Neuron
Article

A Neural Representation of Prior Information during Perceptual Inference

Christopher Summerfield[1,*] and Etienne Koechlin[1,2]
[1]Institut National de la Santé et de la Recherche Médicale, Département des Etudes Cognitives, Ecole Normale Supérieure, 29, rue d'Ulm, Paris 75005, France
[2]Center for NeuroImaging Research, University Pierre and Marie Curie (Paris 6), Groupe Hospitalier Pitie-Salpetriere, Paris 75005, France
*Correspondence: summerfd@paradox.columbia.edu
DOI 10.1016/j.neuron.2008.05.021

SUMMARY

Perceptual inference is biased by foreknowledge about what is probable or possible. How prior expectations are neurally represented during visual perception, however, remains unknown. We used functional magnetic resonance imaging to measure brain activity in humans judging simple visual stimuli. Perceptual decisions were either biased in favor of a single alternative (A/∼A decisions) or taken without bias toward either choice (A/B decisions). Extrastriate and anterior temporal lobe regions were more active during A/∼A than A/B decisions, suggesting multiple representations of prior expectations within the visual hierarchy. Forward connectivity was increased when expected and observed perception diverged ("prediction error" signals), whereas prior expectations fed backward from higher to lower regions. Finally, the coincidence between expected and observed perception activated orbital prefrontal regions, perhaps reflecting the reinforcement of prior expectations. These data support computational and quantitative models proposing that a visual percept emerges from converging bottom-up and top-down signals.

INTRODUCTION

At the heart of psychophysical theory is the notion that there exists a metric linking a physical stimulus with internal percept (Stevens, 1957). Classical models of perceptual decisions incorporate this feedforward model, in that physical sensory input is serially accumulated toward a threshold or criterion for response (Ratcliff, 1978), presumably as stimulus information feeds forward through sensory neocortex and on to the higher brain (Heekeren et al., 2004; Kim and Shadlen, 1999). However, perception is also shaped by *reciprocal* interactions between an internal mental state and the external information impinging upon it. This reciprocity is evident in a range of neurocognitive phenomena in which evoked cortical responses are modulated by prior expectations or motivational states. For example, in the visual domain, prior contextual information is responsible for the elementary constancies that influence our perception of shape and color (Land, 1977), innate statistical biases in the processing of natural images (Kersten et al., 2004), as well as the effects of learning (Gilbert et al., 2001), local environmental context (Bar, 2004; Palmer, 1975), or task set (O'Craven et al., 1999) on the processing of visual objects and scenes.

Although it is well established that prior information influences perceptual inference, consensus has yet to emerge on how this finding should be incorporated into accounts of the decision process. Moreover, while neuroscientists have begun to characterize the forward flow of neural information during decisions with two visual alternatives (Heekeren et al., 2004; Kim and Shadlen, 1999), much less is known about the brain mechanisms that underpin the "top-down" influences on signal detection and object recognition (Frith and Dolan, 1997; Gilbert and Sigman, 2007; Kveraga et al., 2007b). Several quantitative and computational theories have proposed that backward (or "reentrant") connections within the ventral stream allow prior information to flow back to visual regions and guide object decisions as they unfold (Deco and Rolls, 2005; Friston, 2003; Grossberg, 1999; Mumford, 1992; Ullman, 1995). For example, one recent model describes how a percept evolves as sensory information travels through successive stages of a hierarchically organized cortical architecture (Friston, 2003; Mumford, 1992). Under this framework—known as "predictive coding"—the presence of specific prior information relevant to a perceptual decision allows the generation and representation of *conditional expectations* at multiple hierarchical levels in sensory neocortex—the most likely cause of observed sensory information given (or conditioned upon) that input. Expectation-related information is projected backward to the immediately preceding cortical level via reentrant pathways, such that forward-flowing sensory input is interpreted at each cortical stage within the context of a prior expectation (Figure 1A). Quantitatively, one can thus consider these representations (blue circles in Figure 1A) to be empirical priors for a Bayesian inference process occurring at the hierarchical level immediately below (Friston, 2005).

Predictive coding is useful because it allows us to formulate some clear hypotheses about how visual regions should behave during rudimentary decisions about sensory signals. Specifically, evoked neural responses (and effective connectivity) within the sensory neocortex should dissociate decisions informed by specific prior information from judgments that require an unbiased discrimination between two alternatives. To address this hypothesis, we devised a paradigm that draws

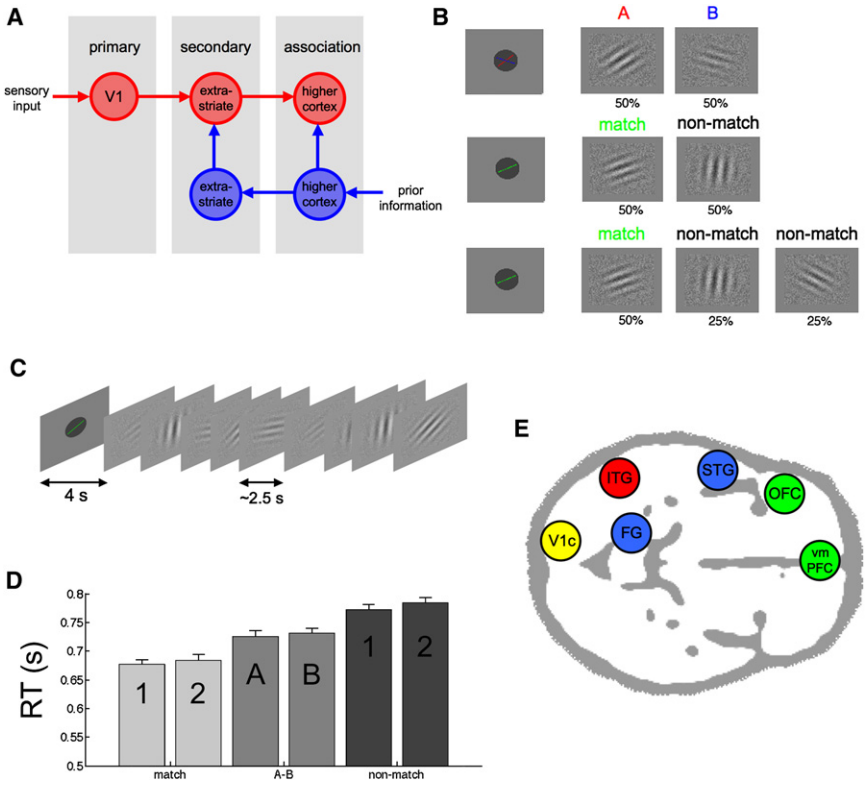


Figure 1. Theory, Paradigm, and Behavioral Data

(A) According to predictive coding, each hierarchical layer of a sensory system contains units for error prediction (red circles) and units for representing prior expectations (blue circles). Information flowing forward during perception (red arrows) evokes a response in error prediction units when it does not coincide with prior expectations. Information about expectations flows backward, gating sensory input at each cortical stage (blue arrows). Note similarities with Friston (2005).

(B) Stimuli. Instruction cues (left column) were gray circles crossed by either a red line and a blue line separated by 60° (A/B condition, top row) or a gray circle crossed by a single green line (A/~A condition, middle and bottom rows). Imperative stimuli were contrast-modulated Gabor patches embedded in visual noise. In the A/B condition, Gabors were always oriented either as the red line (50% trials) or the blue line (50% trials); in the A/~A condition, Gabors were either oriented as the green line (match trials, 50%) or with a 60° deviation from the orientation defined by the green line (nonmatch trials, 50%). In separate versions of the A/~A task, either one or two distracters could appear in a single block.

(C) Task sequence. The instruction cue was followed by nine imperative stimuli. Subjects responded with a button press to each stimulus.

(D) Reaction time data. Reaction times for match (light gray bars), A/B (gray bars), and nonmatch (dark gray bars) trials. For A/~A blocks (match and nonmatch), bars marked with 1 and 2 indicate the number of distracters. For A/B blocks, the response A marks that made with the same finger as a match on the A/~A blocks (counterbalanced across subjects). Error bars indicate SEM.

(E) Summary of regions and abbreviations. A large number of regions are discussed in this report. To help the reader understand our results, we include a schematic diagram of the main regions discussed in the text, indicated by their relevant abbreviation (V1c, primary visual cortex; FG, fusiform gyrus; ITG, inferior temporal gyrus; STG, superior temporal gyrus; OFC, orbitofrontal cortex; vmPFC, ventromedial prefrontal cortex). All regions except the vmPFC had corresponding partners on the opposite hemisphere, which are not shown. Locations are approximate. Colors indicate the contrast with which the region was defined: red, nonmatch > others; blue, A/~A > A/B; green, match > others; yellow, PPI connectivity with FG/MOG (see Results for details).

upon two tasks traditionally employed in visual psychophysics: the presence/absence (or "yes/no") judgment and the forced-choice categorization with two alternatives. Forced-choice discrimination (here, *A/B decisions*) has often been characterized as a race between two competing percepts that occurs without prior bias toward one or the other and that is decided in a winner-takes-all fashion at a downstream processing stage (Ratcliff, 1978). Yes-no judgments (here, *A/~A decisions*), by contrast, have been modeled as a biased decision process in which subjects confirm or disconfirm the match between an internal "perceptual template" and an external stimulus (Dosher and Lu, 1999). We view A/~A decisions as a special case of A/B decisions in which one perceptual alternative is represented in a privileged fashion, allowing perceptual inference to be reduced to the computationally less burdensome task of "matching" external sensory input to an internal template (Dayan et al., 1995). In our task, simple visual stimuli (oriented Gabor patches) were presented in two interleaved blocks, one of which (the A/B condition) forced participants to discriminate between two alternatives (i.e., *is the orientation A or B?*), whereas another (the A/~A condition) explicitly required subjects to match expected and observed information (i.e., *is the orientation A or not?*). Because the A/B and A/~A tasks were carefully tailored not to differ for bottom-up sensory input, we reasoned that comparing brain activity on these two decision tasks would reveal brain regions associated with the representation of conditional expectations—or, a "perceptual template"—during decisions about visual stimuli.

Predictive coding shares with many other theories the idea that evoked cortical responses depend not only on the physical nature of sensory stimulation but on the confluence of bottom-up sensory input with top-down signals encoding prior expectations. Specifically, it proposes that cortical foci falling within the relevant processing stream will respond more robustly when there is a mismatch between the observed sensory signal and the expected perceptual template; these population cortical responses can thus be considered *prediction error* signals, and a single solution for detection/recognition is only recovered once this prediction error has been jointly eliminated at all levels of the hierarchy (Friston, 2005). Here, we sought to identify prediction error responses in visual regions by comparing nonmatch (~*A*) trials, where expected and observed sensory information are at odds, to all other trial types. According to predictive coding, such responses should be observed in visual regions

responsive to the stimulus dimension being judged—in our case, angle of orientation (Friston et al., 2006).

Third, prior information during perceptual decisions should affect not only the height of the evoked response in the sensory neocortex but also the dynamic flow of neural information among visual regions. In particular, it has been argued that expectation-related information flows backward from higher to lower cortical stages, whereas prediction error responses flow forward (Friston, 2005). We subjected this view to empirical scrutiny using dynamic causal modeling (DCM) (Friston et al., 2003), which models the elements of a brain network as a system governed by causal mutual interactions with known inputs (experimental parameters) and outputs (evoked neural data). DCM allowed us to determine how expectation-related information flowed within the ventral stream during decisions informed by prior information (A/∼A decisions).

Finally, how does the brain respond when a "match" has occurred? Detecting a perceptual match is important not just to prompt a rapid and efficient behavioral response but also because it allows the incentive value associated with the expectation to be adjusted appropriately, and the decision criterion to be adjusted accordingly. We let this reasoning draw our hypotheses concerning perceptual matching toward portions of the ventromedial prefrontal (vmPFC) and orbitofrontal cortices (OFC) previously implicated in reward expectation (Kringelbach, 2005), decision making (Bechara et al., 1994; Wallis, 2007), and top-down visual perception (Bar et al., 2006b; Kveraga et al., 2007a; Summerfield et al., 2006). We thus tested our third hypothesis, that the vmPFC and/or OFC are responsible for *perceptual matching*—detecting the match between expected and observed sensation—by searching for voxels that responded uniquely to match (*A*) trials on A/∼A blocks.

In summary, we hypothesized that (1) we would observe neural *prediction error* and *conditional expectation* responses in visual regions, (2) that effective connectivity between these regions would be enhanced on decisions informed by prior information (A/∼A decisions), and (3) that neural correlates of *perceptual matching* would be observed in the ventral prefrontal cortex. Here, we confirm each of these three hypotheses, using voxel-wise, functional, and effective connectivity analyses of functional magnetic resonance imaging (fMRI) data obtained from healthy human subjects performing a simple perceptual decision-making task.

RESULTS

Both A/B and A/∼A blocks comprised an instruction cue followed by nine successive central, singly presented Gabor patches with variable orientation, embedded in visual noise (see Figures 1B–1D and Experimental Procedures). On A/B blocks, the instruction cue comprised two oriented lines; subjects discriminated whether the orientation of each subsequent Gabor patch was the same as line A (button 1; 50% of trials) or line B (button 2; 50% of trials). On A/∼A blocks, a single oriented line was shown at instruction; subjects determined whether each subsequent Gabor patch matched that orientation (button 1; 50% of trials) or not (button 2; 50% of trials). A/∼A blocks included target trials interleaved with either a single type of distracter trial (i.e., nonmatching Gabor) or two types of distracter trial. This manipulation prevented subjects from using the same discrimination strategy on A/B and A/∼A blocks. All orientations presented within a single block differed from each other by 60°. We additionally varied the contrast of the Gabor patches, allowing us to model the visibility of the stimuli (and thus the perceptual demand of the task) as a separate factor.

Behavioral Results

Evidence that simply instructing subjects that one of two stimuli was the "target" was sufficient to bias their judgments in favor of that stimulus is provided by reaction time (RT) data (Figure 1D). Although RTs were overall very well matched between A/B blocks (729 ± 74 ms) and A/∼A blocks with both one (725 ± 72 ms) and two (734 ± 78 ms) distracters (main effect of block, $p > 0.4$), match trials were reliably faster than nonmatch trials on A/∼A blocks, including both those with one distracter ($t_{(19)} = 8.05$, $p < 1 \times 10^{-7}$) and two distracters ($t_{(19)} = 8.09$, $p < 1 \times 10^{-7}$). Indeed, taking A/B blocks as an RT baseline, A/∼A match trials were about 60 ms faster (one distracter: $t_{(19)} = 3.83$, $p < 0.002$; two distracters: $t_{(19)} = 3.94$, $p < 1 \times 10^{-4}$), and nonmatch trials were about 40 ms slower (one distracter: $t_{(19)} = 3.31$, $p < 0.004$; two distracters: $t_{(19)} = 4.98$, $p < 1 \times 10^{-5}$) than responses on A/B trials.

Further analyses indicated that A/∼A blocks with one and two distracters did not differ on any of these comparisons (all p values > 0.1) nor was there a reliable interaction between number of distracters (n = 1, n = 2) and trial type (match, nonmatch) ($p > 0.5$). This prompted us to collapse across the two types of A/∼A block for subsequent analyses of imaging data. Errors remained relatively stable at ∼10% and did not differ between blocks or trial type (all p values > 0.1); neither did sensitivity calculated as d prime ($p > 0.1$).

fMRI Results

In the fMRI results described below, we only report regional effects if they survived correction for multiple comparisons (over the appropriate search volume) at a false discovery rate of 0.05. The p values for regional effects cited below pertain to the mean activity over voxels that were part of a significant cluster.

Expectation Representation

Comparing fMRI responses on A/∼A trials relative to A/B trials revealed a ventral stream network spanning early visual regions and the anterior temporal pole (Figure 2 and Table S1). Two regions falling within the occipital cortex exhibited larger responses on A/∼A than A/B trials, one on the superior/middle occipital gyri in Brodmann's area (BA) 18/19 (MOG; main effect of block: $F_{(1,19)} = 47.8$, $p < 1 \times 10^{-5}$) and another lying ventrally and medially on the fusiform/lingual gyri (FG; $F_{(1,19)} = 32.8$, $p < 1 \times 10^{-4}$). Within these clusters, match and nonmatch trials elicited responses that did not differ reliably (both p values > 0.2).

Outside of the visual cortices, preferential responses for A/∼A relative to A/B trials were identified in voxels falling close to the temporal pole (BA 38) bilaterally ($F_{(1,19)} = 25.47$, $p < 1 \times 10^{-5}$). This regions fell on the medial aspect of the superior temporal gyrus (STG). As for visual regions, no difference between match and nonmatch trials was observed here ($p > 0.1$).

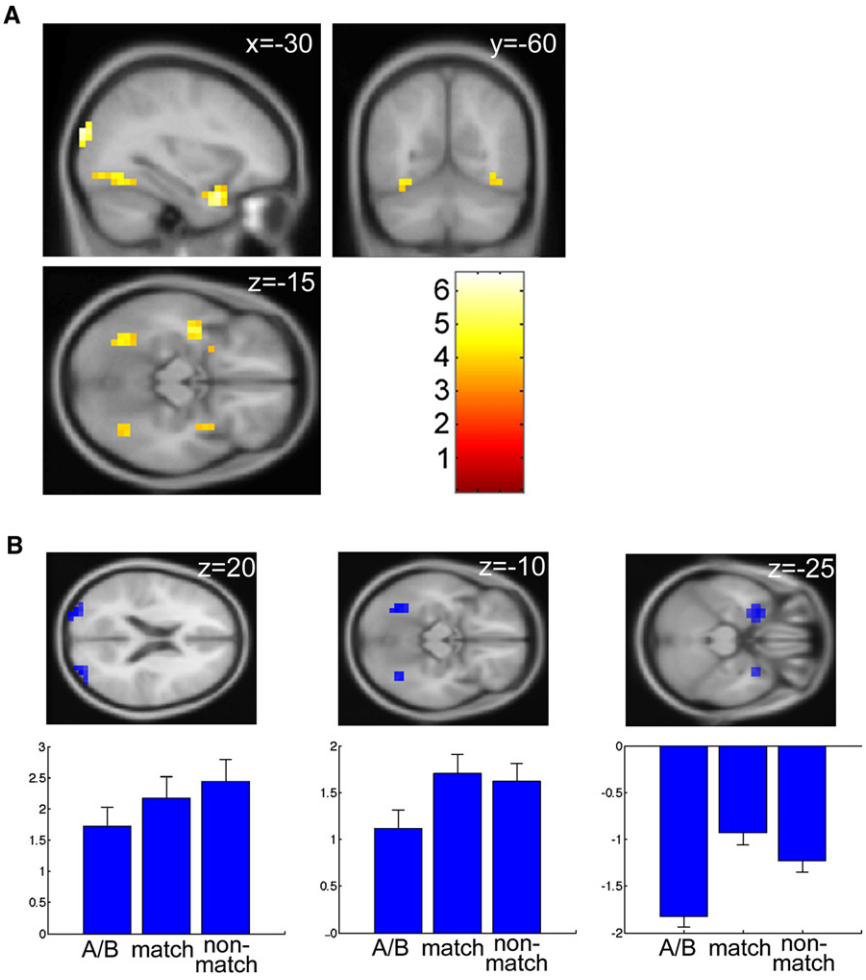


Figure 2. Brain Regions Involved in Representing Prior Expectations

(A) Voxels responding to the contrast A/∼A > A/B, thresholded at $p < 0.001$, uncorrected. The red-white scale refers to t values.

(B) Parameter estimates for A/B, match, and nonmatch trials averaged within ROIs corresponding to MOG (left), FG (middle), and STG (right). Error bars indicate SEM.

Prediction Error

We identified maximal prediction error responses bilaterally on the middle/inferior temporal gyrus complex (ITG), straddling BA 19 and 37 (ITG; Figure 3 and Table S2). In ITG, A/∼A nonmatches simultaneously elicited larger responses than match trials ($t_{(19)} = 4.54$, $p < 0.001$) and larger responses than A/B trials ($t_{(19)} = 4.29$, $p < 0.001$), whereas fMRI responses on match and A/B trials did not differ ($p > 0.6$). In the primary visual cortex (V1p), on the medial aspect of the lingual gyrus (responsible for representing the peripheral visual field), nonmatch trials similarly elicited larger fMRI responses than A/B trials ($t_{(19)} = 5.09$, $p < 1 \times 10^{-4}$) and match trials ($t_{(19)} = 3.24$, $p < 0.005$), which did not differ from one another ($p > 0.6$).

Additional visual prediction error responses were observed bilaterally on the dorsal aspect of the superior occipital gyrus, where BA 19 meets the parietal lobe (Figure 2). In this region, the advantage for nonmatch trials was largely carried by its difference from A/B trials ($t_{(19)} = 5.80$, $p < 1 \times 10^{-4}$); the comparison nonmatch > match achieved a more modest level of statistical significance ($t_{(19)} = 2.86$, $p < 0.01$). Indeed, unlike in V1 or ITG, the main effect of block (A/∼A trials > A/B trials) was also statistically reliable $F_{(1,38)} = 9.67$, $p < 0.006$), suggesting that this region was engaged by both prediction error and expectation representation.

Functional Interactions among Visual Regions during Perceptual Inference

Having identified a brain network for predictive visual inference that exhibited characteristic responses in the visual cortices and ventral stream, we turned to functional connectivity to characterize regional interactions during perceptual decisions of an A/∼A and A/B nature (Figure 4). Our analysis of functional coupling comprised two steps: first, we used a psycho-physiological interaction (PPI) analysis to identify cortical regions whose responses depended upon the activity of other reference regions and then assessed the directed effective connectivity among them using dynamic causal modeling.

Functional Connectivity (PPI Analyses)

Predictive coding allowed us to hypothesize that, during A/∼A relative to A/B blocks, information should flow from regions involved in expectation representation (e.g., MOG, FG) to regions demonstrating error prediction responses (e.g., V1, ITG). Accordingly, within the ITG ROI as defined above, A/∼A blocks led to significantly increased functional connectivity with seed regions placed at MOG ($t_{(19)} = 3.22$, $p < 0.005$) and FG ($t_{(19)} = 3.89$, $p < 0.001$). The specificity of this effect to ITG is illustrated in Figure 4A, where all voxels exhibiting functional connectivity with both MOG and FG are rendered onto the relevant axial slice ($z = 0$) at a joint threshold of $p < 0.005$, uncorrected. Although neither MOG nor FG showed increased connectivity with V1p (all p values > 0.2), a more posterior portion of V1 (involved in representation of the central visual field; here, V1c) was also jointly targeted by MOG and FG during A/∼A blocks, with high statistical reliability (Figure 4A).

Second, on nonmatch trials, increased information should flow between early visual regions involved in orientation representation (such as V1) and ITG, reflecting the higher weight afforded to bottom-up information when conditional expectations are not sufficient to explain the causes of sensation. We thus estimated a second PPI, searching across the brain for voxels whose correlation with a seed region planted at ITG increased on nonmatch trials relative to other trial types (Figure 4B). As

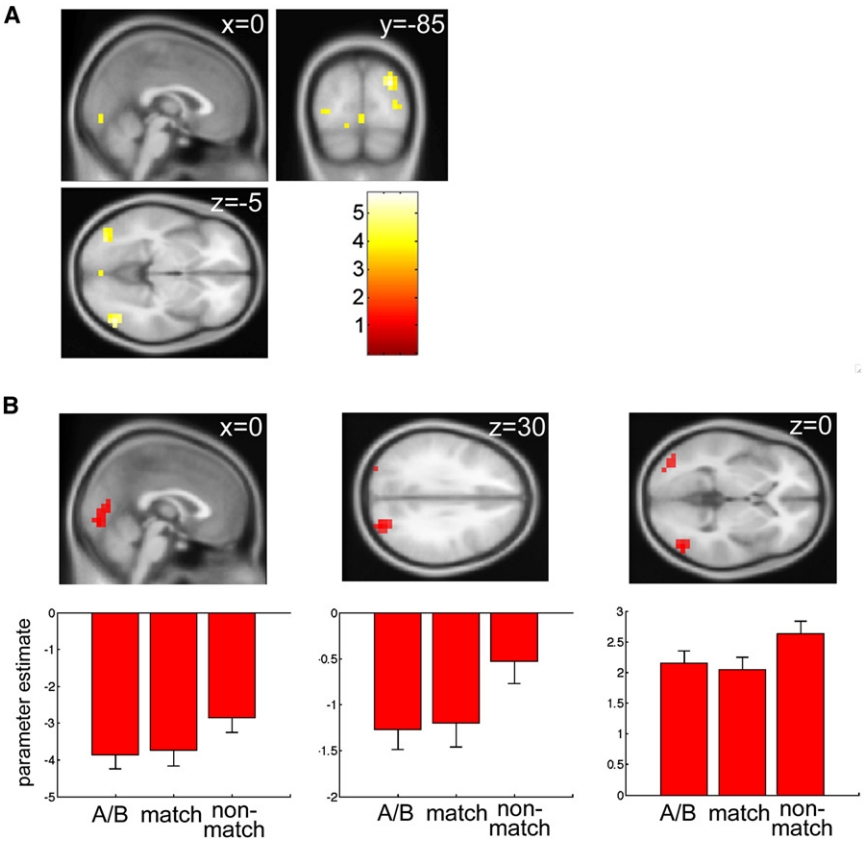


Figure 3. Neural Prediction Error Responses

(A) Voxels responding to the contrast nonmatch > (A/B and match), thresholded at $p < 0.001$, uncorrected. The red-white scale refers to the t value at each significant voxel.

(B) Parameter estimates for A/B, match, and nonmatch trials averaged within ROIs corresponding to V1p (left), sV3 (middle), and ITG (right). Error bars indicate SEM.

predicted, this analysis revealed increased connectivity between ITG and V1c ($t_{(19)} = 3.12$, $p < 0.006$).

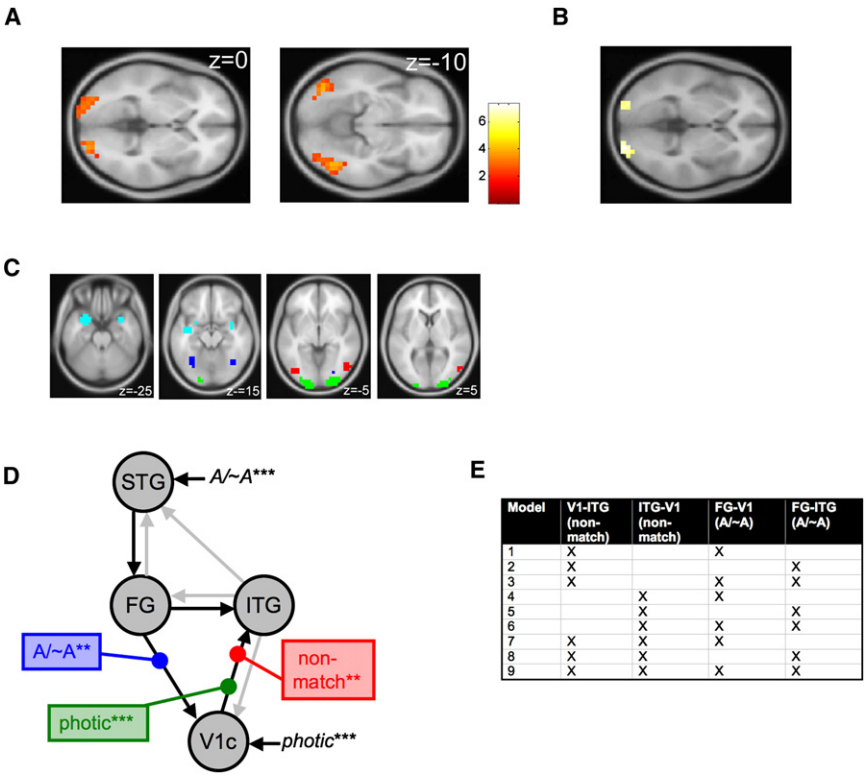
Forward and Backward Connectivity (DCM Analyses)

Predictive coding argues that whereas sensory evoked responses—reflecting prediction error—flow *forward* within the cortical hierarchy, information about prior expectations flows *backward*. Because PPIs do not carry information about the direction of flow of information between brain regions, we next turned to DCM, which draws upon concepts from systems theory to model how observed fMRI activations depend on both the experimental variables *and* the circulation of neural information within an anatomically plausible brain network (Friston et al., 2003). DCM calculates the statistical likelihood that an evoked response is driven by the *flow of information* from another brain area with which it is interconnected (monosynaptically or otherwise), leading to the derivation of modulatory parameters (bilinear terms) expressing the extent to which unidirectional effective connectivity between these regions varies as a function of stimulus inputs or experimental context. We equate information "flow" with the influence that activity in a source region has on the rate of change of neuronal states in a target region, encoded directly by the parameters of the DCM that are estimated below.

Building upon the PPI results, our DCM analyses posed two questions: first, does information flow forward or backward or reciprocally between V1c and ITG on nonmatch trials? Second, does expectation-related information flow from FG to V1c, or from FG to ITG, or both? (The selected regions of interest are shown in Figure 4C. We focused on FG because it yielded more robust effects in the PPIs and because the parameter estimates observed here were positive-going and thus more easily interpretable.) Combining these questions created a family of nine possible DCMs, and we used Bayesian model comparison to obtain statistical estimates of which model offered the optimum balance between simplicity and fit to the data (Penny et al., 2004).

We began with a standard model describing the influence of three parameters on effective connectivity: (1) simple visual stimulation, independent of condition (*photic*), (2) the main effect of the A/∼A task, relative to the A/B task (*A*/∼*A*), and (3) the effect due to A/∼A nonmatch trials, relative to other trial types (*nonmatch*). In the standard model, visual information (*photic*) entered the model at V1c, and expectation-related information (*A*/∼*A*) entered at the STG; intrinsic connectivity was specified in reciprocal, hierarchical fashion (between primary and extrastriate cortices, within the extrastriate cortex, and from the extrastriate cortex to the STG); forward modulatory connectivity from V1c to ITG increased with *photic*, and backward connectivity from the STG to FG increased with *A*/∼*A*. Nine model variants were created by building upon this standard model with combinations of the modulatory effective connections described above. The different patterns of modulatory connectivity (on a full reciprocal intrinsic connectivity) expressed in each variant is reported in Figure 4E.

Bayesian model comparison preferred the model in which nonmatch trials led to enhanced forward connectivity from V1c to ITG, whereas A/∼A trials led to increased backward connectivity from FG to V1c (Figure 4D and Tables S4 and S5). In this model, *photic* enhanced forward connectivity from V1c to ITG ($t_{(19)} = 4.68$, $p < 0.001$, 1.91% increase over baseline), A/∼A significantly enhanced backward connectivity from FG to V1 ($t_{(19)} = 2.67$, $p < 0.02$, 2.48% increase), and nonmatch increased connectivity from V1c to ITG ($t_{(19)} = 3.32$, $p < 0.004$, 2.23% increase). Intrinsic but not modulatory connectivity was significant between STG and V4 ($t_{(19)} = 3.27$, $p < 0.004$), and driving input to V1 and STG was, as expected, highly significant (all p values < 0.001). This model thus supports the view espoused by



E

Model	V1-ITG (non-match)	ITG-V1 (non-match)	FG-V1 (A/~A)	FG-ITG (A/~A)
1	X		X	
2	X			X
3	X		X	X
4		X	X	
5		X		X
6		X	X	X
7	X	X	X	
8	X	X		X
9	X	X	X	X

Figure 4. Functional and Effective Connectivity among Visual Regions

(A) PPI results 1: voxels exhibiting reliable increases in connectivity with *both* MOG and FG during A/~A blocks relative to A/B blocks, rendered at a joint threshold of $p < 0.005$, uncorrected. Voxels in the right panel overlap heavily with ITG.

(B) PPI results II: voxels exhibiting reliable increases in functional connectivity with ITG on nonmatch trials (relative to match and A/B trials).

(C) ROIs used for DCM analyses. V1c (green) was defined by its overlapping sensitivity to the PPIs described in (A) and (B). V1c exhibited highly significant positive-going responses on all trials that did not differ as a function of condition. ITG (red), FG (blue), and STG (cyan) are defined by the analyses described in Figures 2 and 3.

(D) The optimal DCM. BMC indicated that the optimal DCM, chosen from a space of nine models, was characterized by enhanced forward connectivity from V1c to ITG following all visual events (green box, photic) and nonmatch trials (red box) in particular. By contrast, backward connectivity from FG to V1c increased on A/~A trials (blue boxes). Intrinsic connectivity is signaled by black (significant, $p < 0.05$) or gray (nonsignificant) intrinsic connectivity. Driving input from photic to V1 ($p < 1 \times 10^{-12}$) and from A/~A to STG ($p < 1 \times 10^{-3}$) were also significant. Asterisks refer to significance for one-sample t tests conducted over subjects: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

(E) All DCMs. Each of the nine DCMs was characterized by a different pattern of modulatory connectivity. The winning model was model 1.

predictive coding that information about prior expectations (specific to A/~A trials) flows *backward* (e.g., from FG to V1), whereas nonmatch trials lead to increased *forward* flow of information from lower to higher regions (V1 to ITG).

Perceptual Matching

We predicted that the vmPFC/OFC would respond to perceptual matches, that is, would exhibit the largest responses whenever there was a coincidence between expected and observed perceptual information. Correspondingly, the vmPFC was more robustly activated by A/~A match than nonmatch trials ($t_{(19)} = 5.13$, $p < 1 \times 10^{-4}$) and by A/~A match than A/B trials ($t_{(19)} = 4.63$, $p < 0.001$), whereas nonmatch and A/B trials elicited similar responses ($p > 0.3$) (Figure 5). Statistical significance for the OFC was more modest (match > A/B trials, $t_{(19)} = 3.8$, $p < 0.002$; match > nonmatch trials, $t_{(19)} = 3.33$, $p < 0.004$; A/B > nonmatch, $p > 0.7$). The posterior cingulate cortex (PCC) responded in a similar fashion (match > nonmatch: $t_{(19)} = 4.37$, $p < 0.001$; match > A/B: $t_{(19)} = 4.36$, $p < 0.001$), with no difference for nonmatch > A/B trials ($p > 0.9$). Control analyses rule out task difficulty as an alternative explanation for the results obtained in vmPFC and PCC (Supplemental Results 1).

DISCUSSION

Psychophysical detection (yes-no) and discrimination (forced choice) judgments have historically been considered indices of a common decision process (Swets, 1964). However, we observed striking differences in the brain activity elicited by these two types of perceptual judgment. Brain regions within the visual cortex (MOG and FG) and at the apex of the ventral stream (STG) were more active during A/~A decisions, that is, judgments that were informed by specific prior information about forthcoming stimulation. Other visual regions showed *prediction error* responses and ventral frontal and medial parietal sites responded to *perceptual matching* when expected and observed sensory information coincide. Together, these results describe a network for "predictive" perceptual decision making that links the visual cortex with ventral prefrontal sites previously implicated in decision making and reward (Bechara et al., 1994; Rushworth et al., 2007; Wallis, 2007).

How, precisely, do decision strategies differ on A/~A and A/B judgments? Behavioral data revealed that, on A/~A blocks, match trials elicited very rapid responses, whereas nonmatch trials elicited slower responses relative to the baseline set by A/B responding. This phenomenon replicates the classically described RT advantage for "same" over "different" judgments in perceptual comparison tasks, known as the "fast-same" effect (Egeth, 1966; Farell, 1985). Conceiving of the decision as a diffusion (or "random walk") process in which sensory information in favor of one alternative or another is serially accumulated toward a decision threshold—a quantitative description of the decision process that has been successful in explaining both behavioral (Ratcliff, 1978) and neural (Kim and Shadlen, 1999)

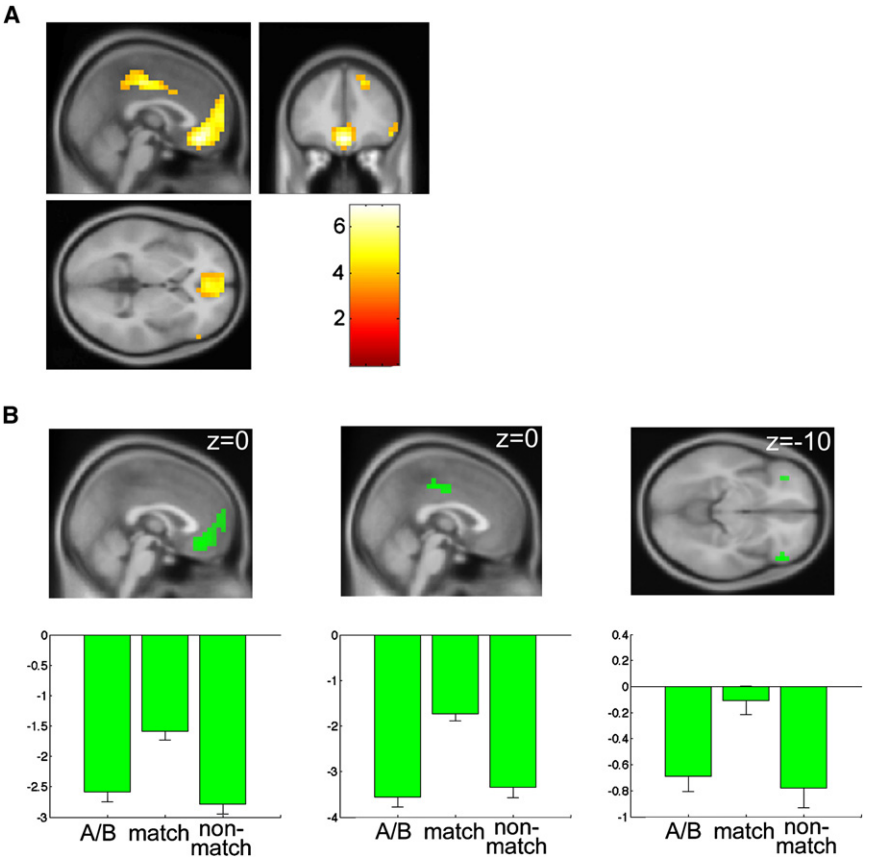


Figure 5. Brain Regions Involved in Matching Expected and Observed Perceptual Information

(A) Voxels responding to the contrast match > (A/B and nonmatch), thresholded at $p < 0.001$, uncorrected. The red-white scale refers to t values.

(B) Parameter estimates for A/B, match, and nonmatch trials averaged within ROIs corresponding to the PCC, vmPFC, and OFC.

Error bars indicate SEM.

data obtained in perceptual choice tasks—we consider two explanations for this effect. First, it could be that, on A/∼A blocks, the mere presence of prior information biases the origin of the random walk toward the threshold for "same" judgments (i.e., toward A and away from ∼A), rendering RTs for A shorter, and for ∼A longer, than for A/B judgments where the diffusion processing begins equidistant from the thresholds for A and B. By this account, because *conditional expectations* (reflecting the anticipation of one favored perceptual alternative) are already present in the visual system, less bottom-up sensory information has to be accumulated for perception to be completed and the decision threshold breached (Carpenter and Williams, 1995).

Another possibility is that the acquisition of sensory information is actively modulated by expectations, for example as processing in lower visual region is biased in favor of the conditionally expected target represented in higher regions (Friston, 2005). This top-down biasing in the visual hierarchy can be seen as a gain control mechanism, whereby "same" (or "A") information is accumulated more rapidly toward threshold once the diffusion process has begun, upon receipt of priors from higher stages of the processing hierarchy. Below, we discuss or fMRI data in the light of these two putative mechanisms of biased decision making, arguing that both processes may be occurring in our task.

Expectation Representation

Two extrastriate cortex regions were more active during A/∼A decisions than A/B decisions. Critically, these activations are not attributable to differences in the type or nature of sensory stimulation, as the train of stimuli presented in A/∼A and A/B blocks was carefully equated for its physical characteristics. Nor is this effect likely to be due to differences in the deployment of attention or the level of cognitive demand between the two conditions, as behavioral data indicated that—on average—accuracy and RTs were statistically indistinguishable between blocks. Our study thus differs from previous research in which attention was biased toward one task-relevant visual feature or dimension and away from another (Desimone and Duncan, 1995; Kastner et al., 1999) in that both A/∼A and A/B conditions required subjects to attend equally to the same stimulus dimension (angle of orientation). Instead, we argue that these MOG and FG activations reflect the privileged maintenance of prior information—in this case, an anticipated target orientation—against which each incoming sensory stimulus is compared. This view is consistent with the neurophysiological recordings demonstrating that, in the visual cortex, extrastriate regions including MOG and FG contain orientation-tuned neurons (Desimone et al., 1985; Pasupathy and Connor, 2002) and fMRI studies revealing orientation-specific adaptation effects in these regions (Boynton and Finney, 2003).

However, our connectivity analyses also revealed that direction-specific backward projections from FG to V1 were enhanced on A/∼A relative to A/B blocks. One interpretation of this finding is that these regions also contribute actively to the decision process by projecting back expectation-related signals to guide processing of orientation-related information in earlier visual regions, such as V1 (Hubel and Wiesel, 1968). A plausible anatomical basis for such backprojections has been described in the macaque (Felleman and Van Essen, 1991; Shipp and Zeki, 1989), and a causal role for reentrant connections between cortical stages in the formation of a visual percept has been previously demonstrated in humans (with TMS) and other primates (with regional cooling), where temporarily inactivating visual area MT has been shown to disrupt figure-ground segmentation, in the latter case by weakening center-surround interactions in hierarchically lower regions, V1, V2, and MOG (Hupe et al., 1998; Pascual-Leone and Walsh, 2001). These findings thus contribute to a growing literature emphasizing the importance

of reentrant connectivity in visual perception (Gilbert and Sigman, 2007).

Another consequence of our tight control over the sequence of stimulation in A/∼A and A/B blocks is that statistically both A and ∼A, and A and B were equiprobable, and thus the neural and behavioral biases observed on A/∼A blocks are not attributable to the brain tracking higher-order likelihoods of occurrence of a given stimulus. Rather, these biases probably occur because prior knowledge is available about the possible presence of a single stimulus exemplar. This bias may constitute an innate tendency to anticipate the persistence of a currently available percept, an ecologically plausible assumption in particular in the visual domain, where objects exhibit a gestalt constancy and sensory signals tend to be highly autocorrelated.

Although we argue that A/∼A blocks engendered a heightened expectation of a specific orientation, there are other processes that were probably common to both blocks. For example, both blocks required information about the appropriate response contingencies to be sustained. One possibility is that these more general short-term storage mechanisms are associated with the engagement of parietal and prefrontal cortex (PFC) regions known to be tonically active when information is held active across delay (Fuster, 1973). Indeed, fMRI activity that diverged strongly from the resting baseline but that did not differ between A/B or A/∼A conditions was observed in parietal and prefrontal regions (Figure S1 and Table S6). These regions may be particularly important for maintenance processes that guard the instructed target representation from potentially interfering information arising from the distracter stimuli or other sources (Sakai et al., 2002). By contrast, the sustained signals occurring uniquely during A/∼A decisions may be closely allied to that which accompanies mental imagery or during autobiographical reminiscence (Buckner and Carroll, 2007; Hassabis and Maguire, 2007), cognitive processes that may also require the creation and monitoring of a "generative model" of visual scenes or objects.

Prediction Error

During an orientation judgment task, we observed perceptual mismatch responses in ITG, a visual region known to be involved in processing simple stimuli defined by their orientation (Larsson and Heeger, 2006). We interpret these activations as reflecting "prediction error" signals occur in task-specific visual cortical regions, akin to those described in prefrontal and midbrain structures following the surprising presence or absence of a reward during reinforcement learning (O'Doherty et al., 2003). Prediction error signals might reflect the expanded neurocognitive processing needed to reconcile two sources of information: those arising from prior expectations and those from observation. Indeed, it has been argued that prediction error responses are ubiquitous throughout the brain, and the monitoring of sensory surprise may be a cardinal feature of many biological systems (Friston et al., 2006). Computational models have proposed plausible neural architectures giving rise to sensory error prediction responses, for example, those in which population cortical responses represent a posterior probability distribution (Deneve, 2008).

How might prediction error signals contribute to perceptual decision making? One possibility is that (consistent with the contention that on A/∼A blocks, the origin of a diffusion process is biased toward the match threshold) nonmatch trials simply require more information to be accumulated in order to confirm that a nonmatch has occurred and, consequently, more accompanying processing in extrastriate visual regions sensitive to the critical dimension (orientation). If ITG is participating in the accumulation of information about orientation, this would explain its enhanced responsivity to nonmatch over match trials. Our connectivity analyses additionally suggest that, where a nonmatch is detected, greater weight may be placed upon forward-flowing sensory evidence and less upon backward-flowing, expectation-related signals in order to rapidly reconcile disparities between what was anticipated and what was observed.

Previous studies have reported that, even when visual stimuli are presented rapidly or are degraded to optimally tax the perceptual decision system, neural signals distinguishing match and nonmatch trials can diverge as early as 150 ms following stimulus onset (Bar et al., 2006b; Thorpe et al., 1996). Perceptual decisions thus can occur very rapidly—perhaps too rapidly for a prolonged, iterative adjustment of predictions and evidence. Expectations thus might impinge on lower hierarchical stages in the form of a rapid "initial guess," which under normal viewing conditions might critically contain low spatial-frequency information, leaving perception of finer detail to bottom-up mechanisms (Bar et al., 2006b). Recent studies have raised the intriguing possibility that the ventral prefrontal cortex might be the origin of such signals (Bar et al., 2006b; Kveraga et al., 2007a).

The observed differences in the neural concomitants of match and nonmatch events complement an earlier study in which subjects viewed more complex visual objects, such as faces and buildings. Face-responsive portions of the fusiform gyrus exhibited stronger responses to faces during a face-matching task than a control task, i.e., *enhanced* match responses (Summerfield et al., 2006). Because predictive coding suggests that, during a hierarchical recognition process, activations at a higher levels "complete" lower-level perceptual signals with prior perceptual codes, it follows that match-suppression (prediction error) responses should be observed in earlier visual regions and match- enhancements in higher regions (such as those involved in the processing of facial identity). Further support for this view comes from studies demonstrating reduced fMRI responses in early regions and enhancements to those in later regions during perception of coherent relative to incoherent objects (Murray et al., 2002) and moving dot patterns (Harrison et al., 2007).

Perceptual Matching

Finally, in a third major result, we demonstrate that the ventral prefrontal cortex (including the OFC and the VMPFC) and the PCC respond to the match between expected and observed information. Control analyses demonstrated that these activations were not an artifact of match trials simply requiring less effort and/or shorter processing times. How, then, might these regions contribute to perceptual decision making?

Ventral prefrontal and medial parietal regions are known for their sensitivity to "old" over "new" items in recognition memory tasks (Wagner et al., 2005), a distinction that is very similar to that made between match over nonmatch trials here. Moreover, they are known to respond in a directionally signed fashion to the

surprising presence or absence of a reward that is predicted by an earlier cue (O'Doherty et al., 2003; Wallis, 2007), suggesting a role in processing the likely outcome encoded by a stimulus-stimulus association (Rushworth et al., 2007). One possibility is that the presence of a perceptual match—a "hit"—is a *reinforcer* in that it confirms the value and/or validity of the conditional expectations as a feasible internal model of the external world. Under conditions in which matches are scarce or absent, the perceptual template is a poor reflection of actual sensory information, and one would correspondingly expect it to be maintained less robustly, with concomitant behavioral effects (such as a reduction in the RT benefit for matches over nonmatches). One interpretation, thus, is that ventral PFC/ PCC activations on match trials may thus reflect the receipt of prediction error signals from visual regions, in the service of reinforcing internally represented expectations during perceptual matching.

One line of evidence favoring this view is that tasks that are most impaired following damage to the vmPFC/OFC tend to be those that benefit from the creation, monitoring, and updating of an internal or "model-based" representation of the external world. For example, state-based accounts of probabilistic reversal learning of visual discriminations outperform simple reinforcement-learning accounts in describing human behavioral performance on the task, and the strength of the "prior" associated with one model over another increases in a fashion correlated with vmPFC activity (Daw et al., 2006; Hampton et al., 2006). The vmPFC may thus be responsible for assigning value to internal representations, such as that for the perceptual "template" corresponding to the anticipated orientation in our study. The response on match trials may thus reflect a boost to subjects' confidence in their template upon receiving coinciding sensory input (Koechlin et al., 2002), which will in turn minimize the surprise associated with future repetitions of that sensory event (Friston et al., 2006).

Another interesting possibility is that the ventral prefrontal regions (in particular the OFC) contribute actively to the recognition process and are themselves the source of top-down signals responsible for guiding object recognition on the basis of expectations. Bar and colleagues have argued that incoming visual information is routed via orbitofrontal structures before being projected back to the extrastriate visual cortex, this backward connection offering an "initial guess" as to the identity of an object that consists principally of low spatial frequency information (Bar et al., 2006a). Although we did not include any of the ventral prefrontal ROIs in our DCM (to preserve model simplicity and because fMRI activity showed the opposite pattern to visual regions, making mutual interdependency in their evoked BOLD responses unlikely), we have previously shown that top-down connectivity from vmPFC to face-responsive voxels on the fusiform gyrus is enhanced by matching judgments about faces (Summerfield et al., 2006). Similarly, a DCM proposing top-down connectivity from the OFC was the best explanation of visual activations where magnocellular (low spatial frequency) information was emphasized (Kveraga et al., 2007a). Indeed, it is noteworthy that fiber bundles linking the ventral visual stream with the vmPFC are known to pass through the superior medial aspect of the anterior temporal cortex (Saleem et al., 2008), precisely where an additional cluster responding to prior visual information was observed in the current study. The possibility that the ventral PFC receives prediction error signals and/or intervenes directly in top-down object recognition is likely to constitute a fertile topic for future investigation (Kveraga et al., 2007b).

As is often observed in fMRI studies, activity in the vmPFC, OFC, and PCC deviated negatively from resting baseline, that is, although these regions were "less deactivated" by matches than nonmatches, signal levels were yet higher in the quiet rest periods between blocks (Gusnard et al., 2001). Although this observation does not necessarily alter the functional significance of the cognitive subtraction described above (Morcom and Fletcher, 2007), by way of explanation, we speculate that the higher activity levels observed in the vmPFC, OFC, and PCC at rest could reflect a constant fulfillment of sensory expectations that occurs whenever stimulus entropy is low and the unfolding sequence of sensory events is highly predictable—at rest, the brain is in a state of constant matching.

Although our study was carried out with simple, artificial stimuli, we believe that our results are relevant to a wider literature concerning the perception and recognition of complex objects. For example, the biasing mechanisms described above might come into play when perceiving objects in a congruent context. In support of this view, the PCC is more activated by objects with an unambiguous contextual association (Bar and Aminoff, 2003), leading to the suggesting that it contributes to the encoding of associative information about visual objects and scenes (Bar, 2007).

Conclusions

We report evidence that visual cortical regions represent expectations and exhibit "prediction error" responses during decisions of an A/~A type—those with an explicit matching component. Connectivity analyses suggest that prediction error responses flow forward and expectation-related information flows backward, allowing incoming sensory information to be constrained by prior information at each stage of the cortical hierarchy. These results offer empirical support for an emerging quantitative theory of perceptual inference (Friston, 2005). Finally, we argue that ventral prefrontal and posterior cingulate regions also contribute to perceptual inference by signaling the presence of a match between expected and observed perceptual information. Collectively, these data describe a cortical network for decisions informed by prior information that links visual regions with prefrontal cortical sites implicated in decision making and reward (Bechara et al., 1994; Rushworth et al., 2007; Wallis, 2007).

EXPERIMENTAL PROCEDURES

Subjects

Twenty neurologically normal individuals between the ages of 20 and 26 participated in the experiment. Subjects all had normal or corrected-to-normal vision and were recruited on campus at the Université Pierre et Marie Curie in Paris, France. Subjects all gave informed consent during an interview with our on-site physician and were paid 120 Euros for their participation.

Stimuli

All stimuli were generated and presented using PsychToolBox (Brainard, 1997) and appeared on a uniform gray background. The experiment consisted of 48 blocks of nine stimuli divided into four experimental runs of ~8 min. Each run

began with 10 s lead in and ended with ~20 s lead out. Each block began with an instruction cue for 3 s, followed by the gray background screen for a randomly determined interval of ~1.5 s (range 1–2 s). Instruction cues were darker gray circles crossed by a single green line (A/~A condition) or by a red and a blue line (A/B condition). A rest period of ~6 s (range 4–8 s) was interposed between blocks. Stimuli were Gabor patches (sine wave gratings enveloped by a Gaussian) of 1.6 cycles/°, subtending 3.8° visual arc, with nine equally distributed contrast levels ranging from 1.2% to 11.0%, added to a noise background. Gabor patches were presented for 1.5 s with an interstimulus interval of ~1.5 s (drawn from a uniform distribution with range 1–2 s). Subjects made an unprompted response during stimulus presentation. No feedback was given.

Design

A/B blocks (n = 24) and A/~A blocks (n = 24) occurred in alternation with the first block randomly determined in each run. On A/B blocks, orientation A was drawn from a uniform random distribution and differed from orientation B by exactly 60°. Subjects were instructed to memorize the instruction cue and respond with button 1 if the Gabor patch was oriented identically to the blue line and button 2 if it was oriented identically to the red line. On half of the A/~A blocks (n = 12), the stimulus train was identical to A/B blocks, i.e., two Gabor patches of orientation that differed by 60°. On the remaining A/~A blocks, two distracter Gabor patches, each differing from the target patch by 60°, were presented. Subjects were instructed to memorize the instruction cue and respond with button 1 if the Gabor patch was oriented identically to the green line and button 2 if it was not. On average across all A/~A blocks, half of the trials were A and half were not; on average across all A/B blocks, half of the trials were A and half were B. Responses were made with the index (left) and middle (right) finger of the right hand. The response keys used by the subject (index, middle finger) were fully counterbalanced within and between both conditions.

fMRI Data Acquisition

Magnetic resonance images were acquired with a Siemens (Erlangen, Germany) Allegra 3.0T scanner to acquire gradient echo T2*-weighted echo-planar images with blood oxygenation level-dependent contrast as an index of local increases in synaptic activity. The image parameters used were as follows: matrix size, 64 × 64; voxel size, 3 × 3 mm; echo time, 40 ms; repetition time, 2000 ms. A functional image volume comprised 32 contiguous slices of 3 mm thickness (with a 1 mm interslice gap), which ensured that the whole brain was within the field of view.

Behavioral Analyses

Behavioral data were analyzed with t tests and ANOVAs, with an α of $p < 0.05$. Post hoc comparisons between match and A/B trials, and nonmatch and A/B trials, were always made by comparing responses made with the same finger.

fMRI Analyses: Preprocessing

Imaging data were analyzed with SPM2 (Wellcome Department of Imaging Neuroscience, University College London, UK, (http://www.fil.ion.ucl.ac.uk/spm/spm2.html). Image treatment followed a standard preprocessing stream in which functional T2* images were slice-timing corrected and spatially realigned to the first volume acquired. The first five functional scans from each task were discarded prior to the subsequent analyses. Transformation parameters were derived from normalizing the coregistered mean echo planar image to a corresponding template brain within the stereotactic space of the Montreal Neurological Institute, and the derived parameters were then applied to normalize the remaining echo planar volumes for that subject. Normalized images were resampled at 5 × 5 × 5 mm and then smoothed with a Gaussian kernel of 10 × 10 × 10 mm full-width half-maximum. A 128 s temporal high-pass filter was applied in order to exclude low-frequency artifacts. Temporal correlations were estimated using restricted maximum likelihood estimates of variance components using a first-order autoregressive model. The resulting nonsphericity was used to form maximum likelihood estimates of the activations.

Conventional SPM Analyses

Conventional SPM analyses included seven task regressors: the instruction cues for (1) A/~A and (2) A/B; (3) match and (4) nonmatch trials for the A/~A blocks; corresponding (5) A and (6) B responses for A/B blocks, and (7) errors. All regression analyses also included a parameter encoding the mean signal from 1000 randomly selected voxels from the space outside the brain in a further attempt to eliminate scanner noise. SPMs were obtained at the second (between-subject) level with one sample t tests on the following contrasts: prediction error (Figure 2): [0 0 −1 3 −1 −1 0]; expectation representation (Figure 3): [0 0 1 1 −1 −1 0]; match detection (Figure 5): [0 0 3 −1 −1 −1 0]. SPMs in Figures 2A, 3A, and 5A are visualized at $p < 0.001$, extent threshold 5 voxels (~0.625 ml) uncorrected for multiple comparisons, but all clusters reported in the text survived false discovery rate (FDR) correction for multiple comparisons (Genovese et al., 2002). The clusters shown above each bar graph are at variable thresholds and are for display purposes only. The cluster maxima (Montreal Neurological Institute coordinate system) and associated FDR corrected p values are reported in accompanying Supplemental Tables. Statistical values reported in the text were obtained by extracting second-level contrast values from the mean of the relevant cluster from the SPM (thresholded at uncorrected $p < 0.001$) and performing conventional ANOVA or t tests on these data. A summary of the major regions discussed is shown in Figure 1E.

PPI Analyses

Seed voxels for MOG, FG, ITG, and more anterior brain regions were defined individually for each subject as the peak voxel sensitive to the appropriate contrast falling within the relevant cluster (defined at $p < 0.001$, uncorrected, at the group level). Using standard analysis techniques, the "physiological" time series extracted at this voxel was corrected for variance associated with parameters of no interest, deconvolved with the haemodynamic responses, multiplied by a parameter encoding the relevant "psychological" contrast (e.g., A/~A > A/B), and reconvolved to form a "psychophysiological interaction" (PPI) regressor. This regressor was entered into a design matrix alongside parameters encoding the main effects of the contrast and time series independently, as well as nuisance regressor encoding instructions and errors (i.e., regressors 1, 2, and 7, see above). Results are reported with an α of $p < 0.05$, uncorrected, within ROIs defined in an a priori fashion by conventional SPM analyses.

DCM Analyses

A new design matrix with six regressors was constructed for DCM analyses, with nuisance regressors 1, 2, and 7 corresponding to the instructions and error trials, plus three regressors of interest: *photic* (all trials), *A/~A* (all trials in A/~A blocks), and *nonmatch* (all nonmatch A/~A trials). New SPMs associated with these three regressors were used to define individual subject peaks for V1c, ITG, and FG, all of which nevertheless fell within the bounds of their corresponding cluster from conventional SPM analyses. Voxels of interest were extracted as for PPI analyses.

We constructed nine DCMs encapsulating variable patterns of forward, backward, and reciprocal connectivity among V1c, FG, ITG, and STG during *photic*, *A/~A*, and *nonmatch*. These models were then compared in a pairwise fashion using Bayesian model comparison in order to identify the model offering an optimal tradeoff between complexity and goodness of fit. Bayesian model comparisons calculates Bayes factors reflecting the optimality of one of two competing models, corresponding to the ratio of the probability of the data y given the model 1, or $p(y|m_1)$, to the probability of the data given model 2, or $p(y|m_2)$. Model evidence is adjusted according to Bayesian (BIC) and Akaike (AIC) criteria, the former of which favors simpler models and the latter more complex models: by an established convention, comparisons for which *both* of these Bayes factors exceed the exponential of 1 (~2.718) reflect positive evidence in favor of one model over another. At the group level, a group Bayes factor can be established by taking the product of minimum Bayes factor (AIC or BIC) across the cohort. Confirmation that the group Bayes factor is not driven by outlying data is offered by the "positive evidence ratio," the ratio of subjects displaying evidence (BF > 1) in favor of model 1 to those displaying evidence in favor of model 2 (BF < 1). In Supplemental Data, we describe all

nine models, along with their Bayes factors and positive evidence ratios and in pairwise competition with each of the eight other models.

SUPPLEMENTAL DATA

The Supplemental Data include tables, figures, and Supplemental Results and can be found with this article online at http://www.neuron.org/cgi/content/full/59/2/336/DC1/.

ACKNOWLEDGMENTS

We thank Eric Bardinet, Kevin Nigaud, Romain Valabregue, and Eric Bertasi for technical assistance and Tobias Egner, Hakwan Lau, and Jennifer Summerfield for comments on earlier versions of the manuscript. This work was supported by a European Young Investigator award to E.K.

Accepted: May 23, 2008
Published: July 30, 2008

REFERENCES

Bar, M. (2004). Visual objects in context. Nat. Rev. Neurosci. *5*, 617–629.

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. Trends. Cogn. Sci. *11*, 280–289.

Bar, M., and Aminoff, E. (2003). Cortical analysis of visual context. Neuron *38*, 347–358.

Bar, M., Kassam, K.S., Ghuman, A.S., Boshyan, J., Schmid, A.M., Dale, A.M., Hamalainen, M.S., Marinkovic, K., Schacter, D.L., Rosen, B.R., and Halgren, E. (2006a). Top-down facilitation of visual recognition. Proc. Natl. Acad. Sci. USA *103*, 449–454.

Bar, M., Kassam, K.S., Ghuman, A.S., Boshyan, J., Schmidt, A.M., Dale, A.M., Hamalainen, M.S., Marinkovic, K., Schacter, D.L., Rosen, B.R., and Halgren, E. (2006b). Top-down facilitation of visual recognition. Proc. Natl. Acad. Sci. USA *103*, 449–454.

Bechara, A., Damasio, A.R., Damasio, H., and Anderson, S.W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. Cognition *50*, 7–15.

Boynton, G.M., and Finney, E.M. (2003). Orientation-specific adaptation in human visual cortex. J. Neurosci. *23*, 8781–8787.

Brainard, D.H. (1997). The psychophysics toolbox. Spat. Vis. *10*, 433–436.

Buckner, R.L., and Carroll, D.C. (2007). Self-projection and the brain. Trends Cogn. Sci. *11*, 49–57.

Carpenter, R.H., and Williams, M.L. (1995). Neural computation of log likelihood in control of saccadic eye movements. Nature *377*, 59–62.

Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., and Dolan, R.J. (2006). Cortical substrates for exploratory decisions in humans. Nature *441*, 876–879.

Dayan, P., Hinton, G.E., Neal, R.M., and Zemel, R.S. (1995). The Helmholtz machine. Neural Comput. *7*, 889–904.

Deco, G., and Rolls, E.T. (2005). Attention, short-term memory, and action selection: a unifying theory. Prog. Neurobiol. *76*, 236–256.

Deneve, S. (2008). Bayesian spiking neurons I: inference. Neural Comput. *20*, 91–117.

Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. Annu. Rev. Neurosci. *18*, 193–222.

Desimone, R., Schein, S.J., Moran, J., and Ungerleider, L.G. (1985). Contour, color and shape analysis beyond the striate cortex. Vision Res. *25*, 441–452.

Dosher, B.A., and Lu, Z.L. (1999). Mechanisms of perceptual learning. Vision Res. *39*, 3197–3221.

Egeth, H. (1966). Parallel versus serial processing in multidimensional stimulus discrimination. Percept. Psychophys. *1*, 245–252.

Farell, B. (1985). Same-different judgments: a review of current controversies in perceptual comparisons. Psychol. Bull. *98*, 419–456.

Felleman, D.J., and Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex *1*, 1–47.

Friston, K. (2003). Learning and inference in the brain. Neural Netw. *16*, 1325–1352.

Friston, K. (2005). A theory of cortical responses. Philos. Trans. R. Soc. Lond. B Biol. Sci. *360*, 815–836.

Friston, K.J., Harrison, L., and Penny, W. (2003). Dynamic causal modelling. Neuroimage *19*, 1273–1302.

Friston, K., Kilner, J., and Harrison, L. (2006). A free energy principle for the brain. J. Physiol. (Paris) *100*, 70–87.

Frith, C., and Dolan, R.J. (1997). Brain mechanisms associated with top-down processes in perception. Philos. Trans. R. Soc. Lond. B Biol. Sci. *352*, 1221–1230.

Fuster, J.M. (1973). Unit activity in prefrontal cortex during delayed-response performance: neuronal correlates of transient memory. J. Neurophysiol. *36*, 61–78.

Genovese, C.R., Lazar, N.A., and Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage *15*, 870–878.

Gilbert, C.D., and Sigman, M. (2007). Brain states: top-down influences in sensory processing. Neuron *54*, 677–696.

Gilbert, C.D., Sigman, M., and Crist, R.E. (2001). The neural basis of perceptual learning. Neuron *31*, 681–697.

Grossberg, S. (1999). The link between brain learning, attention, and consciousness. Conscious. Cogn. *8*, 1–44.

Gusnard, D.A., Raichle, M.E., and Raichle, M.E. (2001). Searching for a baseline: functional imaging and the resting human brain. Nat. Rev. Neurosci. *2*, 685–694.

Hampton, A.N., Bossaerts, P., and O'Doherty, J.P. (2006). The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. J. Neurosci. *26*, 8360–8367.

Harrison, L.M., Stephan, K.E., Rees, G., and Friston, K.J. (2007). Extra-classical receptive field effects measured in striate cortex with fMRI. Neuroimage *34*, 1199–1208.

Hassabis, D., and Maguire, E.A. (2007). Deconstructing episodic memory with construction. Trends Cogn. Sci. *11*, 299–306.

Heekeren, H.R., Marrett, S., Bandettini, P.A., and Ungerleider, L.G. (2004). A general mechanism for perceptual decision-making in the human brain. Nature *431*, 859–862.

Hubel, D.H., and Wiesel, T.N. (1968). Receptive fields and functional architecture of monkey striate cortex. J. Physiol. *195*, 215–243.

Hupe, J.M., James, A.C., Payne, B.R., Lomber, S.G., Girard, P., and Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. Nature *394*, 784–787.

Kastner, S., Pinsk, M.A., De Weerd, P., Desimone, R., and Ungerleider, L.G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. Neuron *22*, 751–761.

Kersten, D., Mamassian, P., and Yuille, A. (2004). Object perception as Bayesian inference. Annu. Rev. Psychol. *55*, 271–304.

Kim, J.N., and Shadlen, M.N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. Nat. Neurosci. *2*, 176–185.

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

Kringelbach, M.L. (2005). The human orbitofrontal cortex: linking reward to hedonic experience. Nat. Rev. Neurosci. *6*, 691–702.

Kveraga, K., Boshyan, J., and Bar, M. (2007a). Magnocellular projections as the trigger of top-down facilitation in recognition. J. Neurosci. *27*, 13232–13240.

Kveraga, K., Ghuman, A.S., and Bar, M. (2007b). Top-down predictions in the cognitive brain. Brain Cogn. *65*, 145–168.

Land, E.H. (1977). The retinex theory of color vision. Sci. Am. *237*, 108–128.

Larsson, J., and Heeger, D.J. (2006). Two retinotopic visual areas in human lateral occipital cortex. J. Neurosci. *26*, 13128–13142.

Morcom, A.M., and Fletcher, P.C. (2007). Does the brain have a baseline? Why we should be resisting a rest. Neuroimage *37*, 1073–1082.

Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. Biol. Cybern. *66*, 241–251.

Murray, S.O., Kersten, D., Olshausen, B.A., Schrater, P., and Woods, D.L. (2002). Shape perception reduces activity in human primary visual cortex. Proc. Natl. Acad. Sci. USA *99*, 15164–15169.

O'Craven, K.M., Downing, P.E., and Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. Nature *401*, 584–587.

O'Doherty, J.P., Dayan, P., Friston, K., Critchley, H., and Dolan, R.J. (2003). Temporal difference models and reward-related learning in the human brain. Neuron *38*, 329–337.

Palmer, S. (1975). The effects of contextual scenes on the identification of objects. Mem. Cognit. *3*, 519–526.

Pascual-Leone, A., and Walsh, V. (2001). Fast backprojections from the motion to the primary visual area necessary for visual awareness. Science *292*, 510–512.

Pasupathy, A., and Connor, C.E. (2002). Population coding of shape in area V4. Nat. Neurosci. *5*, 1332–1338.

Penny, W.D., Stephan, K.E., Mechelli, A., and Friston, K.J. (2004). Comparing dynamic causal models. Neuroimage *22*, 1157–1172.

Ratcliff, R. (1978). A theory of memory retrieval. Psychol. Rev. *85*, 59–108.

Rushworth, M.F., Behrens, T.E., Rudebeck, P.H., and Walton, M.E. (2007). Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. Trends Cogn. Sci. *11*, 168–176.

Sakai, K., Rowe, J.B., and Passingham, R.E. (2002). Active maintenance in prefrontal area 46 creates distractor-resistant memory. Nat. Neurosci. *5*, 479–484.

Saleem, K.S., Kondo, H., and Price, J.L. (2008). Complementary circuits connecting the orbital and medial prefrontal networks with the temporal, insular, and opercular cortex in the macaque monkey. J. Comp. Neurol. *506*, 659–693.

Shipp, S., and Zeki, S. (1989). The organization of connections between areas V5 and V1 in macaque monkey visual cortex. Eur. J. Neurosci. *1*, 309–332.

Stevens, S.S. (1957). On the psychophysical law. Psychol. Rev. *64*, 153–181.

Summerfield, C., Egner, T., Greene, M., Koechlin, E., Mangels, J., and Hirsch, J. (2006). Predictive codes for forthcoming perception in the frontal cortex. Science *314*, 1311–1314.

Swets, J.A. (1964). Indices of signal detectability obtained with various psychophysical procedures. In Signal Detection and Recognition, J.A. Swets, ed. (New York, NY: John Wiley), pp. 164–171.

Thorpe, S., Fize, D., and Marlot, C. (1996). Speed of processing in the human visual system. Nature *381*, 520–522.

Ullman, S. (1995). Sequence seeking and counter streams: a computational model for bidirectional information flow in the visual cortex. Cereb. Cortex *5*, 1–11.

Wagner, A.D., Shannon, B.J., Kahn, I., and Buckner, R.L. (2005). Parietal lobe contributions to episodic memory retrieval. Trends Cogn. Sci. *9*, 445–453.

Wallis, J.D. (2007). Orbitofrontal cortex and its contribution to decision-making. Annu. Rev. Neurosci. *30*, 31–56.