mapping* 2001;13(1):26–33. [PubMed: 11284044]
- Chein J, Fiez J. Dissociation of verbal working memory system components using a delayed serial recall task. *Cerebral Cortex* 2001;11:1003–1014. [PubMed: 11590110]
- D’Esposito M, Postle B, Ballard D, Lease J. Maintenance versus manipulation of information held in working memory: An event-related fMRI study. *Brain and Cognition* 1999;41(1):66–86. [PubMed: 10536086]
- Dobbins IG, Foley H, Schacter DL, Wagner AD. Executive control during episodic retrieval: multiple prefrontal processes subservise source memory. *Neuron* 2002;35(5):989–996. [PubMed: 12372291]
- Dobbins IG, Han S. Isolating rule- versus evidence-based prefrontal activity during episodic and lexical discrimination: a functional magnetic resonance imaging investigation of detection theory distinctions. *Cerebral Cortex* 2006;16(11):1614–1622. [PubMed: 16400153]
- Duncan J. An adaptive coding model of neural function in prefrontal cortex. *Nature Review Neuroscience* 2001;2(11):820–829.
- Duncan J, Seitz RJ, Kolodny J, Bor D, Herzog H, Ahmed A, et al. A neural basis for general intelligence. *Science* 2000;289(5478):457–460. [PubMed: 10903207]
- Endl W, Walla P, Lindinger G, Laluschek W, Barth F, Deecke L, et al. Early cortical activation indicates preparation for retrieval of memory for faces: an event-related potential study. *Neuroscience Letters* 1998;240(1):58–60. [PubMed: 9488174]
- Fleck MS, Daselaar SM, Dobbins IG, Cabeza R. Role of prefrontal and anterior cingulate regions in decision-making processes shared by memory and nonmemory tasks. *Cerebral Cortex* 2006;16(11):1623–1630. [PubMed: 16400154]
- Fletcher PC, Henson RNA. Frontal lobes and human memory. *Brain* 2001;124(5):849–881. [PubMed: 11335690]
- Fockert J, Rees G, Frith C, Lavie N. The role of working memory in visual selective attention. *Science* 2001;291:1803–1806. [PubMed: 11230699]
- Gray JR, Chabris CF, Braver TS. Neural mechanisms of general fluid intelligence. *Nature Neuroscience* 2003;6(3):316–322.
- Han SH, Kim MS. Visual search does not remain efficient when executive working memory is working. *Psychological Science* 2004;15(9):623–628. [PubMed: 15327634]
- Hayama H, Johnson J, Rugg MD. The relationship between the right frontal old/new ERP effect and post-retrieval monitoring: Specific or non-specific? *Neuropsychologia*. (in press)
- Henson RN, Rugg MD, Shallice T, Dolan RJ. Confidence in recognition memory for words: dissociating right prefrontal roles in episodic retrieval. *Journal of cognitive Neuroscience* 2000;12(6):913–923. [PubMed: 11177413]
- Hester R, D’Esposito M, Cole MW, Garavan H. Neural mechanisms for response selection: comparing selection of responses and items from working memory. *Neuroimage* 2007;34(1):446–454. [PubMed: 17071112]
- Huettel SA, McCarthy G. What is odd in the oddball task? Prefrontal cortex is activated by dynamic changes in response strategy. *Neuropsychologia* 2004;42(3):379–386. [PubMed: 14670576]
- Incisa della Rocchetta A, Milner B. Strategic search and retrieval inhibition: The role of the frontal lobes. *Neuropsychologia* 1993;31(6):503–524. [PubMed: 8341411]

- Kerns JG, Cohen JD, MacDonald AW III, Cho RY, Stenger V, Carter CS. Anterior Cingulate Conflict Monitoring and Adjustments in Control. *Science* 2004;303(5660):1023–1026. [PubMed: 14963333]
- Kirino E, Belger A, Goldman-Rakic P, McCarthy G. Prefrontal activation evoked by infrequent target and novel stimuli in a visual target detection task: An event-related functional magnetic resonance imaging study. *Journal of Neuroscience* 2000;20(17):6612–6618. [PubMed: 10964966]
- Lau HC, Rogers RD, Ramnani N, Passingham RE. Willed action and attention to the selection of action. *Neuroimage* 2004;21(4):1407–1415. [PubMed: 15050566]
- Levine B, Stuss DT, Milberg WP. Effects of aging on conditional associative learning: Process analyses and comparison with focal frontal lesions. *Neuropsychology* 1997;11(3):367–381. [PubMed: 9223141]
- Liu X, Banich MT, Jacobson BL, Tanabe JL. Functional dissociation of attentional selection within PFC: response and non-response related aspects of attentional selection as ascertained by fMRI. *Cerebral Cortex* 2006;16(6):827–834. [PubMed: 16135781]
- Macmillan, NA.; Creelman, C. *Detection theory: A user's guide*. New York, NY: Cambridge University Press; 1991.
- Mitchell KJ, Johnson MK, Raye CL, Greene EJ. Prefrontal cortex activity associated with source monitoring in a working memory task. *Journal of Cognitive Neuroscience* 2004;16(6):921–934. [PubMed: 15298780]
- Owen AM, Morris RG, Sahakian BJ, Polkey CE, Robbins TW. Double dissociations of memory and executive functions in working memory tasks following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Brain* 1996;119:1597–1615. [PubMed: 8931583]
- Raye CL, Johnson MK, Mitchell KJ, Greene EJ, Johnson MR. Refreshing: a minimal executive function. *Cortex* 2007;43(1):135–145. [PubMed: 17334213]
- Reynolds JR, McDermott KB, Braver TS. A direct comparison of anterior prefrontal cortex involvement in episodic retrieval and integration. *Cerebral Cortex* 2006;16(4):519–528. [PubMed: 16049191]
- Rowe JB, Passingham RE. Working memory for location and time: activity in prefrontal area 46 relates to selection rather than maintenance in memory. *Neuroimage* 2001;14:77–86. [PubMed: 11525340]
- Shallice T, Burgess PW. Deficits in strategy application following frontal lobe damage in man. *Brain* 1991;114:727–741. [PubMed: 2043945]
- Slotnick SD, Moo LR, Segal JB, Hart J Jr. Distinct prefrontal cortex activity associated with item memory and source memory for visual shapes. *Cognitive Brain Research* 2003;17(1):75–82. [PubMed: 12763194]
- Smith EE, Jonides J. Storage and executive processes in the frontal lobes. *Science* 1999;283(5408):1657–1661. [PubMed: 10073923]
- Smith R, Keramatian K, Christoff K. Localizing the rostrolateral prefrontal cortex at the individual level. *Neuroimage* 2007;36:1387–1396. [PubMed: 17532648]
- Thompson-Schill S, D'Esposito M, Aguirre G, Farah M. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *PNAS USA* 1997;94:14792–14797. [PubMed: 9405692]
- Unterrainer J, Rahm B, Kaller C, Ruff C, Spreer J, Krause B, et al. When Planning Fails: Individual Differences and Error-related Brain Activity in Problem Solving. *Cerebral Cortex* 2004;14(12):1390–1397. [PubMed: 15217897]
- Wager TD, Nichols TE. Optimization of experimental design in fMRI: a general framework using a genetic algorithm. *Neuroimage* 2003;18(2):293–309. [PubMed: 12595184]

	Which FEMALE?			Different Gender? (Yes / No)		Female? Female? (Y / N) (Y / N)	
	2AFC 	3AFC 	4AFC 	2ASD 	3ASD 	2AIC 	
# of Motor Responses	1	1	1	1	1	2	
# of Criterial Classifications	1	1	1	2	3	2	
# of Competing Stimuli	2	3	4	2	3	2	
	A. Forced-choice			B. Same-Different		C. Independent-classification	

Experiment 1 a, b, c, d & e for gender perception and verbal memory tasks

Experiment 2 a, d & f for gender perception task

Figure 1.

Examples of decision rules for perceptual gender judgments used in Experiments 1 and 2. Discussion of rules can be found in the text. Rows beneath boxes indicated how the tasks differed in motor responses, anticipated criterial classifications, and number of competing stimuli. AFC = Alternative Forced-Choice. ASD = Alternative Same-Different. AIC = Alternative Independent Classification.

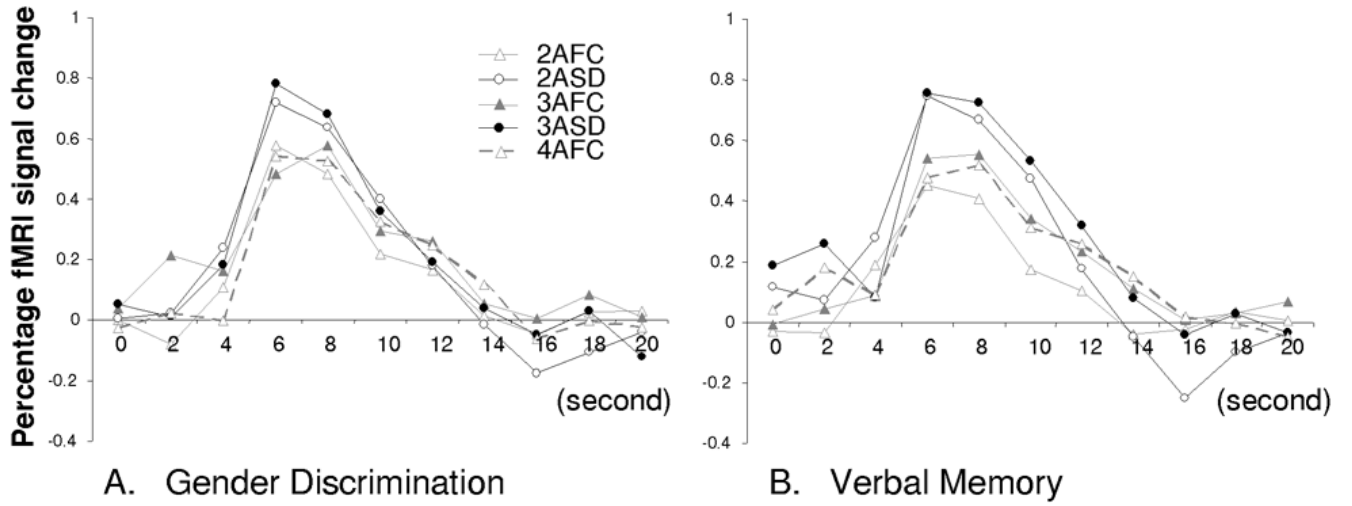
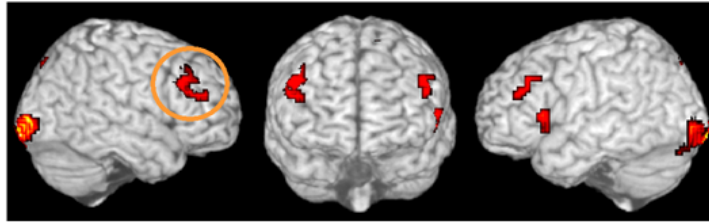


Figure 2.
Rule modulated activation in lateral PFC: Experiment 1

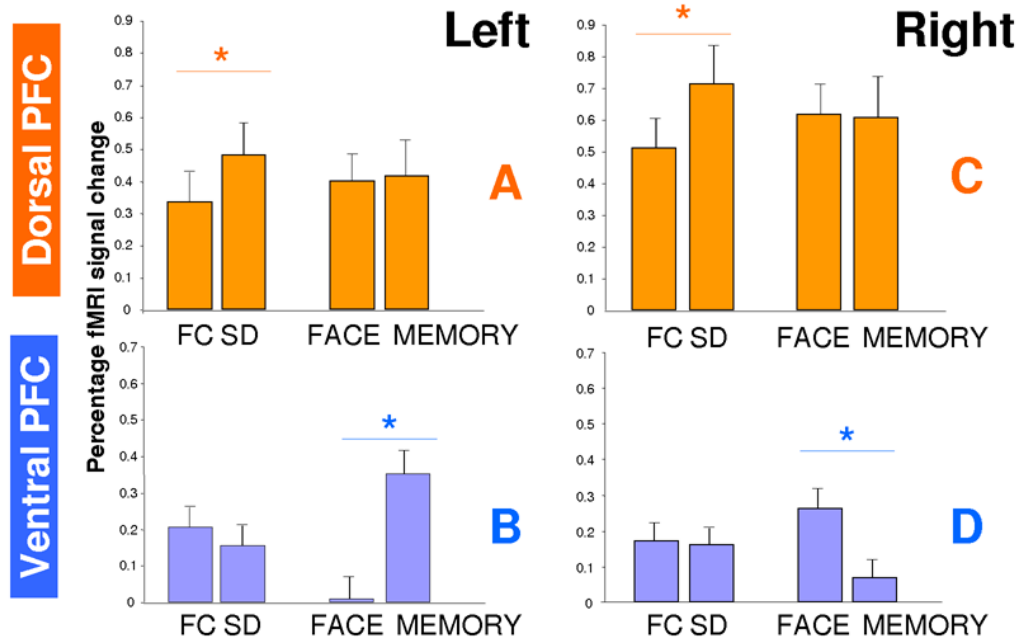
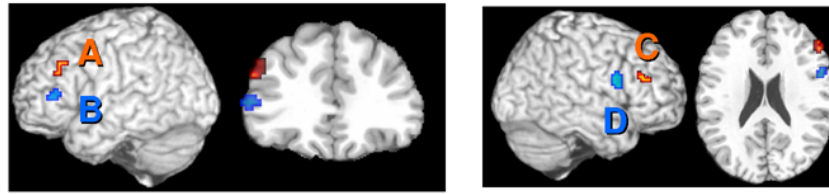


Figure 3. Double dissociation within each hemisphere: Experiment 1. Error bars represent standard errors of the means. FC = Forced-Choice, SD = Same-Different

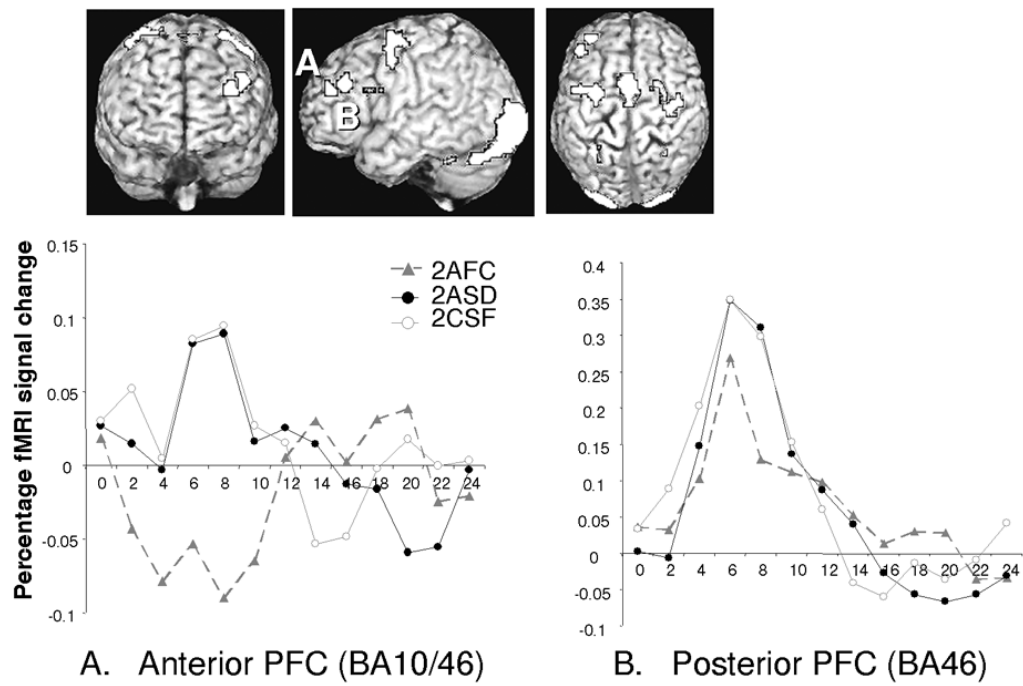


Figure 4. Lateral PFC activation for SIMULTANEOUS conditions: Experiment 2. Three dimensional overlay showing the conjunction of SD > FC and IC > FC. AFC = Alternative Forced-Choice, ASD = Alternative Same-Different, AIC = Alternative Independent-Classification

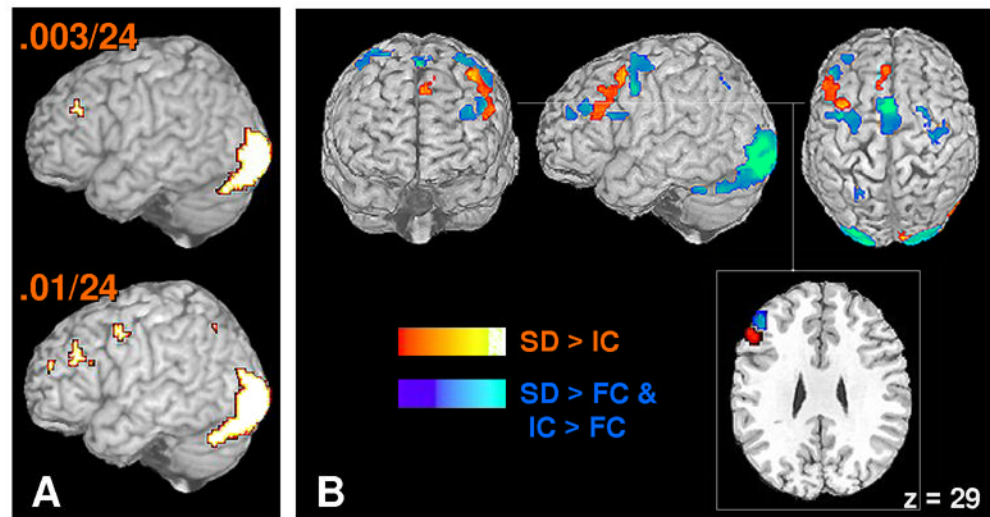


Figure 5. Conjunction and differential activation maps for same-different and independent classification tasks: Experiment 2, SIMULTANEOUS condition. (A) Conjunction between Exp.1 and 2 demonstrating greater activation in LPFC during same-different and independent classifications tasks in comparison to forced-choice judgments. See text for further information. (B) Differential activation between Same-Different versus Independent-Classification (RED), and the regions implicated in conjunction analysis (BLUE) showing the same level of activation for the tasks.

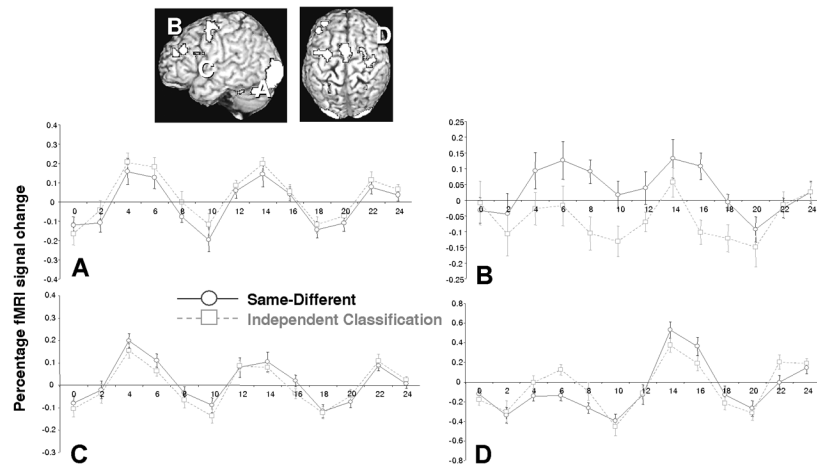


Figure 6. Experiment 2: DELAYED condition. Reconstructed hemodynamic responses for Same-different (solid lines) versus Independent Classification (dotted lines) for representative ROIs. (A) Extrastriate visual areas (BA18). (B) left lateral PFC (BA 46). (C) more posterior lateral PFC (BA 44). (D) right motor cortex (BA 6).

Table 1
Accuracy and reaction time for gender and memory discrimination, Experiment 1

Domain	Rule	Measure	Alternatives			
			Two	Three	Four	
Memory	FC	% Corr	0.87 (.08)	0.57 (.10)	.49 (.12)	
		RT	2390 (336)	2965 (473)	3133 (612)	
Perception	SD	% Corr	0.57 (.13)	0.57 (.12)	N/A	
		RT	3396 (302)	3401 (487)	N/A	
Perception	FC	% Corr	0.86 (.12)	0.84 (.10)	0.85 (.08)	
		RT	2071 (283)	2452 (383)	2754 (403)	
Perception	SD	% Corr	0.64 (.15)	0.75 (.16)	N/A	
		RT	2916 (324)	3419 (488)	N/A	

Note: Values in parentheses indicate standard deviations.

% Corr: Percent correct.

FC: Forced-Choice

SD: Same-Different

RT: Reaction times (ms)

Table 2
Regions demonstrating greater activation during same-different versus forced-choice verbal memory and the analogous contrast during perceptual gender discrimination

Region	BA	x	y	z	Vox.	Z score
Left hemisphere						
Extrastriate	18	-33	-93	-3	21	4.51
	17	-21	-102	-18	8	3.97
	18	-33	-93	-21	11	3.96
Medial Superior frontal	6	-9	12	57	21	4.2
Ventrolateral prefrontal	45	-57	21	6	20	4.18
Dorsolateral prefrontal	9/46	-48	33	33	15	3.94
	46	-51	39	24	11	3.81
Right hemisphere						
Extrastriate	18/17	21	-102	-12	13	5.72
	18/19	30	-96		24	5.47
	18	21	-102	-3	17	4.89
	18	27	-96	-9	23	4.66
	18	39	-93	-9	13	4.6
Dorsolateral prefrontal	46	57	36	24	13	4.15
	46	42	27	24	26	3.56
Dorsal frontal	9	45	27	33	31	3.86
	9	45	30	42	12	3.19
Parietal	7	15	-78	51	11	3.39
	7	18	-72	39	11	3.35

Note: BA, approximate Brodmann's locations. x,y,z corresponds to the MNI coordinates of the maximum voxel. Vox., number of significant voxels from SPM map within an 8-mm radius of maximum.

Table 3
Stimulus type dependent responses showing increases for Verbal memory versus Perceptual gender discrimination *and* Perceptual gender discrimination versus Verbal memory

Regions	Lat.	BA	x	y	z	Vox.	Z score
Verbal memory > Perceptual gender discrimination							
Extrastriate	L	19	0	-90	33	48	4.55
	L	17/18	-6	-78	9	50	4.44
	L/R	18/19	0	-81	27	74	4.38
Parietal	L	39	-36	-63	39	7	3.62
Superior parietal	L	19/7/39	-36	-72	42	4	2.93
	L	19/7	-33	-81	45	9	2.9
Midventrolateral	L	45/46	-54	39	6	17	3.4
Inferior temporal	L	20/37	-60	-57	-18	18	3.28
	L	20/37/21	-60	-42	-21	3	2.77
Dorsal	L	9	-54	12	33	4	3.24
Perceptual gender discrimination > Verbal memory							
Extrastriate	R	17	9	-93	-3	32	4.92
	R	17	12	-87	-9	55	4.81
	L	17/18	-6	-96	-3	24	3.85
	L	18	-18	-90	-12	27	3.81
	L	18	-9	-102	9	24	3.7
	L	18	0	-105	6	19	3.66
	R	18/19	18	-72	-9	15	3.53
Dorsal	R	19	18	-60	-12	10	3.22
	R	9	15	54	30	58	4.27
	R	9/6	51	3	30	5	3.26
Medial Frontopolar	L	10	-9	69	3	20	4.12
	R	10	9	54	3	45	4.1
	R	10	9	66	3	45	3.96
	R	10	12	66	18	57	3.9
Inferior Temporal	R	37/19	54	-66	-3	31	3.94
	R	37	-33	-48	-15	14	3.49
	R	19	39	-75	-6	2	2.81

Regions	Lat.	BA	x	y	z	Vox.	Z score
Ventral	R	44/45	57	12	21	22	3.94
	R	44	48	6	12	21	3.77
Parietal	R	40	45	-39	51	26	3.51
	R	40	54	-30	48	32	3.39
	R	40	39	-48	54	13	3.23
	R		27	-42	45	2	2.97
Fusiform	R	40	33	-45	48	8	2.94
	R	20/36	48	-45	-30	2	3.04
	R	20/37	57	-57	-18	3	3.04
	R	19	42	-87	24	2	3.01
	R	19/37	21	-51	-12	3	2.76
Middle Temporal	R	37/39	36	-72	21	11	3.11

Note: Lat. = Laterality

Table 4
Accuracy and reaction time for gender judgment, Experiment 2

Condition	Rule	Measure	Response	
			First	Second
Simultaneous	FC	% Corr	0.88 (.08)	N/A
		RT	1881 (290)	N/A
	SD	% Corr	0.64 (.15)	N/A
		RT	2468 (258)	N/A
	IC	% Corr	0.83 (.07)	0.84 (.07)
		RT	1984 (260)	773 (197)
Delayed	SD	% Corr	0.70 (.15)	N/A
		RT	1729 (336)	N/A
	IC	% Corr	0.81 (.09)	0.82 (.09)
		RT	1795 (277)	1529 (304)

Note: Values in parentheses indicate standard deviations.

% Corr: Percent correct

FC: Forced-Choice

SD: Same-Different

IC: Independent-Classification

RT: Reaction times (ms)

Table 5
Regions demonstrating greater activation during same-different compared to forced-choice, AND greater responding during independent classification compared to forced-choice

Region	BA	x	y	z	Vox.	Z score
Left hemisphere						
Superior Frontal Gyrus	6	-3	9	63	53	5.81
	6	-27	0	66	37	4.71
	6/8	-27	-9	54	33	4.09
Middle Frontal Gyrus	6/9	-39	3	39	22	4.75
	6	-48	3	48	31	4.46
		-39	18	24	26	4.31
	46	-45	36	30	25	4.20
	10	-33	48	24	27	3.59
Inferior Frontal Gyrus		-45	6	24	8	3.70
Middle Occipital Gyrus	18/19	-30	-93	0	78	5.24
Inferior Occipital Gyrus	18	-24	-90	-18	75	5.31
	18	-18	-90	-12	77	5.15
	18	-3	-90	-12	54	4.39
Fusiform Gyrus	19	-30	-75	-18	42	4.10
Precuneus	7	-21	-66	51	31	3.91
Cuneus	18	-18	-105	-6	51	4.79
	17	-6	-90	3	17	3.96
Parahippocampal Gyrus	28	-21	-27	-6	26	4.71
Caudate		-21	-33	12	7	3.73
		-18	-18	21	18	3.48
Ventral Lateral Nucleus		-12	-18	6	10	3.30
		-12	-9	12	7	3.17
Culmen		-3	-72	-18	26	4.62
		-36	-45	-30	30	4.56
		-27	-48	-33	23	4.30
Declive		-42	-63	-33	34	4.50
	19	-18	-75	-24	31	4.19
		-30	-63	-30	24	3.97

Region	BA	x	y	z	Vox.	Z score
Right hemisphere						
Medial Frontal Gyrus	6	6	-6	60	30	4.76
	32	9	21	33	12	3.63
Middle Frontal Gyrus	6	27	6	69	13	4.42
	6	45	-3	60	11	3.31
Precentral Gyrus	4/6	39	-15	57	23	4.77
	6	36	-15	72	16	4.63
	4/6	42	-21	69	7	4.41
Inferior Occipital Gyrus		24	-90	-9	79	4.48
	18	30	-93	-15	67	4.40
	17	6	-93	-12	45	4.50
	18/19	24	-78	-18	64	4.39
Cuneus	18	24	-99	-3	78	4.53
	17	3	-93	-3	31	4.16
Thalamus		15	-15	0	26	4.45
		21	-27	3	21	3.82
		21	-21	15	24	3.67
Culmen		42	-42	-36	15	4.23
		30	-45	-30	20	4.01
		36	-54	-30	14	3.73
Declive/Cerebellum		6	-75	-21	32	4.63
		45	-57	-30	14	3.79

Table 6
Regions demonstrating greater activation during same-different versus independent classification

Region	BA	x	y	z	Vox.	Z score
Left hemisphere						
Middle Frontal Gyrus	6	-39	9	54	37	3.67
	46	-51	27	27	31	3.25
	46	-42	27	24	19	3.09
	8/9	-45	24	42	23	3.06
Medial Frontal Gyrus	8/6	-3	39	42	28	3.27
	8	-6	24	51	8	2.98
Right hemisphere						
Middle Occipital Gyrus	19/37	39	-72	0	51	4.73
	19	54	-75	-9	34	3.44
	19/37	57	-72	3	21	3.07
Middle Temporal Gyrus	19/37	42	-75	9	14	3.16
Fusiform Gyrus	37	42	-66	-18	15	3.12
Cuneus	17	15	-99	-3	31	4.12
	18	12	-93	12	11	3.10