FRONTOPOLAR SIGNALING OF INFERRED VALUE

by

Karin Cox

B.A., Macalester College, 2003

M.S., University of Pittsburgh, 2007

Submitted to the Graduate Faculty of

the Kenneth P. Dietrich School of Arts and Sciences

in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Karin Cox

It was defended on

November 30, 2011

and approved by

John Anderson, PhD, Departments of Psychology and Computer Science

Eric Donny, PhD, Psychology

Mark Wheeler, PhD, Psychology

Dissertation Advisor: Julie Fiez, PhD, Departments of Psychology, Neuroscience, and Communication Sciences and Disorders

FRONTOPOLAR SIGNALING OF INFERRED VALUE

Karin Cox, PhD

University of Pittsburgh, 2011

Adaptive decision-making requires that options be weighed according to the predicted value of the outcomes that they are likely to produce. Cognitive neuroscientific research has frequently emphasized the importance of the ventromedial prefrontal cortex (VMPFC) for encoding value information. As measured via functional MRI, VMPFC activity has been shown to correlate with the subjective values of explicitly-presented options (e.g., monetary gambles, offered foods) and also with the learned values of stimuli or actions that have been arbitrarily paired with rewards or punishments. Additional work has confirmed VMPFC sensitivity to abrupt changes in option values, particularly when these may be inferred on the basis of regular fluctuations in and/or interrelationships between option-outcome contingencies (e.g., as in serial reversal learning). The use of fixed option alternatives in these settings (e.g., constant choice stimuli) raises the question of whether such rapid updating of value activity can only occur when inferences can rely upon concretely-encoded sensory and motor information. To determine whether any brain region tracks values that may only be inferred via abstract rules, we scanned 17 participants as they completed a series of unique, two-trial discrimination problems in which the stimulus chosen on each Trial 1 was replaced with a novel stimulus of the same reward status on the following Trial 2. Under this protocol, optimal responding required inferences regarding the valence of each unsampled Trial 1 stimulus. Additional problem-wise manipulation of the magnitude of the available gains and losses created a situation in which the specific values of the unsampled stimuli could also be inferred. BOLD-signal analyses revealed activity in the right

lateral frontopolar cortex (FPC) that varied linearly with the inferable values of the unsampled Trial 1 stimuli; follow-up analyses confirmed that this effect was not attributable to the influence of the recently-delivered Trial 1 outcomes. Therefore, the results indicate that the brain does encode values that may only be inferred via abstract deduction. The frontopolar focus of this inferred-value activity corroborates and expands upon recent accounts of FPC function, which posit a broader role for this region in monitoring the value of re-directing behavior towards postponed response options.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

Adaptive decision-making requires that options be weighed according to their *value* – a measure that, within many theoretical frameworks, is proportional to the likelihood and magnitude of the gains or losses that the option is likely to produce (for reviews, see Glimcher, Camerer, Fehr, & Poldrack, 2009; Sutton & Barto, 1998). In some cases, the pending gain or loss may be obvious, due to the availability of either direct sensory evidence (e.g., the sight of an offered food) or a verbal indication of a potential outcome. In other cases, the values of initially-neutral choice alternatives may be learned through their repeated pairing with motivationally-significant consequences. Given the diversity of options which may inherently carry or ultimately acquire value, a major question for cognitive neuroscience concerns how the brain learns and encodes value information in all of its different possible forms.

In humans, functional MRI (fMRI) investigations of value processing have frequently focused on the ventromedial prefrontal cortex (VMPFC), a broad region that encompasses the inferior sector of the medial PFC and the medial sector of the orbital frontal cortex (Bechara, Damasio, & Damasio, 2000). In decision-making settings, blood-oxygenation-level-dependent (BOLD) signals within the VMPFC reliably track the value of a remarkable variety of choice options. During decisions between explicitly-offered food, non-food, or monetary rewards, VMPFC activation has been shown to correlate with the options' subjective values, with the latter estimated via value functions that have been derived from participants' choice behavior

(e.g., Chib, Rangel, Shimojo, & O'Doherty, 2009; Kable & Glimcher, 2007). VMPFC value signals also emerge as participants learn to associate either distinct motor actions (e.g., Glascher, Hampton, & O'Doherty, 2009; Wunderlich, Rangel, & O'Doherty, 2009) or choice stimuli (e.g., Glascher, et al., 2009; Wunderlich, Rangel, & O'Doherty, 2010) with experimentally-introduced reward rates of varying probability or magnitude.

The observation of VMPFC sensitivity to learned option values raises the question of how this form of value-signaling is acquired. During standard probabilistic discrimination (i.e., two constant options with independent, stable probabilistic return rates), the evolution of both choice behavior and VMPFC value signals can be adequately modeled with basic reinforcement learning (RL) algorithms (e.g., see Palminteri, Boraud, Lafargue, Dubois, & Pessiglione, 2009). Although the available RL algorithms vary in their specific computations, they frequently incorporate an iterative process that consists of (1) the sampling of an option, (2) a generation of a prediction (e.g., V_{option}) of the likely value to be gained from choosing that option, (3) the computation of an error signal δ proportional to the actual reward r minus V_{option}, (4) an adjustment of V_{option} by a quantity proportional to δ times a learning rate η, and (5) the weighting of the next trial's decision by the updated option values (see Daw & Doya, 2006; Sutton & Barto, 1998). The V_{option} term of the algorithm may then be used as a regressor to explain trial-by-trial variability in VMPFC responses to the actions or stimuli under evaluation.

When option values are learned in more structured tasks, the dynamics of VMPFC value activity may be better explained by sophisticated algorithms that exploit any structural information that may be available to the learner (e.g., interdependencies between, or reliable fluctuations in option values). The superior fit that may be obtained with these "model-based" RL algorithms – as contrasted with the "model-free" approach described above – was initially

demonstrated in an fMRI study by Hampton et al. (Hampton, Bossaerts, & O'Doherty, 2006). Hampton et al. were specifically interested in probabilistic reversal learning, a task which extends upon standard probabilistic discrimination by regularly switching the contingencies associated with the choice alternatives. In the specific reversal protocol used by Hampton et al., two fractal stimuli alternated in their status as the "correct" option (70%:30% probability of gain:loss) and the "incorrect" option (40%:60% gain:loss) at a relatively stable, moderate rate. Participants were alerted to the probabilistic and reversing nature of the contingencies, and advised to persistently choose the "correct" stimulus until they had inferred that a reversal had occurred. Additionally, a set of pre-scan practice trials provided exposure to the reversal rate and reward schedules that would be used in the scanner. The subsequently-acquired neuroimaging data were then used to evaluate the ability of a model-free and a model-based RL algorithm to predict VMPFC activity, with the assumption that trial-by-trial responses in this region should reflect the anticipated value of the to-be-selected stimulus.

The key qualitative difference between Hampton et al.'s (2006) model-free and model-based algorithms was illustrated through the divergent predictions that they made for trials that immediately follow a loss outcome. Within model-free RL, the delivery of a loss following the selection of a stimulus S_1 on trial t results in an incremental reduction in the estimated value V_{s1} during the transition to trial t+1 (assuming that $V_{s1}(t) > r(t)$). The estimated value of the unsampled stimulus S_2 remains unchanged. If the participant repeats the S_1 choice on trial t+1, then the predicted magnitude of VMPFC activity is relatively low, given the recent decrease in V_{s1}. If the participant switches to stimulus S_2 -- a choice that, given the instructions, is indicative of an inference that a reversal has occurred -- then the RL model still predicts a relatively low degree of VMPFC activity, since S_2 typically has not been sampled (and V_{s2} has not been

updated) since it was abandoned as a newly "incorrect" stimulus at the time of the preceding reversal.

In contrast to the model-free approach, Hampton et al.'s (2006) model-based RL algorithm took advantage of the available task structure information to simultaneously update the estimated values of both choice options. In the case of the specific algorithm that was employed (formally, a Bayesian Hidden Markov Model), the output that was used to predict VMPFC value activity was *p*(correct), the probability that the current "choice state" of either staying with or switching from the trial *t* choice was the "correct" strategy for trial *t*+1. The trial-by-trial calculation of *p*(correct) relied upon the combination of the trial *t* selection and reward information with the structural details learned during the pre-scan practice, including the reversal rate and probabilistic schedules, and, critically, the "anticorrelated" nature of the reward contingencies (i.e., as implied in the constant presence of one "correct" and one "incorrect" stimulus). The algorithm exploited the anticorrelation property by coupling negative adjustments in *p*(correct) for "staying" with positive adjustments in *p*(correct) for "switching", and vice versa. As a consequence, although both the model-based and model-free algorithms predicted decreased VMPFC activation when a loss was followed by a "stay" response, only the model-based algorithm predicted increased activation when a loss was followed by a response switch. The predictions of the model-based RL algorithm best matched the observed VMPFC activity, which exhibited increased as opposed to decreased response magnitude on the post-loss switch trials (see Hampton et al. Fig. 3A). This specific observation was corroborated by the superior quantitative fit of the model-based algorithm's predictions to overall VMFPC activity, even when fit assessments accounted for the greater number of free parameters utilized in this approach.

In summary, Hampton et al. (2006) demonstrated that VMPFC activity – at least within a reversal setting -- was best explained by a model that made use of known contingency interrelationships in order to infer value changes before they had been directly experienced. As the authors noted, their Hidden Markov Model was just one of many possible algorithms that would have generated such structure-based inferences. Consequently, much uncertainty remains with respect to the precise nature of the operations that ultimately gave rise to their VMPFC results. One important question concerns the degree of abstraction that the inference-guiding operations entailed. In the Hampton et al. model, abstract processing was assumed: The estimated values of the "stay" and "switch" options were encoded and adjusted in a manner that was independent of the specific stimuli to which they were being applied. However, post-reversal inferences are also predicted by models in which options are encoded more concretely. For example, Fuhs and Touretzky (2007) proposed a model of deterministic reversal learning in which the sensory and motor experiences of the task are categorized into two contexts, with each context corresponding to a stable pattern of environmental inputs (i.e., a single pair of stimulus-outcome contingency assignments). A key function of the model is to recognize ongoing experiences as consistent with one of the two contexts, which can in turn be used to predict the current value of each behavioral option. Extended to the probabilistic setting, a two-context model, like the Hampton et al. model, would predict increased value estimates and increased VMPFC activation during post-loss response-switching trials. In this case, the increase in predicted value follows from the fact that the negative outcomes received for the original selection are most consistent with the context in which the value of the alternative choice is high.

The distinction between the Hampton et al. (2006) approach and the two-context model is not trivial. Relative to the latter, concrete experience-based approach, any model that

incorporates an abstract encoding of the task structure is much more powerful in its inferential capabilities. This enhanced power would apply to any situation that involves the generalization of learned structural information to a new set of stimulus- and/or action options. As an example, consider a situation in which extensive training on a single pair of serially-reversing stimuli is followed by transfer to a new reversal problem that uses a new set of stimuli (see Rygula, Walker, Clarke, Robbins, & Roberts, 2010 for an example of the use of such a protocol with marmoset subjects). Models that rely on a concrete versus abstract encoding of the task structure would behave quite differently on the first occasion when the values of the new stimuli are reversed. From the perspective of a two-context model, the new circumstances would not match any experienced context; as a consequence, the model would have no basis for predicting the new stimulus values. In contrast, a model based on abstractly-encoded "switch" and "stay" options would be capable of applying previously-learned structural information to the new stimuli, and could thereby support rapid post-reversal adjustments.

Given the importance of abstract, rule-guided value predictions, the central goal of the present dissertation research is to determine whether activity in the VMPFC – or in any other brain region -- is sensitive to value information that may *solely* be inferred using abstract rules. This goal is accomplished by adapting a task from the non-human primate literature (Lockhart, Parks, & Davenport, 1963) in which optimal performance requires the repeated generation of inferences regarding *never*-sampled stimuli.

The original Lockhart et al. (1963) study was designed to explore the learning capabilities of monkeys that had acquired an object discrimination "learning set." In this context, the term learning set (LS) describes the rapid, often single-trial learning that highly-trained animals may exhibit on deterministic object discriminations (i.e., S+ and S- stimuli rewarded at rates of 100%

and 0%, respectively). In a typical learning set training protocol, monkeys complete hundreds of unique discrimination problems, each comprised of a new pair of S+/S- junk objects presented for two or more successive trials (for examples see Harlow, 1949; Levine, Levinson, & Harlow, 1959). The hallmark of learning set development is the gradual but ultimately dramatic increase in the slope of the single-problem learning curve, such that mean accuracy for the first non-chance problem trial (Trial 2) increases from near-chance to near-ceiling levels. Substantial behavioral and neural evidence suggests that this improved performance reflects a shift from simple associative to strategy-based learning mechanisms (Browning, Easton, & Gaffan, 2007; Deets, Harlow, & Blomquist, 1970; Murray & Gaffan, 2006). The purpose of the Lockhart et al. study was to determine the scope of the information that might be encoded during this strategic learning phase. The authors were specifically interested in the nature of the learning that might occur during the Trial 1-to-Trial 2 transition, at which point monkeys might develop accurate response tendencies directed towards either the chosen Trial 1 object (i.e., "win-stay/lose-shift") or the unchosen Trial 1 object (i.e,, avoid unchosen following a win, and approach unchosen following a loss).

To test this question regarding Trial 1 learning, Lockhart et al. (1963) created a task that selectively enhanced the behavioral relevance of either the initially-chosen or initially-unchosen problem stimuli (Fig. 1). The task consisted of 6-trial discriminations presented according to a largely standard protocol, which included control procedures designed to ensure that Trial 1 S+ and S- selections occurred with equal frequency. The principal experimental manipulation occurred on Trials 2-6, during which either the chosen or unchosen Trial 1 stimulus was replaced with a novel stimulus ("N+" or "N-") that adopted the value of the stimulus that it replaced. Replacement of either the chosen or unchosen stimuli was carried out as a between-subjects

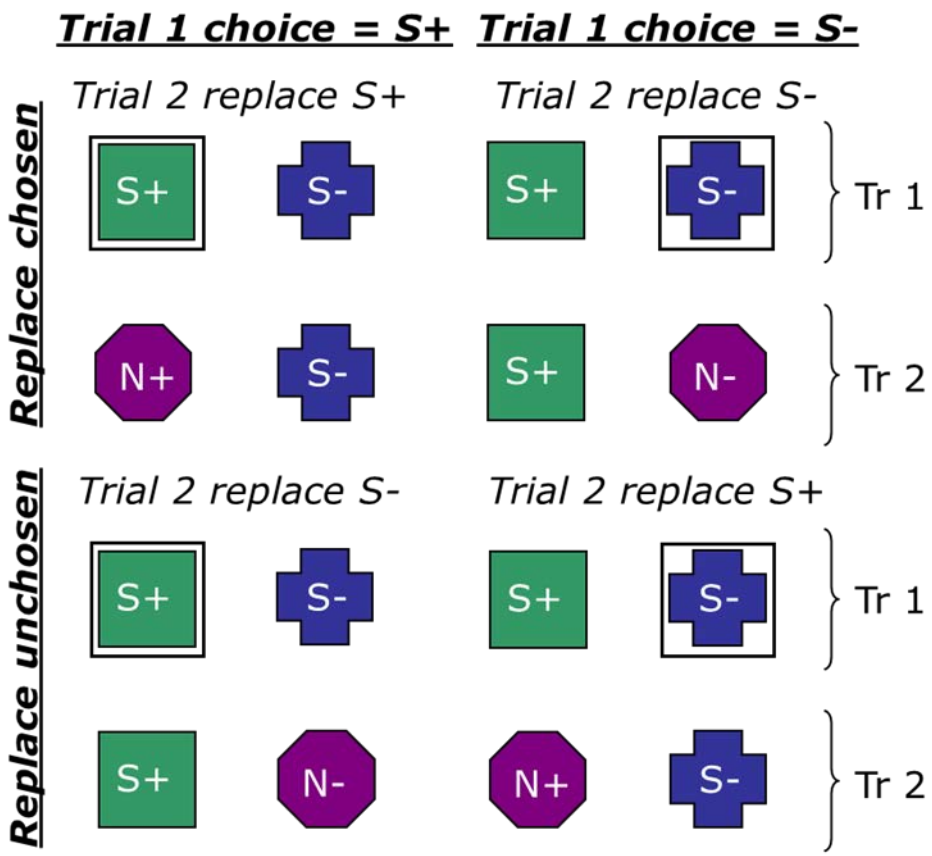


Figure 1. The original Lockhart et al. (1963) task design, Trials 1-2. Actual stimulus position was randomized across trials. Black outlines indicate the sampled Trial 1 stimuli. S/N +/- = original/novel rewarded/nonrewarded objects.

manipulation. Note that for subjects assigned to a "replace-unchosen" condition, optimal Trial 2 performance could be achieved via a strictly-defined application of the win-stay/lose-shift rule – in other words, the approach or avoidance of a previously-rewarded or -nonrewarded object. Under the "replace-chosen" protocol, it is impossible to execute a strict win-stay/lose-shift strategy; successful performance instead requires attention to and generation of an appropriate response towards the unsampled Trial 1 stimulus.

In their small sample of monkey subjects, Lockhart et al. (1963) only observed accurate Trial 2 responding when performance could be guided by sampled as opposed to unsampled objects. By the end of an 8-week testing period, the replace-unchosen group (who could rely on experienced contingencies) achieved ~80% Trial 2 S+/N+ selection rates, with accuracy levels

approaching ceiling across Trials 3-6. Although the replace-chosen subjects also demonstrated strong Trial 3-6 performance, their Trial 2 accuracy never reliably exceeded chance, regardless of whether performance was assessed within or across the two Trial 1 Selection conditions. Thus, at minimum, these results suggest that monkeys do not readily take advantage of the anticorrelation structure of deterministic discriminations in order to generate adaptive responses to unsampled stimuli, even when alterations in task structure encourage this kind of behavioral policy. In the case of the Lockhart et al. study, the replace-chosen subjects instead appeared to base their learning on the outcome of Trial 2, which was received for sampling a stimulus that would continue to appear on the remaining four trials.

Preliminary data from our lab suggested that human participants, when exposed to a computerized version of the Lockhart et al. (1963) task, can spontaneously acquire the optimal Trial 2 behavior in the replace-chosen condition (presented as a conference poster, Cox & Fiez, 2010). Therefore, in spite of monkeys' poor performance in this setting, the replace-chosen protocol provides a viable means for measuring humans' neural sensitivity to values that are inferred via abstract rules. In the present study, we refined our approach from the preliminary work to create both a pre-scan training sequence and an in-scanner test task that incorporated the critical replacement manipulation (see Fig. 2 for a depiction of the fMRI task).

All stages of the experiment were organized around a fundamental two-trial problem structure, with each problem consisting of (1) the free choice, on Trial 1, of either one of two uniquely-patterned "card" stimuli, (2) the delivery of a gain or loss of "points", depending on the predetermined status of the Trial 1 selection, (3) a second free choice on Trial 2, and (4) delivery of a second points outcome, depending on the Trial 2 choice and the active contingency assignments. During an initial, simple-discrimination phase of pre-scan training, the same pair

of stimuli was repeated across the two problem trials, with the goal of inducing our largely-uninstructed participants into acquiring a pattern discrimination learning set. Subsequent training problems incorporated the replace-chosen manipulation, such that each sampled Trial 1 stimulus was replaced with a novel card of the same value on the following Trial 2.

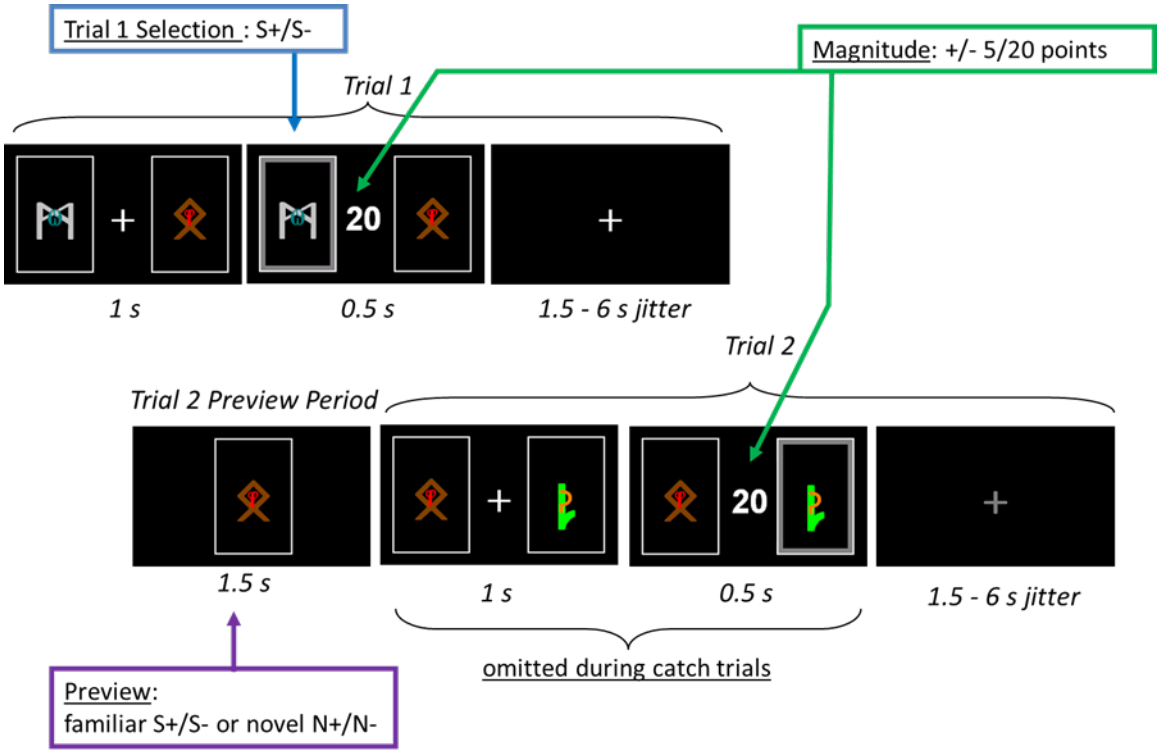


Figure 2. A sample two-trial problem from our human fMRI adaptation of the Lockhart et al. replace-chosen protocol (TR = 1.5 s). Arrows identify the three directly-manipulated factors (see Methods for full explanation). The three factors jointly determined the inferable value of the previewed stimulus. The example depicts the preview of an old, unchosen S- stimulus with a value of -20 points.

In view of our specific neuroimaging aims, the fMRI task was designed to facilitate the detection of hemodynamic signals that track with the inferable values of the unchosen Trial 1 stimuli. In this process, it was necessary to make three important modifications to the original Lockhart et al. (1963) task structure. First, to dissociate the BOLD responses that might be evoked by either of the two Trial 2 stimuli (i.e., the novel card and the old-but-unsampled card), we introduced a brief "preview period" consisting of the display of a single member of the card

pair immediately prior to the onset of Trial 2. Participants were explicitly instructed to refrain from responding during these preview periods. Old-unsampled and novel stimuli were previewed at equal rates, with the "old" stimuli representing the primary condition of interest. The novel previews were included both to prevent the success of any inference-generation rules that might be exclusively guided by the Trial 1 outcomes and to address any alternative interpretations of preview-period activity as directly reflective of these outcomes. We did not make any formal predictions as to whether the participants would spontaneously infer the values of the previewed novel stimuli, since adaptive behavior can be guided through reliance on the old stimuli alone.

Second, to allow for a more rigorous test of value-related signaling than would be afforded by a simple valence manipulation (i.e., S+ versus S-), we added a categorical, problem-level manipulation of outcome magnitude. Across each two-trial pair, the delivered gains and losses were either consistently "high" (20 points) or "low" (5 points), thus providing a basis for inferring the magnitude of the outcomes associated with the previewed stimuli. Third, to facilitate deconvolution of the hemodynamic responses evoked by the Trial 2 preview and Trial 2 proper, we applied a "catch trial" manipulation to a small percentage of the problems (Ollinger, Shulman et al., 2001). In this context, a catch trial referred to a preview period that was immediately followed by fixation as opposed to the corresponding Trial 2; points were neither earned nor lost for these trials.

Using this design, we sought the existence of brain regions in which preview-period activation was directly correlated with the values of the old-but-unsampled previewed stimuli. Such a response pattern would be consistent with the interpretation that a region signals value information that has been inferred via abstract rules. In light of the past neuroimaging research

described above, we anticipated that a primary locus of value-related signaling might be found in the VMPFC. Although our analysis differed from that of Hampton et al. (2006) in its inclusion of both to-be-chosen and to-be-unchosen stimuli, our prediction of VMPFC value activity across all "old" previews was supported by data suggesting that this region signals stimulus values regardless of ultimate selection status during "pre-decision" intervals (Wunderlich, et al., 2010).

Although our *a priori* predictions were limited to the VMPFC, other prefrontal regions have exhibited inferred-value activity in specialized settings, and therefore could be speculated to be potentially responsive to values that must be inferred via abstract rules. Two such regions are the lateral PFC (LPFC, including both dorsal and ventral sectors) and frontopolar cortex (FPC). Electrophysiological data from the monkey LPFC have suggested a potential role for this region in signaling "category-based" inferred values (Pan, Sawa, Tsuda, Tsukada, & Sakagami, 2008). In a task in which value information regarding the terminal member of a stimulus chain ("C" in "ABC") could be inferred to apply to the chain as a whole, the pairing of a new reward contingency with the terminal stimulus was reflected in subsequent LPFC responses to the chain-initial stimulus ("A"). This observed LPFC sensitivity to higher-order structure suggests that value activity in this region might be similarly influenced by other, more abstract task rules. With respect to the FPC, recent fMRI evidence (Boorman, Behrens, Woolrich, & Rushworth, 2009) raised the intriguing possibility that this region specifically tracks value information related to recently-*unchosen* options. In a task in which left- and right-button presses were rewarded according to independent, randomly- and gradually-drifting probabilities, Boorman et al. found that FPC activity tracked the difference in the estimated reward probabilities associated with the trial-by-trial unchosen and chosen options. A finding of such FPC relative unchosen-value activity in an abstract rule-bound context would be consistent with models that place this

region at the apex of a rostrocaudal, abstract-to-concrete processing gradient within the PFC (Badre, 2008).

1.1 METHODS

1.1.1 Participants

Healthy adults between 18-35 years of age were recruited from the University of Pittsburgh and surrounding community. Prior to enrollment, prospective participants were screened according to the inclusion criteria (right-handedness, MRI compatibility, English proficiency) and the exclusion criteria (self-report of learning or reading disabilities, psychiatric or neurological illness, or usage of psychoactive medication). Thirty-eight participants were initially enrolled in the study, which consisted of two experimental stages (behavioral training, fMRI acquisition) scheduled to occur on two consecutive days.

A subset of the enrolled participants were excluded from the final dataset due to problems that occurred during either the behavioral or fMRI stages. During the behavioral training stage, participants were excluded on the basis of either unsatisfactory performance (n = 6, see section 1.1.3), voluntary withdrawal (n = 1) or software malfunction (n = 2). Participants who were excluded during behavioral training did not participate in the fMRI stage, and their data were omitted from all behavioral analyses. Behavioral data were also omitted for participants for whom scheduling conflicts prevented the fMRI session from occurring as planned (n = 3). During the fMRI stage, additional participants were excluded on the basis of either unsatisfactory performance (n = 5, see section 1.1.4.2), visible head movement during the structural series (n =

1), head movement > 4 mm during the functional series (*n* = 1), early termination due to participant discomfort (*n* = 1), or an error in the acquisition protocol (*n* = 1). Both behavioral and fMRI data were discarded for these participants. The final analyzed dataset consisted of 17 participants (10 male, 18-22 *y*, mean +/- s.d. = 19.82 +/- 1.29 *y*). For a single analyzed participant, the second half of the fMRI dataset was omitted due to movement that crossed the 4 mm threshold during these final five runs. Behavioral data acquired during these runs were also eliminated from all analyses.

1.1.2 Behavioral task

The behavioral task was generated using the E-Prime software package (v. 1.1.4.1., Psychology Software Tools, Pittsburgh, PA). During the behavioral training, stimuli were presented via an LCD desktop computer monitor, and responses were collected via keyboard input. During the fMRI scans, stimuli were projected onto a mirror affixed to the scanner head coil, and responses were registered via an MRI-compatible response unit worn on the right hand.

1.1.2.1 Stimuli and general task structure

The basic task display features and response requirements remained constant throughout the pre-scan training and fMRI stages. For each two-trial pattern discrimination problem, a trial included, at minimum, a choice period, an outcome period, and an inter-trial (or inter-problem) fixation period. On each trial, two playing cards marked with unique multi-colored patterns flanked a central fixation cross (see Fig. 2). During the choice period, the left and right cards (position randomized across trials) were selected using either the “1” and “2” keys on the keyboard or the index- and middle-finger buttons of the MRI response unit. During the outcome

period, the selected card was highlighted with a gray box, and the fixation cross was replaced by the number of points earned for the current selection (bold text for rewards, italics for losses). The inter-trial/problem interval was marked by the return of the central fixation cross, which appeared in white between Trials 1 and 2 of a single problem, and in gray before the start of a new problem.

The procedure for generating the individual card patterns was designed to fulfill two goals: (1) to permit the creation of a very large number of unique stimuli, and (2) to present abstract patterns which were unlikely to be familiar to the participants. The latter goal was intended to minimize the inclusion of patterns towards which participants might possess pre-existing biases, and also to discourage the formulation of irrelevant hypotheses based on meanings that the images might carry outside of the task context. Our task stimuli were generated according to a method introduced by Gaffan and Harrison (1988) and subsequently used in several non-human primate studies. Each pattern consisted of two superimposed characters, with the superimposed character appearing as half the size of the base character (52 pt and 104 pt font for the top and bottom characters, respectively). Characters were randomly sampled from a fixed pool of 36 total characters that had been selected from three obscure writing systems (the Shavian, Deseret, and Elder Futhark scripts, see http://www.omniglot.com). On each occasion when a character was sampled, it was additionally assigned a color, which was randomly drawn from a pool of 12 total colors. All random sampling occurred without replacement, such that no individual character (or color) was eligible to re-appear until each of the characters (or colors) had been sampled at least once during a single cycle through the corresponding list.

1.1.2.2 Experimental Factors

The replace-chosen problems presented during fMRI testing and the final phases of behavioral training were manipulated according to three within-subjects factors (see Fig. 2). The first factor, Trial 1 Selection (S+, S-), determined the valence of the initially-chosen stimulus, which in turn determined the replacement which would occur on Trial 2. The second factor, Preview Familiarity (old, novel), determined whether the old, unsampled Trial 1 stimulus or the novel replacement stimulus was previewed before Trial 2. The third factor, Problem Magnitude (high, low), varied the size of the rewards and punishments that could be delivered on each problem (+20/-20 points or +5/-5 points for selection of a positive/negative stimulus).

The experimental design may also be described according to a scheme based on the values of the old- and novel-previewed stimuli; that is, with respect to a 2(Preview Familiarity) *x* 4(Preview Value: +20, +5, -5, -20) design. The two different design descriptions are emphasized for different aspects of the experiment. Because the initial behavioral training phases did not include a preview period, and because training-stage performance was evaluated according to criteria that were jointly applied to the two Trial 1 Selection conditions (see section 1.1.3), Trial 1-centered designs are emphasized throughout the majority of the discussion of the training. A Trial 1-centered design is also relevant for some aspects of the fMRI data, as Trial 1-based performance criteria were also applied to this stage of the experiment, and because the Trial 1 Selection factor figures into the potential interpretation of some of our observed BOLD effects. The preview-centered design is primarily relevant for the fMRI stage, during which our principal experimental question focused on BOLD (and to some extent, behavioral) effects that were related to the values of the previewed stimuli. In some instances we also report analyses organized according to a 3-way preview-centered design in which the statuses of the previewed

cards are defined categorically as opposed to quantitatively (i.e., 2(Preview Familiarity) *x* 2(Preview Valence) *x* 2(Problem Magnitude)).

1.1.3 Behavioral training

Study participation began with a single behavioral training session lasting up to 2.5 hours. Training consisted of five separate phases designed to gradually introduce the deterministic contingencies, the replacement manipulation, and a set of additional features that supported the transition of the task to an fMRI-appropriate design. At the start of training, the experimenter informed participants that they would earn "points" which would not carry any monetary value during the first four phases of the task, but would acquire a small amount of monetary value during the fifth phase (1 point = 1 cent). Participants were also instructed that their goal should be to earn as many points as possible, and that the rate at which they earned points would govern the rate at which they progressed through the task phases. These verbal instructions were followed by a set of instructions presented on the computer screen:

> On each trial of this training, you will see two playing cards, and you will be able to choose one of them. You can choose the left card using the "1" key, and the right card using the "2" key. For the initial components of the training, there will be no time limit on how long you can take to make your choice.
>
> On each trial, when you make a choice, you will earn or lose a certain number of points. After you make your selection, the number of points that you have earned or lost will appear in the center of the screen.
> For the initial phases of training, the points do not correspond to any actual monetary earnings or losses. However, the rate at which you progress through the training is influenced by the rate at which you earn points. So try to perform well!

> The training will consist of five total phases. Some phases may be preceded by a set of special instructions. Please read these instructions carefully! The information that they provide will help you to maximize your point earnings.

The omission of any reference to the relevant structural features of the task (e.g., the deterministic, opposing stimulus-outcome contingencies, or the replacement manipulation) was motivated by two general considerations. First, the sparse instructions created a learning situation which approximated that which would be presented to nonhuman subjects, thereby implying that our training data provide at least an initial basis for cross-species comparison on this task. Second, we anticipated that the requirement for self-discovery of the structure would add a degree of uncertainty that might otherwise be lacking in a deterministic setting. The presence of uncertainty figures importantly in the engagement of several of the regions that comprise the brain's valuation circuitry (e.g., Bhanji, Beer, & Bunge, 2010; Delgado, Miller, Inati, & Phelps, 2005).

1.1.3.1 Phase 1: Discrimination learning set training

Phase 1 introduced the deterministic contingencies and the problem-level magnitude manipulation. The phase consisted of one or more consecutive blocks of 20 two-trial deterministic discrimination problems. Problems were defined according to two factors: Trial 1 Selection (S+, S-) and Problem Magnitude (high, low). Within each block, five problems were assigned to each of the resulting four conditions and presented in a random sequence. During Trial 1 of each problem, participants were free to choose either the left or right stimulus. All responses were self-paced. Immediately after a response, an outcome display appeared for 0.5 s, with the outcome value determined by the current Trial 1 Selection and Problem Magnitude

assignments. In accordance with the deterministic contingencies, the displayed outcome became permanently associated with the chosen stimulus, and the oppositely-valenced outcome became associated with the unchosen stimulus. Following a 1 s ITI, the two cards re-appeared until a Trial 2 selection was registered. The Trial 2 outcome was determined by the current Problem Magnitude condition and the contingencies that were set on Trial 1. Following the 1 s inter-problem interval, a new pair of cards appeared and a new Trial 1 began.

Although participants were told that progress through the phases was determined by the points earned per block, the actual performance criteria were more subtle and focused exclusively on Trial 2 accuracy. Progression to Phase 2 required the joint achievement of ≥ 90% Trial 2 accuracy on the 10 problems that begin with an S+ selection and with an S- selection, respectively. These joint accuracy criteria were motivated by pilot data which revealed a greater errors-to-criterion rate for problems that begin with the selection of an S- as opposed to an S+ (or, phrased differently, for problems which required a "switch" as opposed to a "stay" response on Trial 2, Cox & Fiez, 2010).

1.1.3.2 Phase 2: Replace-chosen training

Phase 2 introduced the Trial 2 replacement manipulation; all other factors and timing parameters were carried over from Phase 1. Participants were informed that they were proceeding to Phase 2 but were not alerted to the specific change in the task. As in Phase 1, the Trial 1 Selection and Problem Magnitude factors were crossed to form four problem conditions, each of which were represented by five problems in each 20-problem block. For the second trial of each problem, the computer replaced the initially-sampled card with a novel card, with the novel card inheriting the value of the replaced card. As was the case for Phase 1, the Phase 2 accuracy criteria required the joint achievement of ≥ 90% Trial 2 accuracy for both Trial 1

Selection conditions. Pilot data suggested that, relative to Phase 1, Phase 2 tends to produce a reversed Trial 1 Selection effect (Cox & Fiez, 2010); specifically, we found a qualitative trend for a greater errors-to-criterion rate for problems that began with an S+ as opposed to an S- selection (or, phrased differently, for problems which required the avoidance of an old-unsampled S- and/or approach of a novel N+).

1.1.3.3 Phase 3: Adaptation to the Trial 2 preview and the fMRI timing parameters

Phase 3 introduced the fMRI timing parameters and the Trial 2 preview periods. Phase 3 also marked the first phase during which participants were exposed to the full 2(Preview Familiarity) *x* 2(Trial 1 Selection) *x* 2(Problem Magnitude) design. To accommodate this design, block size was increased to 24 problems, with each of the eight conditions represented three times during each randomly-sequenced block.

The transition to the fMRI timing parameters involved two general changes in the structure of each trial. First, the duration of the choice period was limited to a fixed 1 s, regardless of when the participant responded. The chosen stimulus was highlighted if a response was registered during the first 0.9 s of this time window (with the 0.1 s buffer included due to technical issues concerning the maintenance of precise timing). If a response was not detected by the 0.9 s deadline, then the participant earned 0 points for the trial, and this point value was displayed with the warning message "Too Slow" during the outcome period. Since no stimulus was selected, the S+/S- designations were randomly assigned to the two Trial 1 stimuli. The second general change in trial structure involved the introduction of the jittered inter-trial intervals that were necessary for our fast event-related fMRI design. Within each block, ITI durations of 1.5, 3, 4.5, and 6 s were jittered according to a rough approximation of an

exponential distribution (Ollinger, Corbetta, & Shulman, 2001), such the shortest intervals occurred with the greatest frequency.

The Trial 2 preview, which occurred immediately prior to the Trial 2 choice period, consisted of a 1.5 s period during which the previewed stimulus was displayed in the center of the screen. Participants were notified of the preview period, the response deadline, and the altered timing parameters with the following instructions:

> Phase 3 is different from the previous phases in three important ways.
>
> First, during Phase 3 you will be required to make a response within slightly less than 1 second (0.9 seconds) after a pair of cards appears. If you do not make a response within this time window, then you will see a "too slow" message and will receive zero points for that trial. Second, on every other trial of Phase 3 you will notice that one card will appear 1.5 seconds prior to the other card. You should NOT RESPOND when only one card is on the screen. If you do make a response during this time, then you will see a "too early" message and will receive zero points for that trial. You may respond after the second card appears. You will have slightly less than 1 second (0.9 seconds) to make your response. Third, you will notice that the delays between individual trials will be much longer during Phase 3.
>
> In summary, you can expect the following changes during Phase 3:
> (1) You must respond within 0.9 seconds after a PAIR of cards appears.
> (2) You must NOT respond when only one card is on the screen.
> (3) You can expect the delays between trials to be long.

If a response was detected during the preview period, then a display consisting of the previewed card and the “Too Early” message replaced the Trial 2 choice display, and the outcome display was replaced by 0.5 s of additional fixation. For the purposes of assessing attainment of the performance criteria, these trials were recorded as inaccurate. Satisfaction of

the Phase 3 accuracy criteria required ≥ 11 correct Trial 2 responses out of the 12 problems representing each Trial 1 Selection condition during a single block (i.e., > 91.6% accuracy).

1.1.3.4 Phase 4: Adaptation to catch trials

Phase 4 introduced the Trial 2 "catch trials", which in this context referred to Trial 2 previews that were not followed by a Trial 2 choice or outcome period. The catch trials were included to support the eventual fMRI analysis. Specifically, by decoupling the relationship between the preview and the events that typically followed it, the catch trials would facilitate the estimation of any preview-specific BOLD responses (Ollinger, Shulman et al., 2001). Relative to the Phase 3 block structure, the Phase 4 blocks included eight added "catch trial" problems, with a single problem representing each of the eight cells of our design.

Participants were alerted to the presence of the catch trials with the following instructions:

> Phase 4 is similar to Phase 3, except for one important difference.
>
> During Phase 4, you will sometimes see a single card appear on the screen, and it will not be joined by the other card 1.5 seconds later. Instead, the card will disappear and you will see the "+" sign only, indicating that the trial has ended. These trials are called "partial trials." Since the second card never appears on these trials, you will not have the opportunity to make a response. Instead, you will need to wait for the next trial to begin. The majority of the trials are NOT partial trials, meaning that on most trials, you will have the opportunity to respond. Your performance will only be assessed for those trials on which you were allowed to make a response.

> In summary, during Phase 4 you can expect the following to occur: On some trials, a card will appear alone and never be joined by the other card. On these partial trials, you should not respond, and you should wait for the next trial to begin.

Since the catch trials did not include the 1 s Trial 2 choice period, any required "Too Early" messages replaced the first 1 s of the fixation period that immediately followed the catch trial. The Trial 2 accuracy criteria for Phase 4 were identical to those used for Phase 3 (> 91.6% accuracy per Trial 1 Selection condition), and were only applied to the standard, complete trials.

1.1.3.5 Phase 5: Monetary earnings

The purpose of Phase 5 was to expose the participant to monetary gains and losses that were similar to those which they would earn in the scanner. Only a single block of Phase 5 was presented; no accuracy criteria were applied. Participants were informed of these changes with the following instructions:

> Phase 5 is similar to Phase 4, with two important differences.
>
> First, the points that you earn or lose during Phase 5 DO indicate the number of cents that you have won or lost. Second, you will only complete one block of Phase 5 during this experimental session, regardless of your performance.
>
> In summary, during Phase 5 you (1) will earn or lose money, instead of points, and (2) complete only one block of trials.

The end of Phase 5 was signaled by a display reporting the cumulative block earnings (range = \$1.70-\$3.20 in the reported dataset). If the participant completed the five phases in < 2 h, then s/he was dismissed for the day and paid an amount equal to the Phase 5 earnings plus an \$18.00 base pay. Participants who did not complete the training within 2 hours were offered the

option of either terminating the training immediately (and canceling the scan session) or of continuing with the training for the remaining 30 min. If the participant chose to terminate the training, as was common amongst individuals who found the task very frustrating, then s/he was thanked for his/her participation and provided with the $18.00 base pay. These participants were recorded as having failed to pass the behavioral training stage. A similar protocol was followed with those participants who did opt to continue training but did not successfully finish within the remaining 30 min.

1.1.4 Functional MRI session

1.1.4.1 Pre-scan practice

Two practice periods were provided immediately prior to the acquisition of functional data. Each period consisted of problems structured identically to those viewed during Phases 4 and 5 of the behavioral training. The first, 32-trial block of problems was presented outside of the scanner, with the participant interacting with a desktop computer similar to that used during training. A second, 16-trial block of problems was presented inside the scanner during the initial series of structural data acquisition. This second practice period was included to acquaint the participant to the novel visual and response elements of the scanner-based task (e.g., projected display, MRI response unit) before the actual collection of functional data began. No monetary rewards were gained or lost during the practice periods.

1.1.4.2 Data acquisition

Participants were scanned at the Neuroscience Imaging Center with a 3T Siemens scanner. The ~80 min scanning session included the acquisition of multiple structural and

fieldmap images; only those data used in the reported analyses will be described here. Prior to functional acquisition, T2-weighted structural images (resolution 0.8 *x* 0.8 *x* 3.2 mm) were collected in 51 contiguous axial slices with a target orientation of 30° relative to the AC-PC plane. The 30° tilt has been employed in several previous studies of the VMPFC (e.g., Hampton, et al., 2006) as a strategy for reducing susceptibility-related signal dropout in the orbitofrontal cortex (Deichmann, Gottfried, Hutton, & Turner, 2003). Due to the head positioning of three of our subjects, it was necessary to undershoot this target orientation by 1-2° in order to avoid a shift from axial to coronal output. Immediately following the functional runs, a high-resolution T1-weighted MPRAGE volume (resolution = 1 mm^3) was also acquired.

Functional data (resolution = 3.2 mm^3) were acquired with a one-shot echo-planar imaging (EPI) pulse sequence (TR = 1500 ms, TE = 25 ms, FOV = 205 mm, flip angle = 70°) across a 29-slice subset of the structural field of view. This subset was adjusted on a per-subject basis such that the 2^{nd}-most inferior slice in the acquisition grid corresponded to the most inferior slice on which the orbitofrontal cortex could be visualized. In other words, the inferior-most slice of the grid was left as a "buffer slice" that protected against experimenter error in precisely identifying the inferior border of the OFC. The resulting coverage generally excluded the relatively superior and/or caudal frontal areas, a large portion of parietal cortex, and smaller portions of other brain regions. The functional data were collected across ten 5 min 7.5 s runs, each of which consisted of a 32-problem block structured identically as the block encountered during Phase 5 of training. The end of the 10^{th} functional run was followed by a screen displaying the cumulative session earnings (range = $23.95-$29.30 in the reported dataset).

As in the behavioral training stage, the performance criteria that were used to determine fMRI data inclusion required the joint attainment of specified Trial 2 accuracy rates for problems

associated with the two possible Trial 1 Selection conditions. However, relative to the training stage, the fMRI criteria were relaxed in three ways. First, the accuracy target was reduced to ≥ 85% correct responses following the Trial 1 S+ and S- selections. Second, accuracy rates were computed as the percent of correct responses out of the total number of "valid" problems (no Trial 1 response omissions or preview-period responses) as opposed to all non-catch-trial problems. Finally, Trial 2 accuracy data were evaluated across as opposed to within runs. These revised criteria were motivated by practical considerations; application of a more rigorous accuracy standard (e.g., ≥ 90% of all non-catch problems) would have resulted in a doubling of the already high rate of performance-related fMRI data exclusion (i.e., 10 versus 5 subjects). Participants' difficulty in maintaining the near-ceiling performance levels within these conditions – in spite of having undergone a very extensive training protocol – might reflect the challenges that are inherent to the prolonged, continuous performance of any cognitively-demanding task. Previous research on "mental fatigue" supports this interpretation by showing that both response accuracy and speed tend to decline as the time spent on a cognitive-control task increases (e.g., Boksem, Meijman, & Lorist, 2006).

1.1.4.3 Post-scan questionnaires

Two post-scan pen-and-paper questionnaires were administered. The first questionnaire tested participants' explicit, "offline" awareness of the values of the preview stimuli in the absence of significant response speed and memory demands. Eight sets of screenshots, one per task condition, depicted a sample series of events leading up to a Trial 2 preview (see Appendix A). Participants were asked to indicate the values of the previewed cards. The second questionnaire included four items requesting that participants rate the extent to which they "cared" about the different possible outcomes (e.g., "How much did you care about earning 20

cents?"). Responses were marked on a 7-point Likert scale (1 = "didn't care at all", 7 = "cared a lot."). This informal survey of reward- and punishment-related motivation was primarily included as a means of evaluating the effectiveness of the Problem Magnitude manipulation.

1.1.5 fMRI preprocessing

Functional data were preprocessed with Automated Image Registration (AIR v. 3.08, Woods, Grafton, Holmes, Cherry, & Mazziotta, 1998) and NeuroImaging Software (NIS v. 3.6, developed at the University of Pittsburgh and Princeton University). Images were motion-corrected to a target volume representing the midpoint of the functional series; structural inplane data were also aligned to this volume. The motion-corrected data were subsequently detrended such that voxel-wise baseline and linear trends were removed on a run-by-run basis.

In preparation for group analysis, functional data were spatially-normalized to a single reference brain chosen from amongst the analyzed participants. For each participant, the normalization protocol began with two structural alignment steps: (1) alignment of the structural inplanes to the MPRAGE volume using FSL FLIRT (v. 5.4.2, Jenkinson & Smith, 2001), and (2) nonlinear warping of the participant's MPRAGE to the reference MPRAGE with AIR align_warp (v. 5.2.5, 5^{th}-order polynomial model, Woods, Grafton, Watson, Sicotte, & Mazziotta, 1998). Following these steps, the functional data were upsampled to the structural-inplane resolution, subjected to the pre-defined alignment and warp, downsampled to the original 3.2 mm^2 resolution, and spatially-smoothed using the AIR gsmooth program (3D Gaussian filter, 8 mm FWHM).

1.1.6 Hemodynamic response estimation

Within each single-subject smoothed dataset, hemodynamic responses to the major task events were modeled using the AFNI 3dDeconvolve program (Ward, 1998). The baseline component of the regression model consisted of the intercept term and the six head-movement time-series estimates (*x*, *y*, *z*, roll, pitch, yaw) determined during the motion correction step. For the signal component of the model, twenty events of interest were identified, with each event aligned to the onset of Trial 1, the Trial 2 preview, or Trial 2 proper (for non-catch trials). The Trial 1 events were defined according to the four possible Trial 1 outcomes (high/low reward/punishment), and the preview and Trial 2 events were organized by the familiarity, valence, and magnitude of the previewed card (old/novel, high or low S+/S-). Each event was modeled with a series of 10 delta (i.e., "stick") functions time-locked to the 10 functional volumes acquired following event onset. Therefore, the model-estimated shape of each event-related hemodynamic response function corresponded to the estimated beta weights (i.e., the amplitude of the stick functions) for the 15 s following event onset.

The goal of the deconvolution analysis was to obtain a set of hemodynamic response estimates that were specific to the valid and accurately-performed problems. Therefore, it was desirable to reduce the impact of BOLD data associated with invalid or incorrectly-performed problems on the ultimate regression solution. Since the low rate and heterogeneous nature of the behavioral errors precluded the inclusion of a full set of error-related regressors in the regression model, we implemented an alternative strategy that aimed to remove error-associated datapoints from the design matrix submitted to 3dDeconvolve. The removal process consisted of two steps. First, in the original, full matrix of task events, we "zeroed out" the markers that would have otherwise indicated the occurrence of the targeted Trial 1, Preview, and Trial 2 events. This step

effectively rendered the 3dDeconvolve program blind to these events, without removing any information related to the passage of time. Second, we used the 3dDeconvolve "censor" flag to instruct the program to disregard, when computing the regression solution, any rows of the matrix that directly temporally coincided with the actual display of the targeted problems on the screen. For example, for a non-catch problem, this censoring step would instruct the program to disregard those rows that corresponded to the actual Trial 1, inter-trial interval, preview, Trial 2, and inter-problem interval displays.

Since this error-management strategy did present some disadvantages (e.g., disregard of datapoints associated with the tail end of responses to preceding "good" problems, and failure to account for any signal variance that the error-associated problems may have contributed to timepoints that occurred after the actual problem displays), we compared this "censored error" approach to a universal error-coding strategy that assigned all problematic events to a single set of 10 error-specific stick functions. The major analysis results remained largely constant across these two deconvolution strategies.

In preparation for statistical analysis, the response estimates generated for each subject, task component, and task condition were summarized according to a value proportional to the area under the estimated response curve, excluding the contribution of the initial and hemodynamically-irrelevant timepoint (Cox, 2010). The area-under-the-curve (AUC) formula was $\beta_{T2} + \beta_{T3} + \beta_{T4} + \beta_{T5} + \beta_{T6} + \beta_{T7} + \beta_{T8} + \beta_{T9} + (0.5*\beta_{T10})$, where each *B* denotes the estimated beta coefficient for the corresponding response timepoint. The AUC values were the primary dependent measure of interest in all reported fMRI analyses.

1.1.7 Statistical analysis of fMRI-stage BOLD and behavioral data

The principal aim of the majority of our fMRI analyses was to test for hemodynamic effects that co-varied with the values of the old-unsampled preview stimuli, or, in other cases, of the novel preview stimuli or the actual delivered outcomes. In practice, this search for value-related activity was carried out by means of linear trend analyses applied to either the voxelwise data (using the "acontr" option of the AFNI 3dANOVA2 program, Ward, 2006) or to ROI-averaged time-series data (using SPSS statistics software, Release 18.0.0, IBM, Armonk, NY). Trend analysis consisted of the subject-by-subject computation of a weighted sum based on the multiplication of a set of condition-specific contrast weights with their respective AUC values. When testing for a linear effect of value, these weights were set at [20, 5, -5, -20] for the preview (or outcome) conditions associated with these monetary quantities. The resulting sums were then submitted to a single-sample *t*-test to assess the reliability of their difference from a test value of 0. In our fMRI data, the value-based linear contrasts were conducted on both a voxelwise and ROI-focused basis; additional tests of quadratic trends (weights = [1 -1 -1 1]) were included to examine the effects of stimulus (or outcome) magnitude. The fMRI-stage behavioral performance data and a subset of the post-scan questionnaire data were also analyzed according to this general approach. In addition to the primary analyses on the value-related trends, in some cases we performed repeated-measures ANOVA testing on the effects of the categorical variables that were present in our design (e.g., Preview Familiarity, Value, Valence, or Magnitude; or Trial 1 Selection status).

All voxelwise analyses were restricted to a masked region that approximated the common area of brain-related, non-background functional signal across the 17 analyzed subjects (Fig. 3). The group-level mask was constructed according to a multistep process. First, in each individual

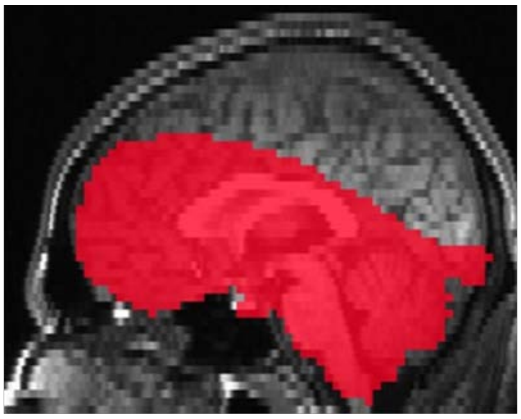


Figure 3. Mid-sagittal view of the masked region included in all voxelwise functional analyses (see section 1.1.7), overlaid on the anatomy of the reference subject.

subject, the AFNI 3dAutomask program was used to generate a mask corresponding to the largest contiguous region of suprathreshold functional intensity. Second, to remove suprathreshold voxels that fell outside of the outer brain contour, we created second-order masks based on the conjunction of the 3dAutomask output and a mask of the subject's brainstripped, warped, and downsampled MPRAGE data. Finally, the group-level mask was defined as the set of nonzero voxels that was common to all 17 subjects' second-order masks.

Within the masked region, the results of the voxelwise trend analyses were thresholded at $p < .001$, and a cluster threshold of 5 contiguous voxels was applied to protect against Type I errors (Forman et al., 1995). For a subset of the significant regions-of-interest (ROI) identified through this approach, the cluster-averaged hemodynamic timecourse data were extracted for the component of the task that contributed to the functional definition of the ROI (e.g., the four old-preview conditions) for submission to follow-up analyses. In cases in which it was of interest to assess the responses of the same ROI to other task components (e.g., the four novel-preview conditions), the ROI voxels identified through the original voxelwise testing process were used as a mask to extract the timecourse data from these other task components. All follow-up hemodynamic timecourse analyses were evaluated against a significance threshold of $p < .05$.

1.2 RESULTS

1.2.1 Pre-scan training

Due to a programming error, two of the 17 participants included in the current dataset were presented with a training sequence that skipped Phases 1-2, such that these individuals only experienced Phases 3-5. This error was not detected until after the analysis process had already begun. Because these participants' latency to complete training and overall performance in the scanner were relatively typical, their data were included in all reported analyses for which data were available (i.e., Phases 3-5 of the behavioral training, and the fMRI analyses).

The phase-by-phase blocks-to-criterion data (Table 1) highlighted a relative peak in the amount of practice required to advance through the phases, with Phase 2 representing the most time-consuming phase. This observation was confirmed by repeated-measures ANOVA, which yielded a significant main effect of phase ($F(3,42) = 10.63$, $p < .00005$), and by significant or marginally-significant comparisons between Phase 2 and the remaining three criterion-dependent phases (p's < .005 for Phase 2 versus Phases 1 and Phase 4, $p = .091$ for Phase 2 versus Phase 3). Note that Phase 2 marks the introduction of the replace-chosen manipulation, and therefore the first point in training during which behavior can no longer rely on experienced stimulus values.

Table 1. Blocks-to-criterion (mean, SE) by pre-scan training phase ($n = 15$).

phase	discrimination (1)	replace-chosen (2)	previews+fMRI timing (3)	catch trials (4)
blocks	1.80 (0.24)	10.20 (1.62)	6.47 (1.18)	3.67 (0.86)

As described in the Methods (Section 1.1.3), in our pilot data we observed that learning difficulty varied across the two Trial 1 Selection conditions, such that the total number of Trial 2 errors to criterion depended on whether the stimulus chosen on Trial 1 had earned a positive or a

negative outcome. The direction of this effect differed across the discrimination and replace-chosen phases of training, such that Trial 1 S- selections were associated with a greater number of Trial 2 errors during the discrimination phase, and Trial 1 S+ selections were associated with a trend for increased Trial 2 errors during the replace-chosen phase. To determine whether these effects were replicated in the present dataset, we specifically focused on the Trial 2 errors-to-criterion data for Phases 1-2 (Table 2). A two-way ANOVA testing the effects of Training Phase and Trial 1 Selection yielded a significant interaction, $F(1,14) = 4.98$, $p = .043$. Consistent with the pilot data, paired *t*-tests performed on the discrimination-phase (Phase 1) error data found a marginally-significant trend for a greater number of errors following Trial 1 S- as opposed to S+ selection ($t(14) = -2.05$, $p = .060$). Likewise, within the replace-chosen phase (Phase 2) error data, the expected trend for a greater number of errors following S+ versus S- selection approached significance ($t(14) = 1.59$, $p = .13$). Therefore, although the phase-specific effects were only modestly supported by the statistics, the training data generally appear to validate the importance of separately evaluating performance for these two problem types.

Table 2. Trial 2 errors-to-criterion (mean, SE) for Phases 1-2 of training, sorted by Trial 1 Selection ($n = 15$).

phase	discrimination (1)		replace-chosen (2)	
trial 1 selection	S+	S-	S+	S-
total errors	1.20 (.43)	2.67 (.78)	36.87 (6.87)	32.73 (6.56)

The many blocks and errors that preceded Phase 2 mastery, coupled with the unconstrained randomization of the card positions, may raise concerns that this phase was passed by means of position-based strategies that were successful when the cards happened to be positioned appropriately. Follow-up analyses (Appendix B) argue against this interpretation of Phase 2 performance.

1.2.2 Functional MRI session behavioral performance

Figures 4(a) and 4(b) present the Trial 2 accuracy and reaction time data for the valid, non-catch trial problems completed during the fMRI session. For the reaction time data, no significant main effects or interactions of preview familiarity, value, valence, or magnitude were detected.

Although, on average, participants maintained generally high accuracy rates, some accuracy differences did emerge amongst the eight preview-organized cells of our design. A two-way repeated-measures ANOVA testing the factors of Preview Familiarity and Preview Value yielded a significant main effect of Familiarity, as reflected in the higher Trial 2 accuracy rates following old as opposed to novel previewed stimuli ($F(1,16) = 10.13$, $p = .006$). This result suggests that participants primarily took advantage of the inferable values of the old, unsampled Trial 1 stimuli, and were less likely to infer (or make use of inferences regarding) the values of the novel replacement stimuli. A significant Familiarity *x* Value interaction was also present ($F(3,48) = 3.20$, $p = .031$); the old- and novel-specific value effects are examined in further detail in the text that follows.

Planned contrasts limited to the old preview data found significant linear (i.e., signed value) and quadratic (i.e., absolute value) effects of previewed stimulus value on subsequent Trial 2 accuracy (linear trend, $t(16) = 2.23$, $p = .040$; quadratic trend, $t(16) = 2.30$, $p = .035$). Both trends were positive, such that increasing stimulus values and high absolute gain/loss magnitudes were both associated with greater Trial 2 accuracy. Further examination of these trends with a two-way Preview Valence (S+, S-) *x* Preview Magnitude (high, low) ANOVA found the expected main effect of magnitude, a main effect of valence that approached significance ($F(1,16) = 3.02$, $p = .10$), and a non-significant interaction effect ($F(1,16) = .087$, p

= .77). Two aspects of these behavioral results are worth noting here. First, although the linear trend may raise concerns that accuracy could serve as an alternative explanation for any value-related effects in our BOLD data, the presence of the quadratic trend suggests that any BOLD activity purely related to accuracy should also be modulated by high versus low stimulus magnitudes. Second, the results of both the trend analysis (combined linear and quadratic effects) and the 2-way ANOVA (significant or near-significant main effects, but no interaction) are consistent with the interpretation that performance largely reflected an additive effect of the two major categorical variables, Preview Valence and Preview Magnitude. An influence of the Preview Valence variable is consistent with the Trial 1 Selection effects reported for the behavioral data, since the old S+ and S- preview stimuli were preceded by Trial 1 Selections of an S- and an S+, respectively (and thereby followed by an optimal policy of choosing either the previewed S+ or the to-be-revealed N+, respectively). An influence of the Preview Magnitude variable is consistent with the assumption that participants appropriately inferred the size of the potential Trial 2 reward that was at stake, which may in turn have created incentive-related effects on performance accuracy (a common phenomenon in rewarded, cognitively-demanding tasks, for examples see Charron & Koechlin, 2010; Gilbert & Fiez, 2004).

In the analyses limited to the novel preview data, an inverse linear relationship with previewed stimulus value approached significance ($t(16) = -1.62$, $p = .12$); no reliable quadratic effect was present. As in the old-preview data, the direction of the linear trend was again consistent with the general observation that Trial 2 was more difficult when the Trial 1 S+ was chosen, thereby requiring the approach of a novel N+ and/or avoidance of a familiar S- on Trial 2. The follow-up valence *x* magnitude ANOVA resulted in a somewhat different pattern than

was seen with the old-preview data, in that only a marginally-significant valence *x* magnitude interaction was found, $F(1,16) = 3.22$, $p = .092$.

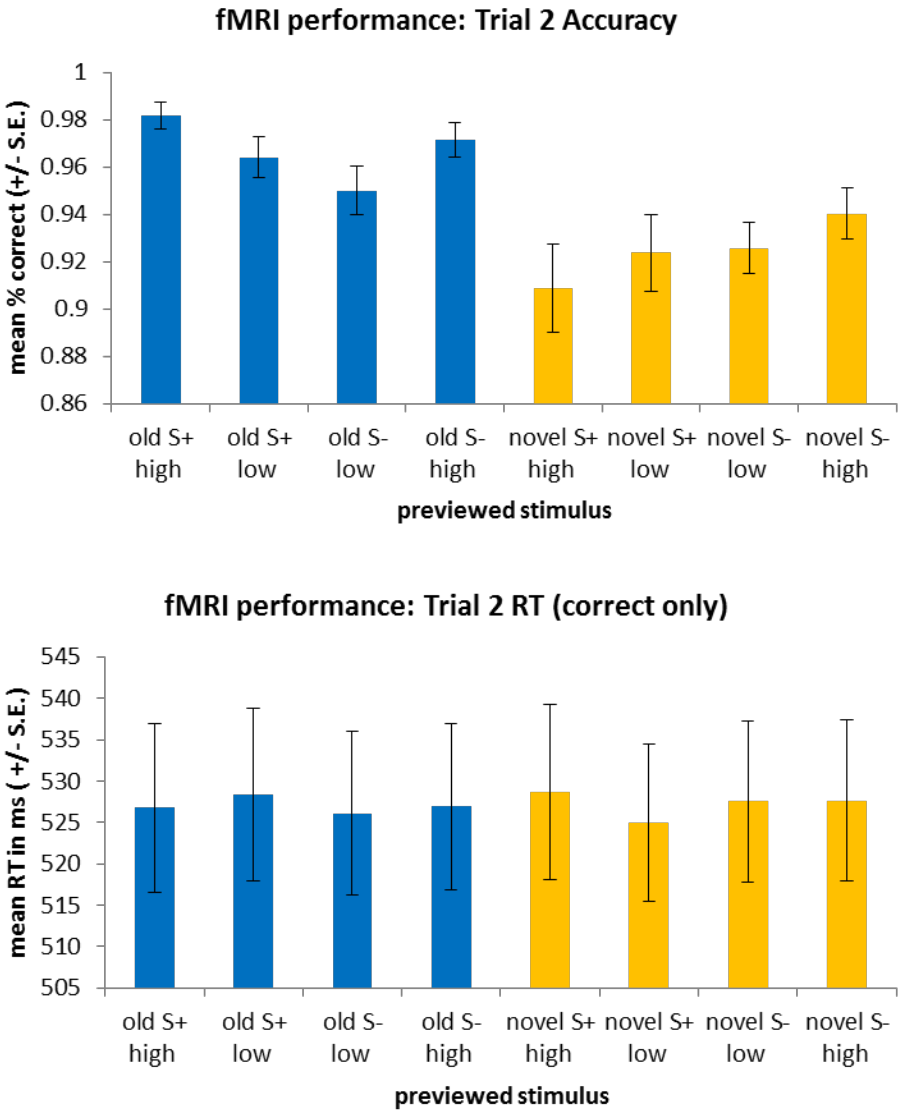


Figure 4. (a) Trial 2 accuracy data for the non-catch trial replace-chosen problems completed within the scanner. Percent correct scores are computed within the set of "valid" problems (no Trial 1 response omissions or preview-period responses). Trial 2 response omissions did contribute to the accuracy scores. (b) Mean Trial 2 reaction time data, computed across valid and accurately-performed problems only.

Appendix C presents the group-summarized data for the total number of invalid response events detected (Trial 1 response omissions and preview-period responses). Some significant

differences between conditions were found (Figure C1(a)); these were largely attributable to patterns that emerged in the preview-period response rates (Figure C1(b)). The most notable effects that were observed included (1) a reliably greater number of preview-period responses during the old versus the novel previews, and (2) a reliable linear relationship between the preview-period response count and the values of the old preview stimuli. This pattern of results is consistent with the assumption that stimulus value is encoded during the old-preview displays, since one might expect that the difficulty of withholding a response should increase in proportion to the predicted value of future approach to the previewed stimulus. The implications of this preview-period effect for the interpretation of our fMRI data are considered in the discussion.

The self-reported degree to which participants "cared about" earning or losing the high or low monetary outcomes is plotted in Figure 5. These outcome-rating data were submitted to a similar sequence of analyses as were applied to the old- and novel-specific Trial 2 accuracy data. Within the self-report data, a positive, U-shaped quadratic effect of value was significant ($F(1,16) = 50.37$, $p = 2.5 \times 10^{-6}$), thus confirming the meaningfulness of our magnitude manipulation within this participant sample. Follow-up *t*-tests confirmed that the greater degree of "caring" about high versus low outcomes remained robust when examined for earnings and losses separately (for earn 20/5, $t(16) = 7.78$, $p = 8.0 \times 10^{-7}$, for lose 20/5, $t(16) = 5.28$, $p = 7.50 \times 10^{-5}$). We additionally observed a linear effect of value that approached significance ($F(1,16) = 2.46$, $p = .137$), and a marginally-significant effect of outcome valence ($F(1,16) = 7.44$, $p = .069$), as reflected in the tendency for the gain outcomes to elicit greater "caring" scores than the loss outcomes. The valence *x* magnitude interaction was non-significant.

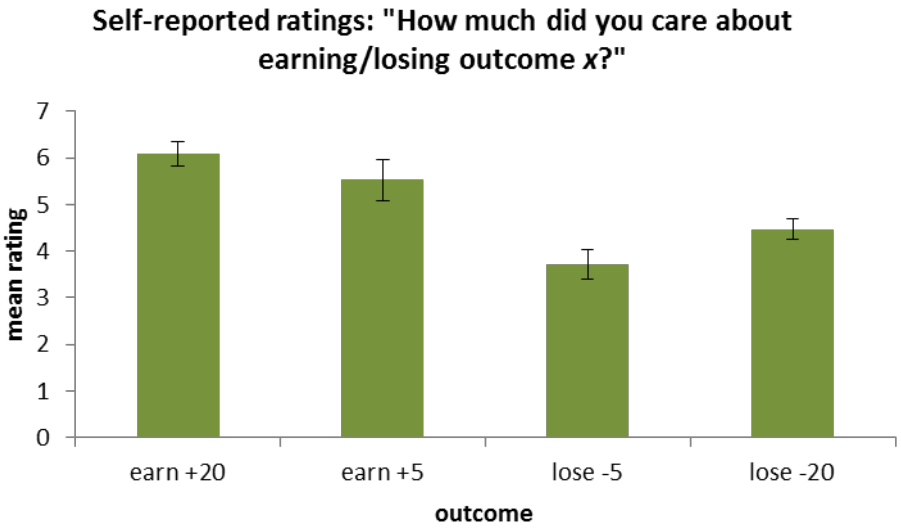


Figure 5. Participants' self-reported attitudes towards each of the possible outcomes, as rated on 7-point Likert scales presented after scan completion.

Participants' responses to the offline sample problems confirmed their ability to infer and explicitly state the values of the old, unsampled preview stimuli. All except one of the 17 participants provided 4/4 correct responses for the questionnaire items that featured old stimulus previews. The single participant who responded incorrectly is likely to have misunderstood the offline task instructions, due to the absence of any sign information (+/-) in his responses, and to his strong performance within the scanner (≥ 90% accuracy for each of the eight possible previews). For the items that featured novel stimulus previews, 8 participants provided 4/4 correct responses, 2 participants provided a single correct response, and the remaining 6 participants did not respond correctly to any of the problems, instead providing responses such as "0" or "don't know." Although these highly disparate response patterns are not of serious concern with respect to our primary experimental question, they do point to the possibility of important individual differences in hemodynamic responses to the novel-preview conditions, particularly amongst those individuals who did correctly predict the novel stimulus values

(hereafter referred to as the "novel-correct" subgroup). This possibility is further explored in the fMRI analyses that focused on the novel preview conditions of the task.

1.2.3 Functional MRI data results

1.2.3.1 Frontopolar BOLD signals vary directly with the values of the old preview stimuli

Our primary experimental question, which concerned the existence of brain regions that signal values deduced via abstract inference, was addressed in the voxelwise test for linear trends that track with the values of the old preview stimuli. At the $p < .001$ threshold, a positive linear effect of old preview value was found in only a single cluster located in the right frontopolar cortex (Figure 6, Table 3). A significant inverse linear relationship with value was found in the lingual gyrus. Notably absent from the voxelwise linear contrast results was the VMPFC. At a relaxed $p < .005$ threshold (cluster threshold = 5 voxels), a second, positive linear effect did emerge in the medial PFC (peak coordinates = [2,63,0]); however, this ROI appeared to be located in the medial frontal pole as opposed to the areas in which VMPFC value activity is typically observed (e.g., the general region identified in the Hampton et al. (2006) study).

Mean hemodynamic timecourse and area-under-the-curve data from the frontopolar ROI are depicted in Figures 7(a) and 7(b). For comparison to data extracted from other task components and other ROIs (see below), trend analysis was repeated on the ROI-averaged AUC data; as expected, the linear effect of value was significant ($t(16) = 5.65$, $p = 3.63 \times 10^{-5}$). The quadratic effect of value was non-significant ($t(16) = .376$, $p = .71$). An alternative analysis organized as a valence *x* magnitude ANOVA found the anticipated significant main effect of valence ($F(1,16) = 12.98$, $p = .004$) and significant valence *x* magnitude interaction ($F(1,16) = 5.99$, $p = .026$).

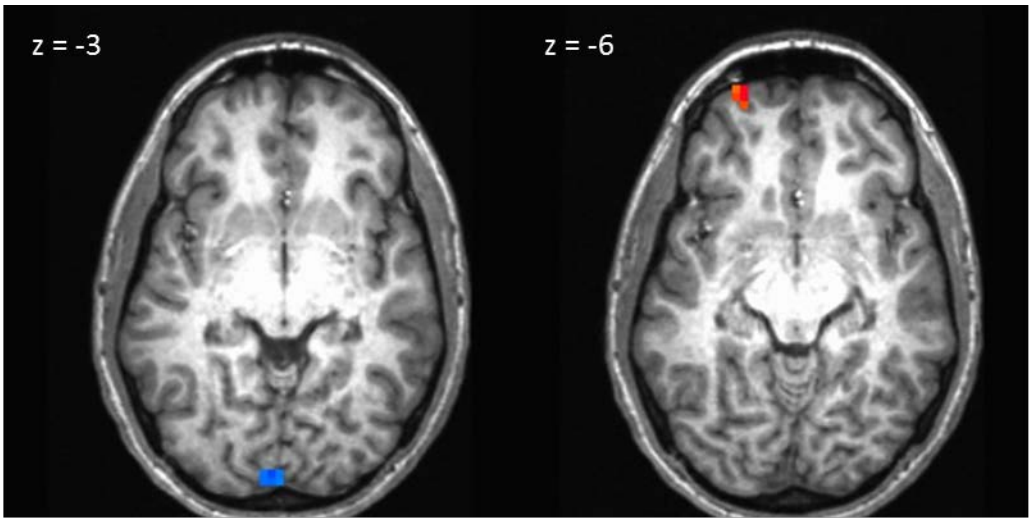


Figure 6. Frontopolar and lingual gyrus regions in which activity exhibited a direct (red) or inverse (blue) linear relationship with the values of the "old" preview stimuli ($p < .001$, contiguity threshold = 5 voxels). All images throughout the text are presented according to radiological convention (right = left).

Table 3. Linear trends related to the values of the old, unsampled preview stimuli, $p < .001$, contiguity threshold = 5 voxels. BA, Brodmann area, TT, Talairach and Tournoux atlas.

ROI Location	Laterality	BA	Peak *t* value	Peak TT coords.	*n* Voxels
Superior frontal gyrus (frontal pole)	R	10	5.25	[22,63,-6]	5
Lingual gyrus	R	17	-4.54	[6,-91,-3]	6

The constant, inverse relationship between the values of the old preview stimuli and of the immediately preceding outcomes (e.g., $S+_{high}$ always follows a 20-cent loss) implies that the observed frontopolar cortex (FPC) value signals may have been attributable to a delayed Trial 1 Outcome effect. According to this interpretation, the FPC does not signal the inferred values of the previewed stimuli, but instead encodes the inverse of the values of the recently-delivered outcomes. If the FPC old-preview response was indeed triggered by the preceding outcomes, as opposed to the preview display *per se*, then this region should exhibit the *inverse* activity pattern in the novel preview conditions, given the direct relationship between the values of the novel preview stimuli and the outcomes that preceded them (e.g., $N+_{high}$ follows a 20-cent reward).

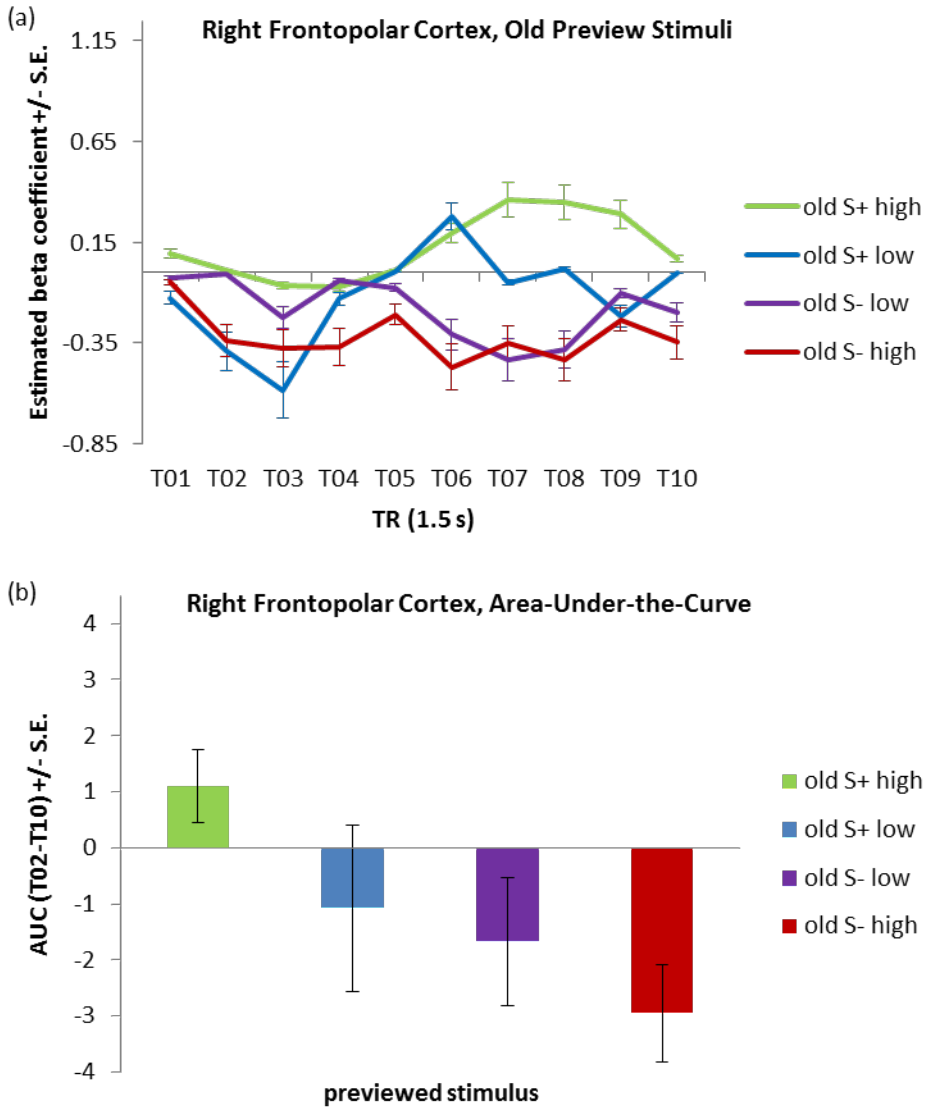


Figure 7. (a) Mean estimated hemodynamic responses extracted from the frontopolar cortex ROI that demonstrated a significant linear effect of old preview stimulus value. (b) Mean area-under-the-curve estimates for the T02-T10 segment of the curves in Figure 7a. Note that the *y* axis scales used for the hemodynamic data vary across figures; however, axes are consistent across the multiple timecourse plots for an individual ROI, and across the multiple AUC plots for an individual ROI.

This possibility was tested in two ways. First, the FPC cluster from the old-preview analysis was applied as a mask for extracting the mean estimated hemodynamic responses to the novel preview stimuli; the resulting timecourses and AUC values are presented in Figures 8(a) and (b). Trend analysis conducted on the AUC data found no significant linear effect of novel preview stimulus value ($t(16) = .47$, $p = .65$; for the novel-correct subgroup, $t(7) = -.24$, $p = .82$).

Additionally, when the novel-preview and old-preview AUC values were compared directly, we found no significant interactions between preview familiarity and any of the other potential effects of interest (i.e, value, valence, magnitude, valence *x* magnitude, all *p*'s > .13).

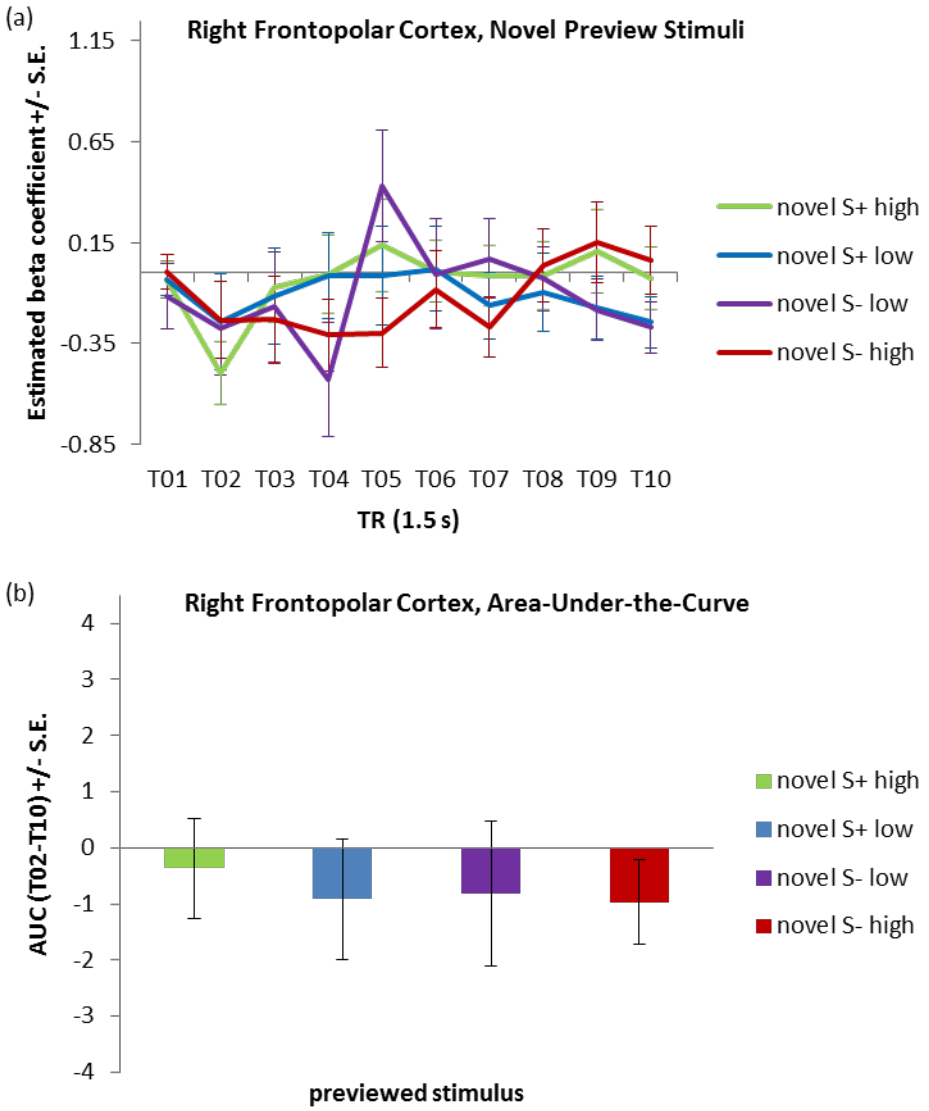


Figure 8. (a) Mean estimated hemodynamic responses to the novel preview stimuli, extracted from the frontopolar cortex ROI defined through the linear contrast of old preview stimulus value. (b) Mean area-under-the-curve estimates for the T02-T10 segment of the curves in Figure 8a.

Consistent with our findings within the pre-defined FPC cluster, voxelwise analysis of the novel-preview data also failed to detect any significant linear effects of value within or near

frontopolar cortex. This null result persisted when the significance threshold was relaxed to $p < .005$ and also when analyses were restricted to the novel-correct subgroup. At the $p < .001$ threshold, positive value-related linear trends were found in two regions located in the right fusiform gyrus and the left middle frontal gyrus. The absence of these regions from the voxelwise old preview analyses, in addition to evidence suggesting that the values of the novel stimuli were not readily inferred (e.g., lower accuracy following novel previews, and for predicting novel stimulus values offline) discourages the interpretation of these regions as playing a value-signaling role. The factors which may be responsible for these novel-preview effects are discussed in Appendix D.

A Trial 1 Outcome-based explanation of our FPC old-preview findings might also predict that significant value-related activity should be seen in this region during the delivery of the outcomes themselves. To test this second prediction, we performed a similar sequence of analyses on the Trial 1 hemodynamic responses as were performed on the novel preview data. Trial 1 timecourses and AUC values from the old-preview-defined FPC cluster are plotted in Figures 9(a) and (b). Trend analysis of the AUC data found no significant evidence for a direct linear relationship with outcome value, but did detect a marginally-significant quadratic effect of Trial 1 outcome, such that the largest gains and losses tended to elicit the greatest outcome-related activity ($t(16) = 1.88$, $p = .079$). This overall pattern of Trial 1 results is consistent with the assumption that the linear FPC effect observed during the old previews is unlikely to be attributable to the immediately preceding outcome displays.

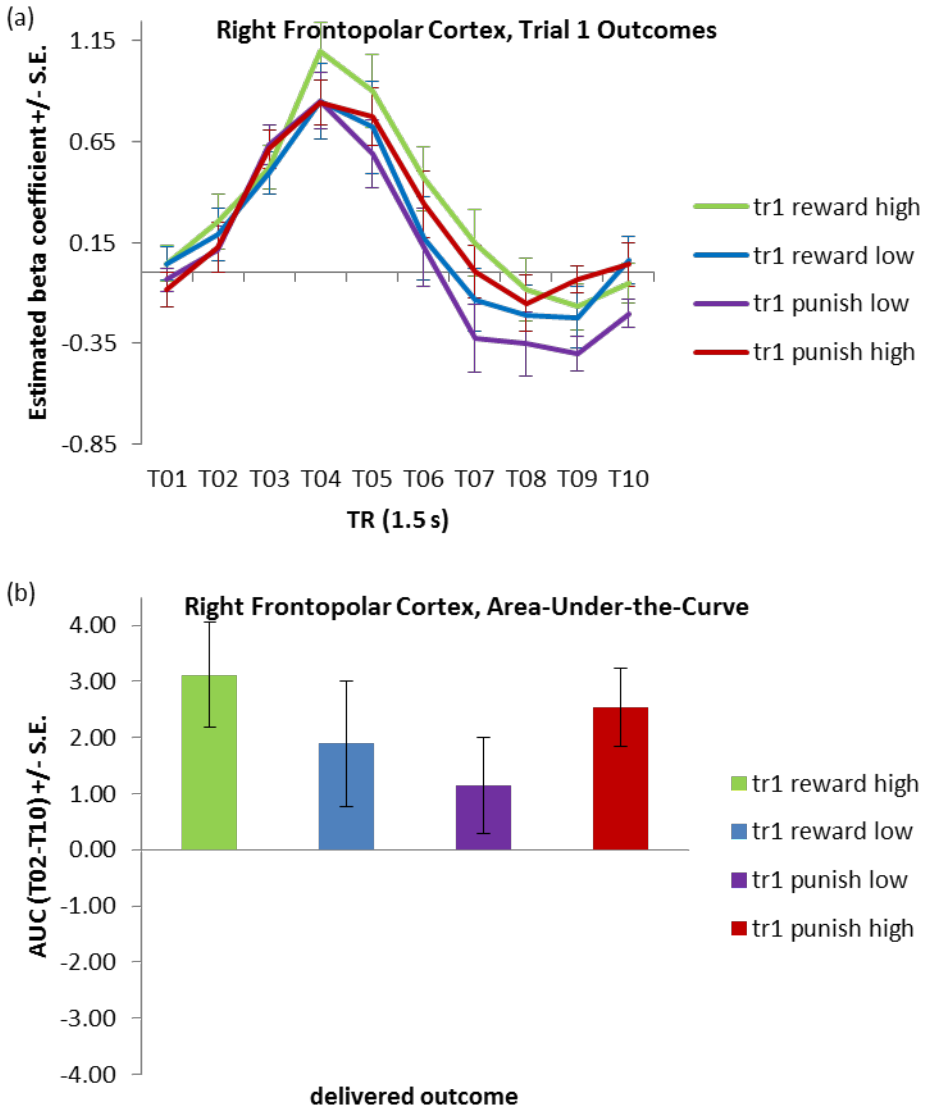


Figure 9. (a) Mean estimated hemodynamic responses to the Trial 1 choice and outcome displays, extracted from the frontopolar cortex ROI defined through the linear contrast of old preview stimulus value. (b) Mean area-under-the-curve estimates for the T02-T10 segment of the curves in Figure 9a.

Follow-up, voxelwise tests of the linear effect of Trial 1 outcome value did not find any significant activation within frontopolar cortex; this null result persisted at the relaxed $p < .005$ threshold. Activation that did track with outcome value was found broadly across the striatum, globus pallidus, and midbrain (Appendix E). This activation pattern is consistent with the extensive existing fMRI and electrophysiological data demonstrating the engagement of these

regions by positive and negative outcome delivery (e.g., Delgado, Locke, Stenger, & Fiez, 2003; Nieuwenhuis et al., 2005; Zaghloul et al., 2009).

1.2.3.2 Engagement of lingual gyrus by the old preview stimuli

As was noted in section 1.2.3.1, the voxelwise analysis did detect a single lingual gyrus cluster in which activity varied inversely with the inferable values of the old preview stimuli. The timecourses and AUC values for this cluster appear in Figure 10.

Trend analyses performed on the ROI-averaged AUC data confirmed the expected, inverse linear effect of value ($t(16) = -4.46$, $p = 3.95 \times 10^{-4}$) and also found a marginally-significant, inverse U-shaped quadratic effect ($t(16) = -2.05$, $p = .057$). The marginal significance of the quadratic trend suggests that the lingual gyrus was responsive to a different feature of our task than the inferred values *per se*. This interpretation is consistent with the results of a secondary valence *x* magnitude analysis, which found a significant main effect of valence ($F(1,16) = 13.24$, $p = .002$), a marginally-significant effect of magnitude ($F(1,16) = 4.19$, $p = .057$), and a non-significant valence *x* magnitude interaction ($p = .22$). Therefore, as was the case for the old-preview Trial 2 accuracy data, the lingual gyrus response may be best described as a categorical effect of viewing either an old-unsampled S+ or S-. As described earlier, participants tended to find the old-preview S- conditions (which evoked greater lingual gyrus activity) more difficult than the old-preview S+ conditions.

To further clarify the nature of the lingual gyrus response, we applied the old-preview-defined ROI as a mask to the novel-preview and Trial 1 data. Lingual gyrus data from the novel-preview conditions appear in Figure 11. Despite the qualitative trend in the AUC means,

such that previews of novel S+ stimuli appear to evoke greater activity than novel S- stimuli, neither trend analyses nor a valence *x* magnitude ANOVA conducted on these data revealed any reliable effects (all *p*'s > .12).

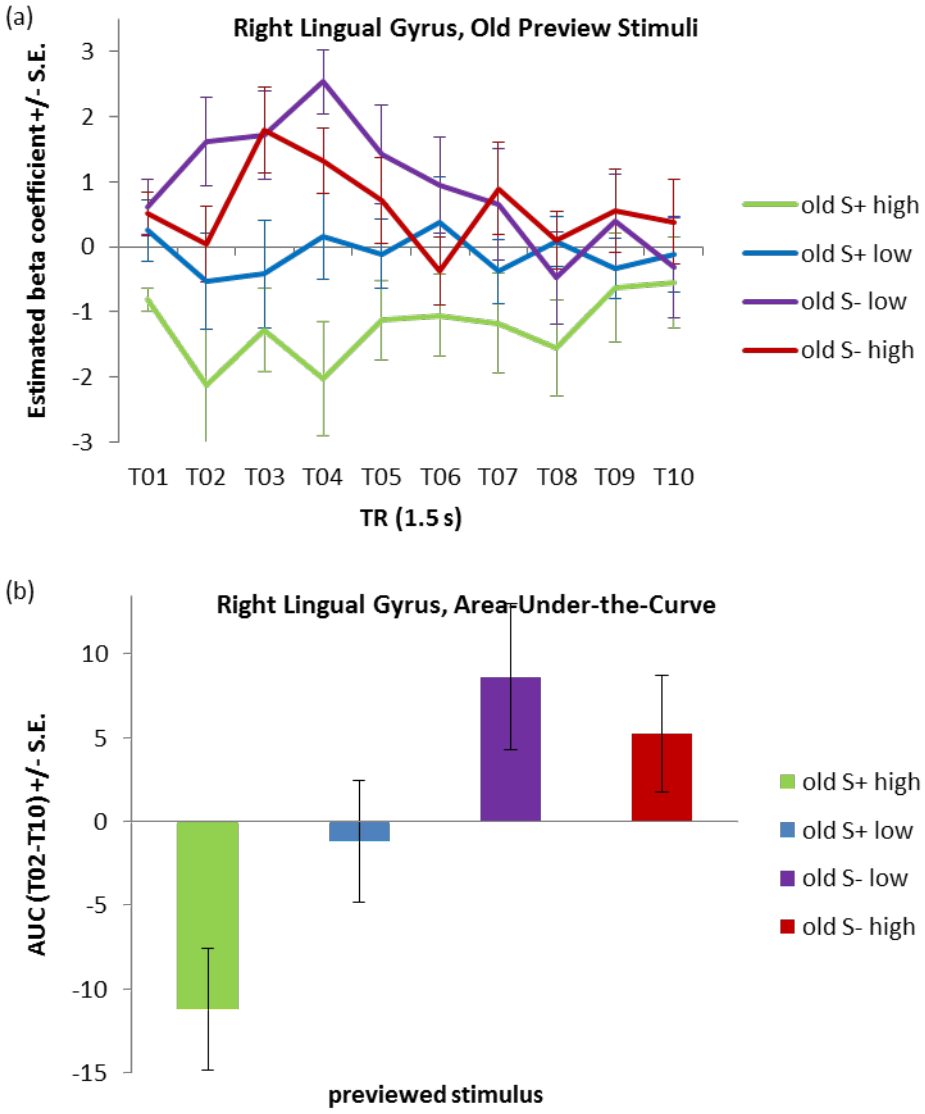


Figure 10. (a) Mean estimated hemodynamic responses extracted from the lingual gyrus ROI that demonstrated a significant inverse linear effect of old preview stimulus value. (b) Mean area-under-the-curve estimates for the T02-T10 segment of the curves in Figure 10a.

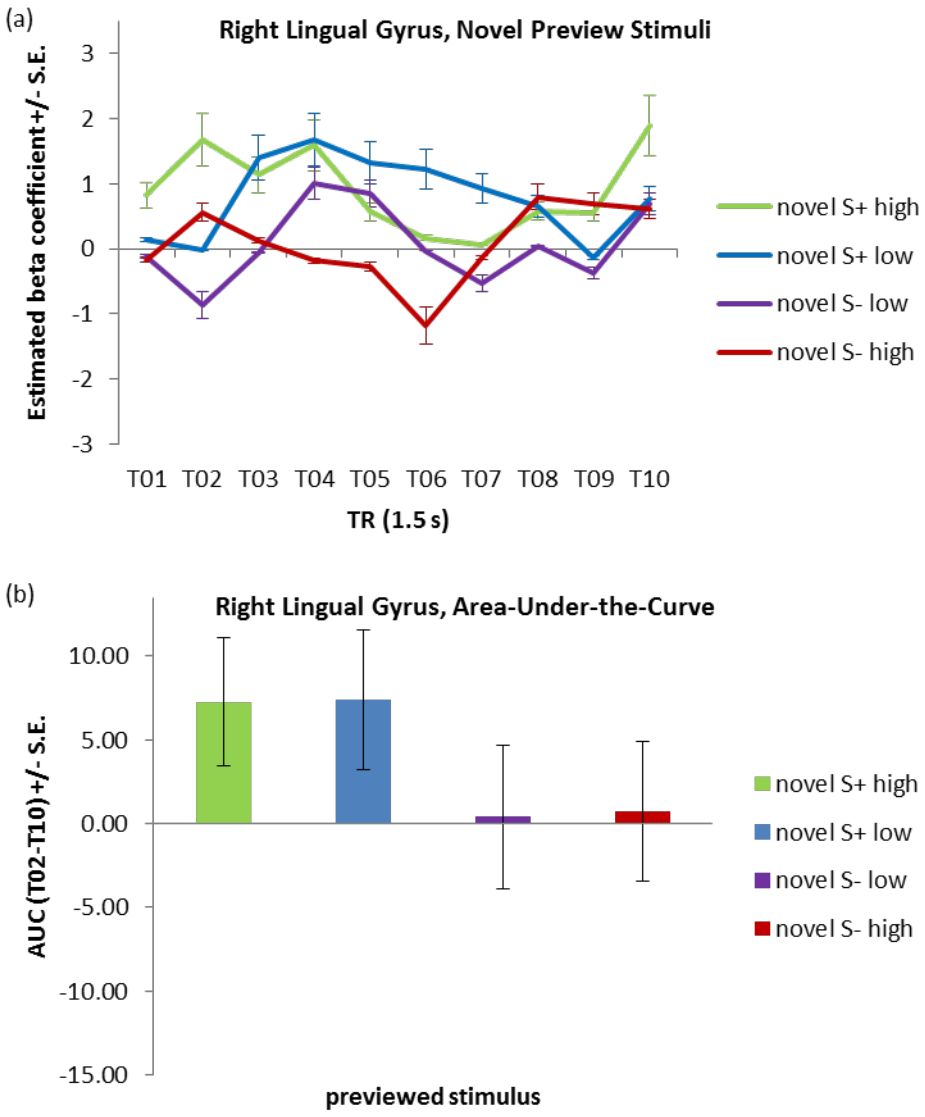


Figure 11. (a) Mean estimated hemodynamic responses to the novel preview stimuli, extracted from the lingual gyrus ROI defined through the linear contrast of old preview stimulus value. (b) Mean area-under-the-curve estimates for the T02-T10 segment of the curves in Figure 11a.

Lingual gyrus response curves and AUC values from the Trial 1 conditions are plotted in Figure 12. During this component of the task, lingual gyrus activity resembled that of the frontopolar ROI, such that responses were primarily sensitive to the greater outcome magnitudes, regardless of outcome valence. Trend analysis found only a significant quadratic trend, $t(16) = 2.79$, $p = .013$. A secondary valence *x* magnitude ANOVA found neither a significant main effect of valence nor a significant valence *x* magnitude interaction (*p*'s > .50).

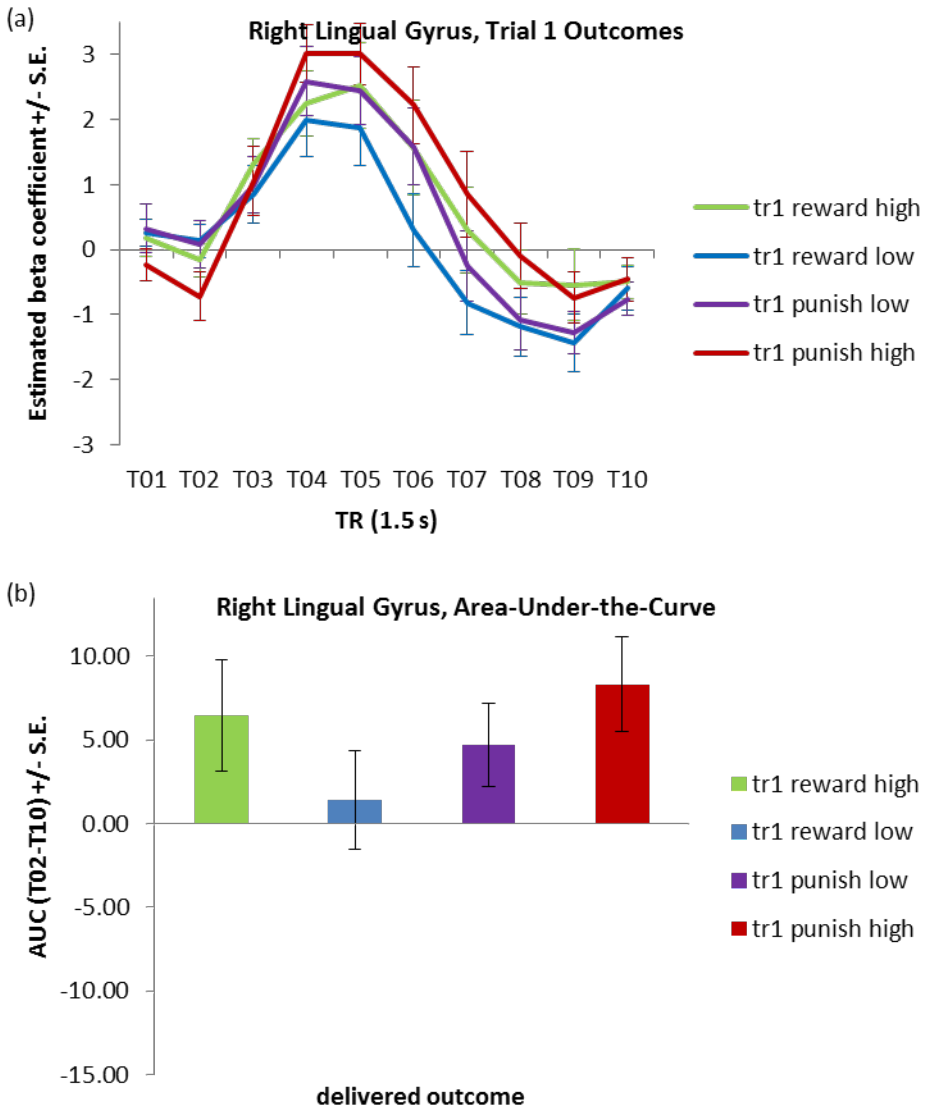


Figure 12. (a) Mean estimated hemodynamic responses to the Trial 1 choice and outcome displays, extracted from the lingual gyrus ROI defined through the linear contrast of old preview stimulus value. (b) Mean area-under-the-curve estimates for the T02-T10 segment of the curves in Figure 12a.

1.2.3.3 Ventromedial prefrontal cortex responses to the old preview stimuli

Given the failure of the voxelwise tests to detect old-preview value effects in the VMPFC, we conducted a series of follow-up analyses in order to determine the nature of the

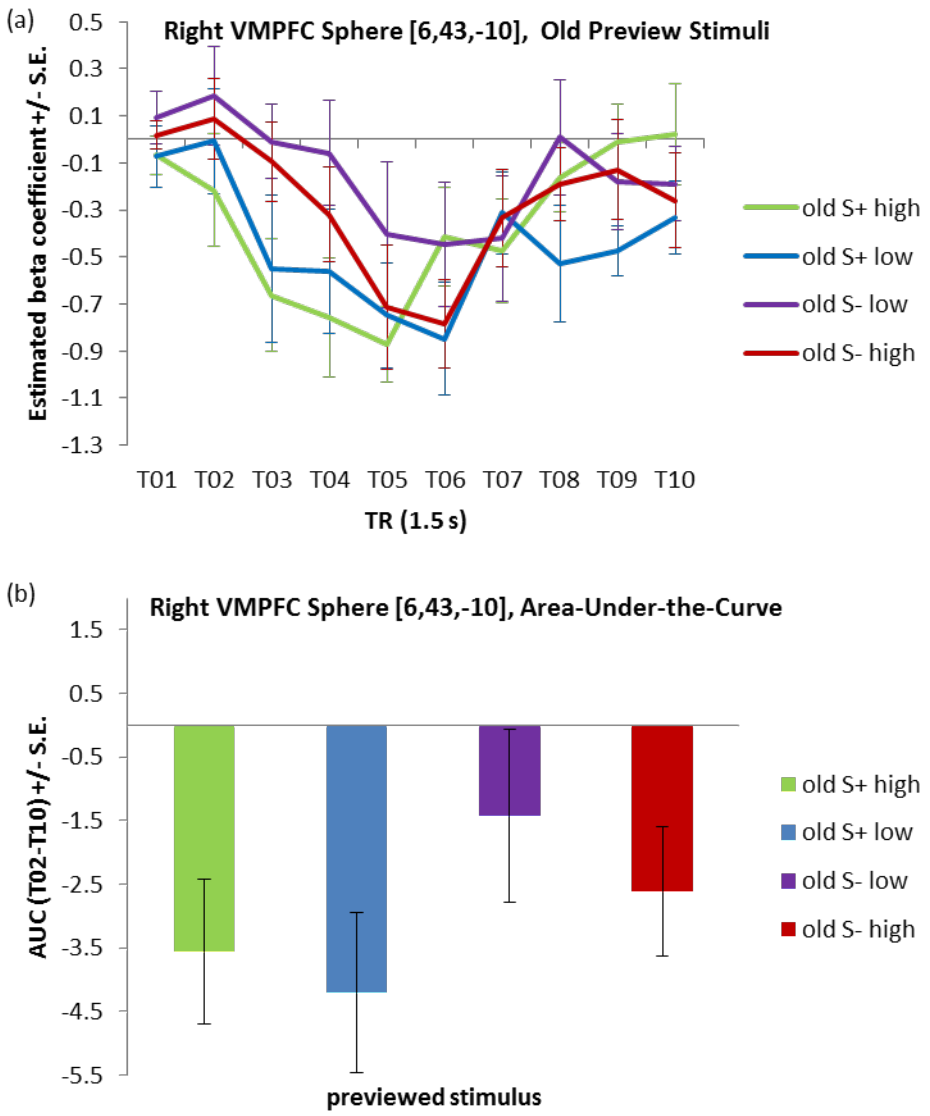


Figure 13. (a) Mean estimated hemodynamic responses to the old preview stimuli, extracted from the a right VMPFC ROI that approximates the region in which activity has been previously best explained by model-based RL (Hampton et al., 2006). (b) Mean area-under-the-curve estimates for the T02-T10 segment of the curves in Figure 13a.

activity pattern that was elicited in this region by the old preview stimuli. As a spatial focus of these analyses, we chose a set of coordinates reported by Hampton et al. (2006) that corresponded to the peak voxel for which the model-based RL algorithm had provided a significantly better account of VMPFC activity relative to model-free RL (MNI coordinates = [6,45,-9]). Old-preview hemodynamic responses and AUC values were extracted from a 7-voxel sphere centered on these coordinates (converted to the Talairach atlas space, [6,43,-10]); these

data are presented in Figure 13. Trend analysis on the VMPFC AUC data found no significant results for either the linear or quadratic value-related trends. A secondary valence *x* magnitude ANOVA yielded only a main effect of valence ($F(1,16) = 6.55$, $p = .021$); VMPFC responses to old-preview S+ stimuli tended to evoke a greater activity decrease relative to the S- stimuli. Therefore, in its differential responsiveness to the previews preceding more difficult (S-) and less difficult (S+) Trial 2 choices, this VMPFC ROI resembles the lingual gyrus ROI, albeit with an overall deactivating as opposed to activating response pattern.

VMPFC timecourse and AUC data from the novel-preview and Trial 1 components of the task are presented in Figures 14 and 15, respectively. Neither trend analyses nor ANOVA tests found any significant value, valence, or magnitude effects within these data.

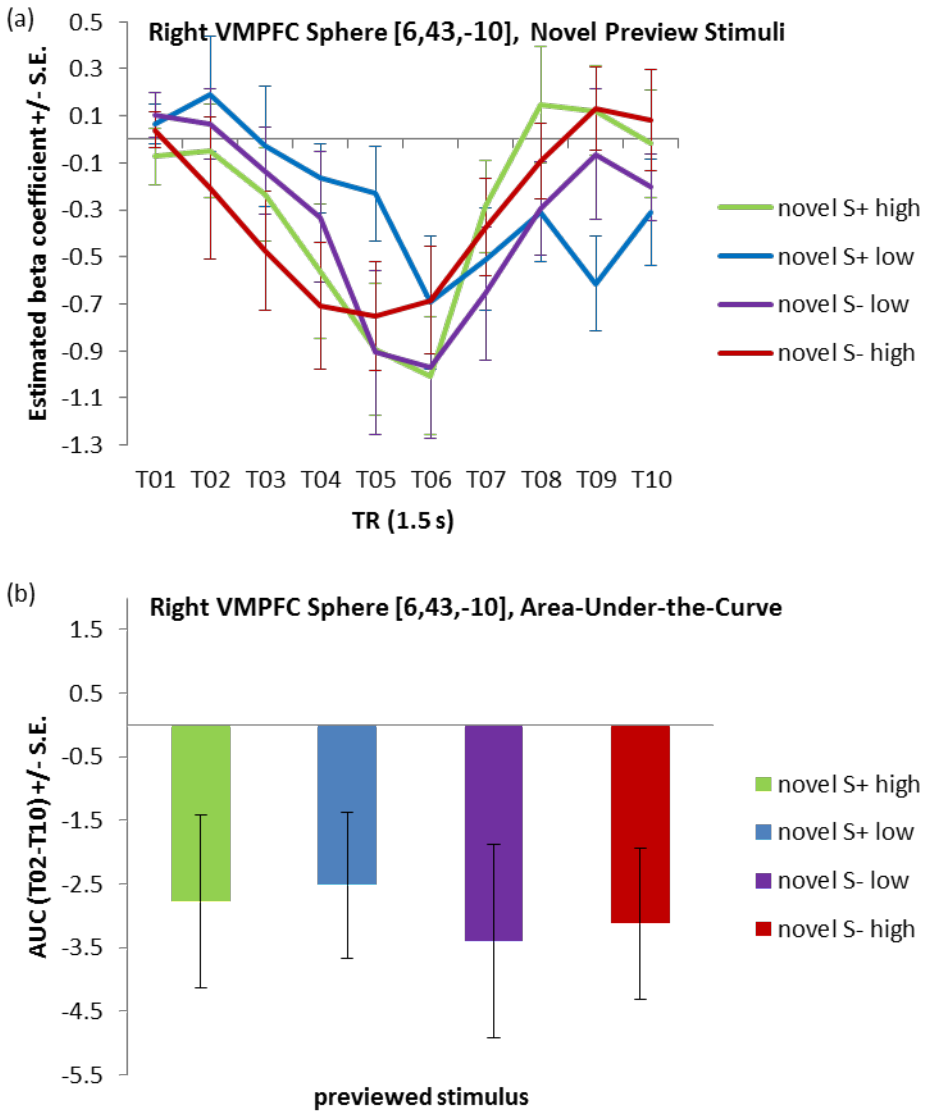


Figure 14. (a) Mean estimated hemodynamic responses to the novel preview stimuli, extracted from the a right VMPFC ROI that approximates the region in which activity has been previously best explained by model-based RL (Hampton et al., 2006). (b) Mean area-under-the-curve estimates for the T02-T10 segment of the curves in Figure 14a.

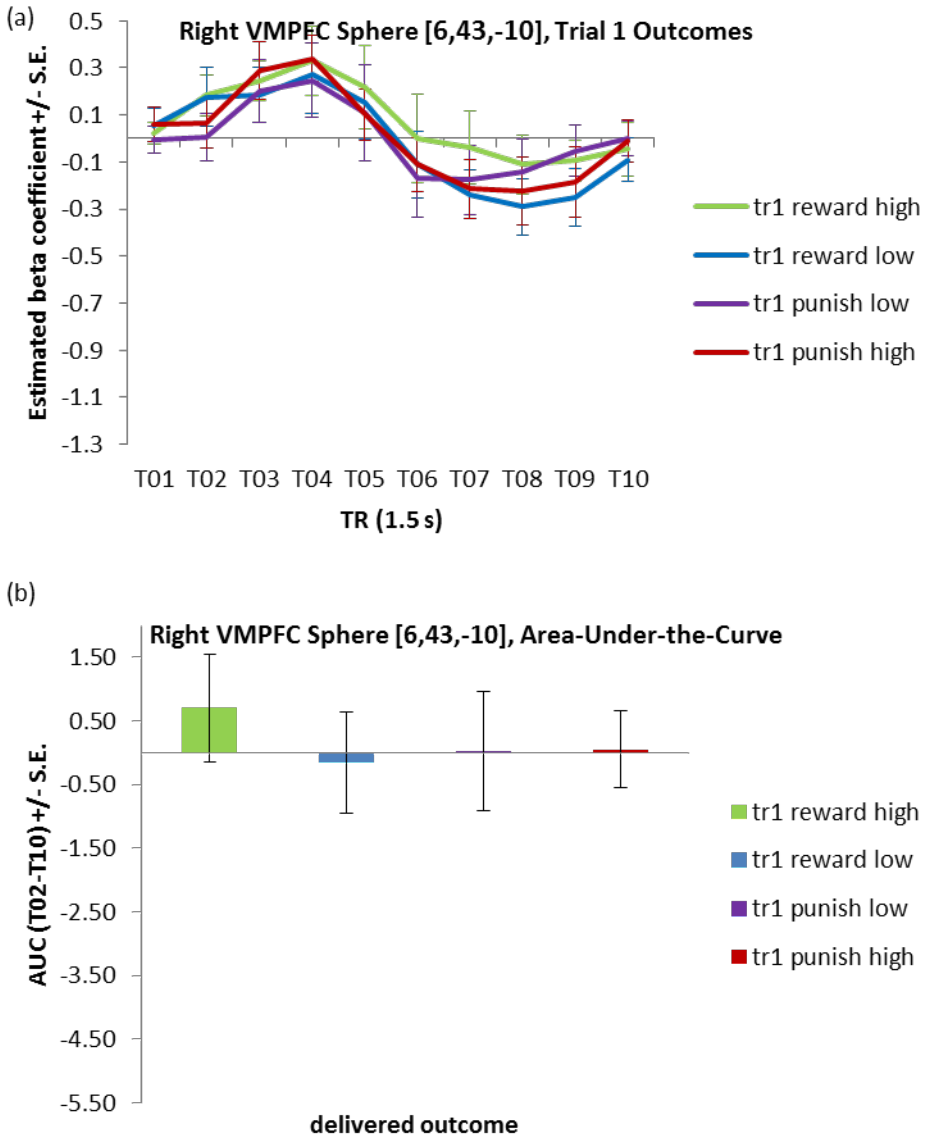


Figure 15. (a) Mean estimated hemodynamic responses to the Trial 1 choice and outcome displays, extracted from a right VMPFC ROI that approximates the region in which activity has been previously best explained by model-based RL (Hampton et al., 2006). (b) Mean area-under-the-curve estimates for the T02-T10 segment of the curves in Figure 15a.

1.3 DISCUSSION

Our functional MRI results confirm that inferred value-related activity can be found in the human brain, even when the underlying inferences must rely exclusively on abstract rules. We specifically focused on a learning context that was based on an adaptation of a task from the nonhuman primate literature (Lockhart et al., 1963). Following each Trial 1 within a series of unique discrimination problems, we replaced the initially-sampled choice stimuli with novel stimuli of the same reward value, thereby creating a situation that encouraged attention to and inferences regarding the initially-unsampled options. BOLD signal analyses focused on the responses evoked by the initially-unsampled stimuli when these were displayed in isolation following Trial 1 outcome delivery. Activity within a region of the right lateral frontal pole was found to vary linearly with the previewed stimulus values. This finding cannot be easily explained as a delayed effect of the recently-delivered outcomes, as similarly-timed displays of the novel replacement stimuli did not evoke any reliable value-related FPC activity.

The finding of an association between FPC activity and the values of the initially-unsampled stimuli is consistent with at least three general interpretations of the function of this region in value-learning settings. One possibility is that the observed patterns reflect the influence of a confounding factor. The presence of positive linear value effects in a subset of our "old-preview" behavioral measures (Trial 2 accuracy and preview-period responses) might motivate such a concern. However, other aspects of our data and the broader frontal pole literature argue against such an interpretation. Explanations of FPC activity in terms of an accuracy-associated factor (e.g., certainty or confidence) are complicated by the additional presence of quadratic value-related trends in the accuracy data (high magnitude > low magnitude) but not in the frontopolar data. Explanations of our value-related FPC activity in

terms of a factor related to the invalid preview-period responses (e.g., increased rates of inappropriate but ultimately suppressed motor preparation in our included trials) cannot be easily addressed on the basis of our data alone, but seem relatively implausible in view of past research on this brain region. Given the typical association of the FPC with complex processes (e.g., relational reasoning, multitasking, and prospective memory, see Ramnani & Owen, 2004 for a review), and the generally preserved motor skills seen in patients with rostral prefrontal damage (Burgess, Dumontheil, & Gilbert, 2007), it would be surprising to find true motor-related FPC activity in the absence of effects in the known sensorimotor regions that fell within our field of view (e.g., sensorimotor striatum).

The remaining two interpretations of our findings both assume that the FPC was indeed responsive to the inferred values, but they make divergent predictions regarding the degree to which this value-signaling function generalizes across all possible abstract rule-bound contexts. On the one hand, one might posit that the FPC is a key and perhaps even critical node for signaling any value inferred via abstract reasoning. This view would lead to the prediction that the FPC should be engaged by and perhaps even necessary for the performance of any value-learning task that involves the transfer of a learned rule to a novel set of stimulus- or action-defined options. Since previous fMRI research of inferred value-signaling employed fixed option sets, it is difficult to locate existing evidence that directly assesses this prediction. However, the monkey literature may provide at least an initial means of addressing this issue, since it has been speculated that monkeys lack a region that is homologous to the human lateral frontal pole (see Tsujimoto, Genovesio, & Wise, 2011). In spite of this anatomical deficit, and in spite of the considerable difficulty that at least one species of monkey has been shown to exhibit with the Lockhart et al. (1963) task, monkeys are generally capable of learning an astounding variety of

abstract performance rules (see Gaffan, 1985; Genovesio & Wise, 2008 for reviews). Therefore, while this observation does not provide a strong refutation of a claim that the FPC is critical for all abstractly-deduced value predictions, it does suggest a reason to be hesitant in promoting such a conclusion.

In view of the above, the third and more plausible interpretation of our results assumes that frontopolar cortex plays an important role in tracking abstractly-deduced values, but only under certain circumstances. Our observations of inferred-value activity most likely represent a special case of a core FPC function that may be defined independently of the degree to which a situation taxes abstract reasoning processes. As briefly noted in the introduction (and reviewed in more detail below), research on the neural signaling of the values of recently foregone option alternatives has provided some insight into how one might begin to describe the FPC contribution within value-learning settings. The relevant findings from the FPC literature are reviewed in further detail below; subsequently, we discuss how our present results might fit into this broader picture of frontopolar function.

1.3.1 Relevant past investigations of FPC function

Research implicating the FPC in value processing has emerged only recently, with three fMRI studies in particular providing much of the evidence relevant to this domain (Boorman, Behrens, & Rushworth, 2011; Boorman, et al., 2009; Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006). The earliest of these studies (Daw et al., 2006) highlighted the potential frontopolar contribution to exploratory decisions, which prioritize the attainment of value-related information over the pursuit of well-established valuable options. In the Daw et al. (2006) task, participants repeatedly chose between four differently-colored “slot machines” which were

rewarded according to independently- and randomly-varying payoff rates (i.e., rates were drawn from normal distributions with means that decayed towards a common value but were also subject to random trial-by-trial fluctuations). Subjective, trial-by-trial predicted values of the slot machines were modeled according to a Bayesian RL algorithm that incorporated a set of assumptions regarding the parameters that governed the variability in the payoffs. Choices were then classified as "exploratory" if they were directed towards any option other than the one associated with the highest predicted value; the remaining choices were labeled as "exploitative." In the BOLD data analysis, contrast of the exploratory versus exploitative trials identified exploration-preferring activity within the frontal pole. The authors offered the preliminary interpretation that the FPC may have been involved in controlling the switch in behavior from exploitative to exploratory modes. As was later noted by Boorman et al. (2009), the Daw et al. results are also consistent with an account in which the FPC principally serves to track the values of unchosen stimuli, if one assumes that participants were most likely to make an exploratory decision when the predicted value of the to-be-explored option was relatively high.

As was briefly summarized in the introduction, the Boorman et al. (2009) study provided direct evidence for unchosen-value activity within the frontal pole. This study is worth reviewing in greater detail, as the behavioral task included several carefully-designed elements that permitted the authors to advance very precise claims regarding the observed FPC signals. Specifically, the authors sought out unchosen-value signals that tracked information relevant for future as opposed to immediate decisions. Within their task, the independently, randomly- and gradually-drifting reward probabilities associated with the left- and right-button press options represented the variables that were most relevant to track as a reference for future decisions. To dissociate the kinds of information that were useful for pending- and current-trial decisions,

Boorman et al. also implemented a "reward magnitude" manipulation: On each trial, randomly-generated numbers appeared on the left- and right-side of the screen, indicating the potential point-total that could be earned for choosing the corresponding button-press. Therefore, on each trial, adaptive decision-making required the integration of the information tracked across past trials (reward probabilities) with that presented on the current trial (reward magnitudes). Participants were explicitly notified of these structural features of the task. Within this context, Boorman et al. found that frontopolar activity tracked with the relative reward probability associated with the unchosen action – that is, the difference between estimated reward probabilities associated with the unchosen and chosen options, with the subjective probability estimates generated according to Bayesian, model-based RL. Activity correlated with the "relative chosen action values" – computed as the difference, between the chosen and unchosen options, of the probability-weighted reward magnitudes – was found within the VMPFC.

Based on their findings, Boorman et al. (2009) proposed that a fundamental function of the FPC is to monitor the evidence in favor switching to a response alternative, and, when this evidence is sufficient, to promote the execution of the switch through interaction with other brain regions. Two additional pieces of evidence support this account. First, results from a follow-up study (Boorman, et al., 2011) confirmed that the FPC signal may be best understood as encoding the value of switching <u>to</u> a particular target, and not solely the value of switching away from the present policy. Specifically, when their basic task protocol was expanded to include three choice options, the FPC signal was most effectively explained as reflecting the reward probability of the highest-valued unchosen option, relative to the reward probabilities of the other-unchosen and chosen options. Second, in the original 2009 study, functional connectivity analyses suggested that high unchosen-value signals in the FPC may promote activity within the mid-intraparietal

sulcus (IPS), a region that has been previously implicated in behavioral switching during rewarded tasks (Glascher, et al., 2009). In the Boorman et al. connectivity analyses, mid-IPS activity was successfully predicted by a regressor encoding the interaction between frontopolar signal intensity and the relative reward probability of the unchosen option; this finding was limited to the periods immediately preceding actual switch execution.

Before considering how the Boorman et al. findings and theory might inform interpretation of the present results, it is important to acknowledge the existence of other views that have been put forth in the literature, and in particular the "domain-general" theories (e.g., Burgess, et al., 2007; Koechlin & Hyafil, 2007; Ramnani & Owen, 2004) which aim to explain the contribution of the FPC across an array of cognitive phenomena. According to one of the most prominent of these accounts (Koechlin & Hyafil, 2007), the FPC plays a fundamental role in "cognitive branching," a process which was summarized as follows:

> "Collectively, the data depicts an anterior prefrontal system in which lateral prefrontal regions select and maintain the task set governing ongoing action, whereas the FPC enables previously selected task sets to be maintained in a pending state for subsequent automatic retrieval and execution upon completion of the ongoing one. This process, which we call "cognitive branching," forms a domain-general core function at the basis of the behaviors and mental activities requiring simultaneous engagement in multiple tasks that are not serially organized into a single, pre-established superordinate plan (p. 595)."

This description of frontopolar function is not incompatible with the Boorman et al. (2009) view, but instead suggests ways in which the Boorman et al. account could be expanded in order to encompass a larger set of learning and decision-making domains. Relative to this goal, two aspects of the Koechlin and Hyafil (2007) view are particularly noteworthy. First, the authors' reference to the maintenance and execution of ongoing and pending task sets suggests

that the FPC might not only signal the value of switching between two, differentially-rewarded, mutually-exclusive decision options, but also between the multiple subtasks that must be executed serially in order to achieve an overarching goal. Second, the authors have used the term "task sets" in a very broad sense, such that it encompasses both overt motor responses (e.g., option selection) and cognitive operations. Notably, this more inclusive conceptualization of that which is "switched" may be similarly applicable to the intraparietal sulcus, as this region has also been argued to mediate both motor and cognitive (e.g., visual attentional) switching (Grefkes & Fink, 2005).

By merging the Koechlin and Hyafil (2007) perspective and that of Boorman et al. (2009), one might propose the very general idea that the FPC serves to signal the value of switching to an alternative operation (cognitive or motor) that must be postponed while a current operation is being executed (either because it is currently associated with greater incentives, or because it holds an earlier position within an operation sequence). The potential validity of this viewpoint can be illustrated within the domain of relational reasoning, which represents an area for which the importance of the FPC has been relatively well-established (see Cho et al., 2010; Christoff et al., 2001; Green, Kraemer, Fugelsang, Gray, & Dunbar, 2010; Waltz et al., 1999). As described by Thompson and Oden (2000), relational reasoning requires, at minimum, the recognition of an abstract relationship between one set of items (e.g., A larger than B) and judgment of whether the same relationship holds for a second item set (C larger than D?). From the perspective of the merged Boorman et al. (2009) and Koechlin & Hyafil (2007) accounts, the positive FPC activations that are observed in fMRI studies of relational reasoning (e.g., see Christoff, et al., 2001) might be understood as signaling the positive value of switching to the

cognitive process of evaluating the second of the two inter-item relationships, pending completion of the evaluation of the first relationship.

1.3.2 Relating existing FPC theories to current results

As stated previously, our observed frontopolar activity might be most plausibly interpreted as reflecting a single instantiation, in an abstract rule-governed task, of an FPC function that transcends the degree of abstraction present in a decision-making context. As summarized above, this core function may involve tracking the value of switching to a cognitive and/or motor operation, pending completion of a current operation. This view can provide a coherent explanation for our fMRI results, including both the Trial 1 and preview period observations. First, consider the situation that is encountered on Trial 1, during which the participant receives a monetary outcome that is anticorrelated with the inferable value of the unsampled Trial 1 stimulus. If frontopolar cortex simply signaled the value of switching to a foregone decision option, then in this case we would have expected to observe an inverse linear relationship between FPC activity and outcome value. Instead, the only effect in our FPC ROI that received any statistical support was a marginally-significant effect of the absolute outcome magnitudes. This pattern of activity is consistent with a view according to which the FPC, on Trial 1, does not signal the value of the alternative decision option, but instead encodes the general advantageousness of the second component of a cognitive plan that includes (1) the initial monitoring of the sampled stimulus-outcome relationship and (2) the subsequent inference of the unsampled stimulus-outcome relationship. The value of switching to the second, inference-generating "task set" should be independent of Trial 1 outcome valence and dependent

on Trial 1 outcome magnitude (due to the problem-level implementation of the magnitude manipulation).

Second, consider the situation that is encountered during the previews of the old-unsampled stimuli. The participant views a card, the value of which can be inferred on the basis of its anticorrelated relationship with the recently-delivered outcome. At first glance, our finding of value-related FPC activity during this period may seem to offer a straightforward confirmation of the unmodified Boorman et al. (2009) account: The FPC is tracking the evidence in favor of switching to a foregone decision option. However, a notable difference between the Boorman et al. study and our task is that our participants evaluated the unsampled stimuli at point at which the sampled stimuli were no longer available, and indeed would never appear again. Therefore, it is difficult to interpret this activity as signaling a value of a response switch relative to "choosing" the initially-sampled stimulus. Instead, an alternative explanation might posit that the FPC is signaling the value of switching responding towards the previewed stimulus, but that this is a switch that is executed relative to the current activity, which consists of waiting for the end of the forced passivity of the preview period. This interpretation of our results does produce a testable prediction; specifically, one would expect to see the value-related FPC activity disappear if the preview period were omitted. Note that confirmation of this prediction would not necessarily raise concerns for the assumption that the FPC is essential for the task, since, as explained above, this region may be playing a very important role in permitting the execution of the inference process on Trial 1.

In summary, our data do appear to support an emerging theoretical view that the frontopolar cortex encodes the benefit to be gained by switching to a pending "task set" following completion of a current task set (Boorman, et al., 2009; Koechlin & Hyafil, 2007, see

also Ramnani & Owen, 2004). This represents a secondary contribution of our results, beyond the primary contribution of confirming that the brain – and in this case, the FPC – is capable of signaling option values that must be inferred via abstract rules.

1.3.3 Ventromedial prefrontal activity and the potential role of choice uncertainty

At the outset of the experiment, we predicted that the ventromedial prefrontal cortex, as opposed to the frontal pole, would be the most likely locus of deduced-value activity. This prediction was motivated by previous demonstrations of value-correlated VMPFC activity, regardless of whether the values were informed by individual subjective preferences (e.g., Chib, et al., 2009), learned probabilistic associations (e.g., Palminteri, et al., 2009), or inference of a post-reversal contingency switch (e.g., Hampton, et al., 2006). However, in our study, we did not find evidence of a reliable linear relationship between VMPFC activity and the values of the old, unsampled Trial 1 stimuli. Signals extracted from a VMPFC region defined through previous research (Hampton et al., 2006) showed only a categorical effect of the old-previewed stimulus' valence, with old S+ stimuli resulting in greater BOLD response decreases relative to old S- stimuli. This region showed no reliable value-, valence-, or magnitude-related effects during previews of the novel replacement stimuli or during Trial 1 outcome delivery.

Interpretation of these results may be informed by existing views regarding the relationship between VMPFC activity and the uncertainty that is present in a task. One recent argument (Bhanji, et al., 2010) states that VMPFC value-prediction signals are most robustly evoked in the presence of "choice uncertainty," as occurs when the actor is unsure of the likelihood that a decision will lead to the desired outcome (see also Wunderlich, et al., 2009). Examples of situations that produce choice uncertainty include the many reinforcement learning

scenarios that have been discussed here (e.g., Boorman, et al., 2009; Daw, et al., 2006; Hampton, et al., 2006; Palminteri, et al., 2009) and also non-learning situations that necessarily involve some degree of "guessing" (Bhanji, et al., 2010).

Based on the assumption that choice uncertainty is an important factor in driving VMPFC value activity, one might explain the present results by appealing to the relative lack of choice uncertainty that was present in our deterministic task, in spite of the measures that were taken to add at least some degree of uncertainty (e.g., omission of explicit instruction regarding the task structure). It is likely that any subjective uncertainty that did persist in our highly-trained subjects would have varied roughly inversely with the Trial 2 accuracy rates that we observed amongst the old-preview conditions, resulting in the observed pattern whereby the VMPFC response exhibits a more modest response decrease during old S- versus S+ previews. In other words, one could interpret the VMPFC deactivations as a reflection of participants' general ability to resolve any choice uncertainty soon after the recognition of an old-unsampled stimulus, and the valence-related effects as a reflection of the residual uncertainty that might have persisted during some of the old S- previews. Note that, from this perspective, it may still be reasonable to predict that the VMPFC may be capable of signaling values that have been generated via abstract rules; however, such an activity pattern might only occur when choice uncertainty is present (e.g., during learning of the rules, or during execution of probabilistic abstract rules).

1.3.4 Lingual gyrus activity and the potential role of color-specific visual imagery

An unexpected effect that emerged in the old preview analyses was an inverse linear relationship between stimulus value and lingual gyrus activity. Follow-up analyses of the ROI-

extracted data suggested that the linear trend may have only captured a component of the meaningful variance in this region. Notably, the additional effects detected through these analyses – including a marginally-significant quadratic trend, and a significant main effect of valence in the absence of a significant valence *x* magnitude interaction – mirrored the effects seen in the Trial 2 accuracy data (although inversely, such that the lingual gyrus trends directly mirrored the trends in the Trial 2 error rates). Similarly, no reliable differentiation was found amongst the lingual gyrus responses to the novel preview conditions, thus mirroring our observation of only approaching-significance or marginal trends in the novel preview accuracy rates. The Trial 1 Outcome data provided additional information regarding the lingual gyrus response profile, with activation during this task component revealing a preference for the greater-magnitude outcomes.

What factors might account for this overall pattern of lingual gyrus effects? One plausible explanation appeals to the important role of the lingual gyrus in color processing, as has been demonstrated via both fMRI and focal-lesion approaches (e.g., Bouvier & Engel, 2006; Cavina-Pratesi, Kentridge, Heywood, & Milner, 2010; Simmons et al., 2007). In addition to permitting the discrimination of presented color stimuli, the lingual gyrus may also contribute to the retrieval of color-associated information in memory. For example, in a recent MEG study that required the identification of celebrities depicted in color photographs (Lindin, Diaz, Capilla, Ortiz, & Maestu, 2010), lingual gyrus was found to be more active when participants indicated a "tip-of-the-tongue" (TOT) state (i.e., the failure to remember a known name), relative to when they successfully produced the name. The authors interpreted the TOT effect in lingual gyrus and other occipital regions as possibly reflecting the increased visual imagery that was evoked during prolonged and unsuccessful retrieval attempts. In light of anecdotal reports from

our participants regarding their use of stimulus color as a retrieval cue, it is conceivable that a similar explanation might apply to our lingual gyrus observations. The old preview stimuli that preceded poorer performance may have induced a state that, like TOT, involves the recognition of a familiar stimulus but an inability to generate an associated label (in this case, its inferred valence, such as "S-"). These TOT-like states may have in turn evoked an increased reliance on visual imagery and therefore a greater lingual gyrus response. A visual imagery interpretation might also explain the enhanced lingual response to the high-magnitude Trial 1 outcomes, particularly if one assumes that high motivation to encode and/or maintain color information results in similarly increased reliance on visual imagery.

1.3.5 Novel contributions, limitations, future directions

In closing, the following paragraphs reflect upon the current study more broadly, and specifically on its relationship to other work on inferred-value signaling, some of the methodological limitations that need to be considered, and the future research directions that our findings motivate.

As stated above, the principal and arguably unique contribution of our study is its demonstration of hemodynamic correlates of values inferred via abstract rules. The necessary reliance on abstract rules marks an important difference between the inferred values in our task and those studied in the Hampton et al. (2006) probabilistic reversal task, since in the latter situation, inferences of a sudden post-reversal value increase can be adequately explained in terms of the relative weighting of previously-experienced contexts that separately encode the two experienced sets of stimulus-outcome contingencies. However, it should be noted that our study is not the first to investigate the hemodynamic correlates of option-outcome contingencies that

have never been previously experienced (Daw, Gershman, Seymour, Dayan, & Dolan, 2011). Therefore, it is important to delineate the differences between our investigation and this previous work.

Daw et al. (2011) recently provided evidence of BOLD signals that were sensitive to values that could be inferred on the basis of regularities in a task's sequential structure. Specifically, the authors sought to examine model-free versus model-based RL predictions of activity within the ventral striatum, a region in which activity is often found to track reward prediction errors (see Daw & Doya, 2006 for a review). Since value predictions necessarily inform the magnitude of a prediction error, this form of reward-related activity can provide indirect information regarding the predicted values that might be encoded elsewhere within the brain. Each trial of the Daw et al. (2011) task consisted of two states, the first of which involved the presentation of two choice stimuli. Choice of either stimulus determined which of two possible successor states was most likely (i.e., with 70% probability) to follow the initial state. Each successor state consisted of its own set of unique choice stimuli, the selections of which were rewarded according to randomly, gradually-varying probabilistic contingencies. The critical distinction between their model-free and model-based accounts of learning within this task hinged upon the degree to which the two algorithms' value predictions were informed by the state transition probabilities. When rewards were gained for choices made following low-probability state transitions, only model-based RL was able to appropriately potentiate the future selection of the *alternative* first-choice option, as opposed to the first-choice option that had been selected immediately preceding the unlikely transition into the rewarded successor state.

Daw et al. (2011) found that activity within the ventral striatum correlated with a regressor that incorporated a combination of the prediction errors generated by the model-free

and model-based RL algorithms (where prediction errors were modeled at both the points of state transition and of outcome delivery). In other words, this region showed a prediction error-related effect that was likely to have reflected the joint operation of both forms of learning during the task. Importantly, the model-based RL effect in this study cannot be explained in terms of the simple recall of a best matching "context" from previous experience. Therefore, it is important to acknowledge that our study is not unique in demonstrating inferred-value activity in the presence of stimulus-outcome contingencies that have never been sampled directly. However, our study results do remain unique in their resistance to explanation by any model that relies upon a fixed set of states defined in terms of their specific sensory features. Although other fMRI investigators have studied the learning and implementation of decision rules that must be transferred across specific stimuli (e.g., Bunge, Kahn, Wallis, Miller, & Wagner, 2003; Kumaran, Summerfield, Hassabis, & Maguire, 2009), these authors were primarily interested in the structures supporting performance more generally, as opposed to identifying regions that signaled the values of particular options.

Some limitations of our study should be recognized and considered in the design of future research. First, due to the relatively restricted sample size and limited fMRI field of view, our BOLD analyses may have been limited in their overall sensitivity in detecting significant value-related effects. Therefore, it is quite possible that future investigations utilizing a greater number of subjects, trials, or functional slices might reveal inferred value-related activity in regions that were not reported here. Second, it would have been desirable to have avoided the observed accuracy differences amongst our experimental conditions, since these effects do add a degree of complexity to the interpretation of the BOLD data. Any future investigations that take advantage of the Lockhart et al. (1963) protocol may benefit from more stringent criteria that eliminate the

persistence of subjects' difficulty with problems that require that the avoidance of an inferred negative stimulus (or, interpreted differently, the approach of a novel positive stimulus).

A more fundamental challenge for future research concerns the identification of which brain circuits, if any, encode value information that is inferred via abstract rules that fundamentally differ from the anticorrelation-based rule utilized here. The process of answering this question need not consist of a mere descriptive exercise of linking different arbitrary rules to different brain regions. Instead, it should be pursued as a means of understanding how the core processing requirements imposed by different decision rules (for example, the nesting of multiple conditional rules, Badre, Kayser, & D'Esposito, 2010, or the maintenance of "temporally-complex" representations, Browning & Gaffan, 2008) might ultimately result in the recruitment of different networks for representing the outcomes that are predicted to be attained through rule execution. Improved understanding of this theoretical issue may in turn point to important clinical implications, such that patients who have sustained damage to particular regions might be expected to exhibit motivational or cognitive control deficits within some rule-governed settings, but not others.

APPENDIX A

POST-SCAN QUESTIONNAIRE #1

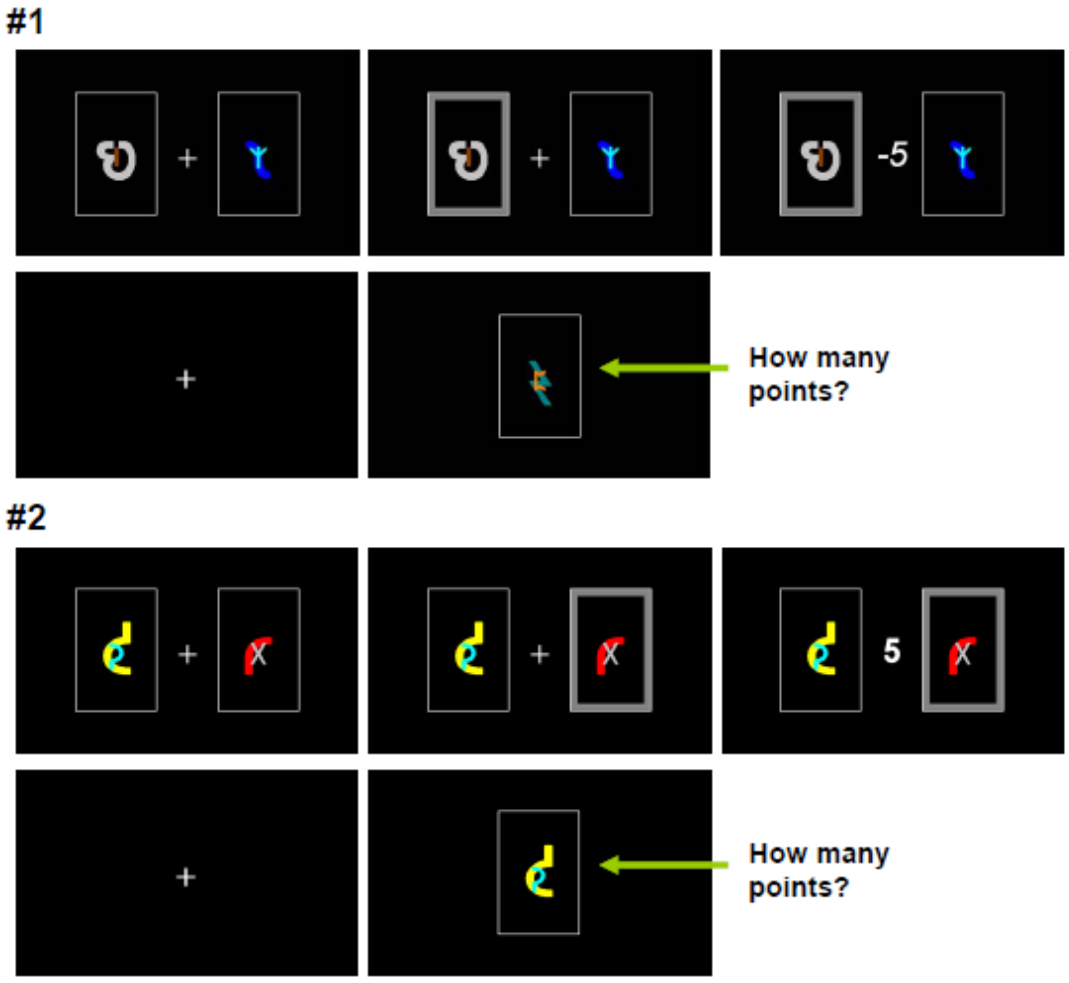


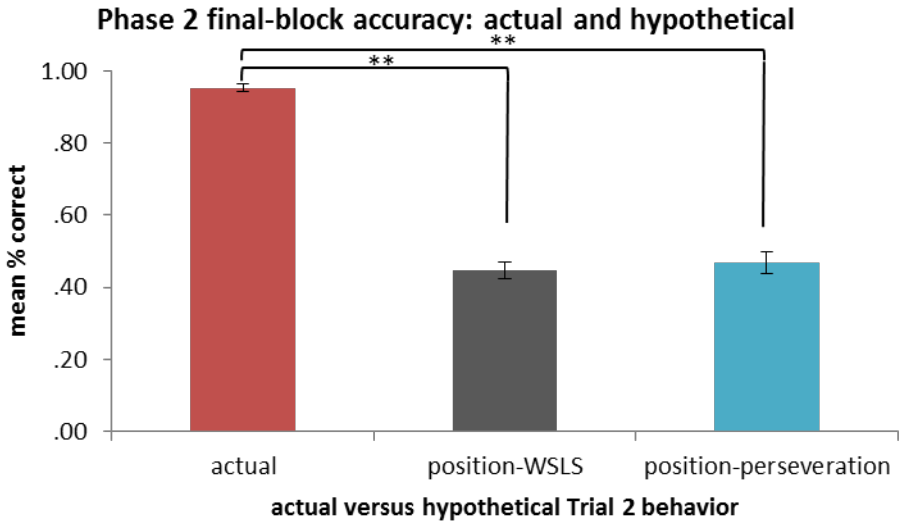
Figure A1. Two sample items from the first post-scan questionnaire. Problems #1 and #2 illustrate previews of a novel, low-magnitude negative stimulus and an old, low-magnitude negative stimulus, respectively.

APPENDIX B

STIMULUS VERSUS POSITION STRATEGIES DURING TRAINING PHASE 2

The randomized nature of the card positions may raise concerns that, by chance occurrence, participants may have satisfied the Phase 2 performance criteria by applying a position- as opposed to a stimulus-based strategy. Two of the position-based strategies that may have been adopted, and which may have been occasionally rewarded, are the position-guided win-stay lose-shift strategy (position-WSLS, c.f. Levine, 1959) and the position-perseveration strategy. Position-WSLS dictates that a left or right keypress should be repeated following a Trial 1 reward, and avoided following a punishment. This strategy would be reinforced at a high rate during blocks for which the relative card positions remain constant across Trials 1-2. A position-perseveration strategy entails the repeated selection of the same keypress across the two problem trials, and would earn the greatest reinforcement during blocks in which position changes more frequently following negative as opposed to positive Trial 1 outcomes. Since Phase 2 was associated with a high blocks-to-criterion count (Table 1), it is conceivable that participants occasionally encountered blocks that were structured in one of these two ways. This occasional position-strategy-favoring structure would be problematic if it allowed for the completion of Phase 2 in the absence of successful learning of a stimulus-based strategy.

To address this concern, we compared participants' actual Phase-2 final-block accuracy with the accuracy rates that would have been observed if the above-described position strategies had been invariably executed. Drawing upon data from each participant's criterion-satisfying block, we used the observed Trial 1 choices and outcomes, combined with the actual Trial 1 and Trial 2 card positions, to determine the expected Trial 2 performance under both the position-WSLS and position-perseveration strategies. The position strategies could plausibly explain participants' satisfaction of the accuracy criteria if they would have indeed predicted highly-successful responding during the final Phase 2 blocks. This did not appear to be the case in our dataset (Fig. B1): The predicted position-strategy accuracy rates approximated ~50%, a significantly lower value than the necessarily high accuracy that was actually observed during the criterion-satisfying blocks (for actual versus position-WSLS, $t(14) = 18.69$, $p = 2.69 \times 10^{-11}$, for actual versus position-perseveration, $t(14) = 15.30$, $p = 3.90 \times 10^{-10}$).



.**Figure B1.** Trial 2 accuracy for the final, criterion-satisfying block of Phase 2, as was observed and as was predicted by two position-based strategies. **, $p < .01$.

APPENDIX C

INVALID RESPONSE TOTALS

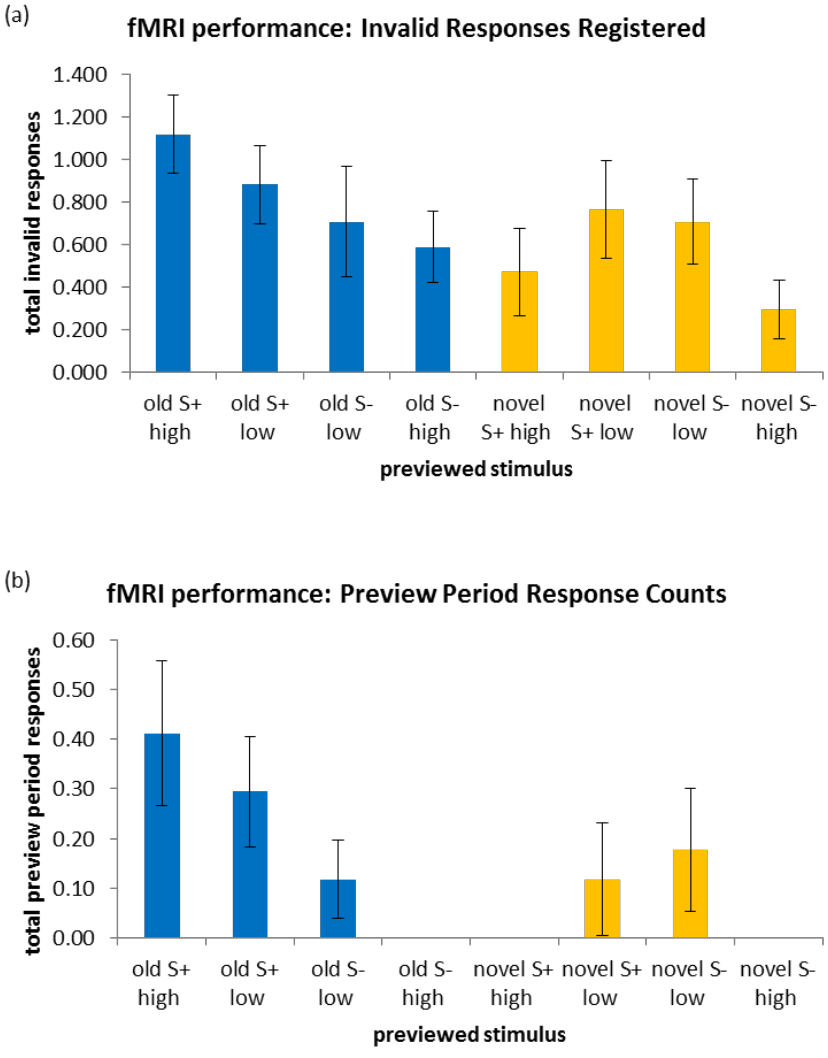


Figure C1. (a) Mean count of invalid responses (Trial 1 response omission or preview-period response) registered during problems associated with the different preview displays. Note that > 1 invalid response can occur during a single problem. The main effect of Familiarity was significant, $F(1,16) = 7.10$, $p = .017$. Across the old-preview stimuli, the positive linear relationship with value was significant, $F(1,16) = 4.84$, $p = .043$. Across the novel-preview stimuli, the negative quadratic relationship with value was significant, $F(1,16) = 4.59$. (b) Mean count of preview-period responses registered in the presence of the different preview displays. The main effect of Familiarity was significant, $F(1,16) = 5.35$, $p = .034$. Across the old-preview stimuli, the positive linear relationship with value was significant, $F(1,16) = 9.67$, $p = .007$. Across the novel-preview stimuli, the negative quadratic relationship with value was marginally-significant, $F(1,16) = 3.13$, $p = .096$

APPENDIX D

VALUE-RELATED ACTIVITY DURING THE NOVEL STIMULUS PREVIEWS

As was briefly noted in the Results (Section 1.2.3.1), the voxelwise search for linear value-related trends in the novel preview data identified two significant clusters at the $p < .001$ level (contiguity threshold = 5 voxels, see Fig. D1 and Table D1). Peak Talairach coordinates suggested that the first of these two clusters primarily mapped onto the right culmen; however, the overlay of adjoining significant voxels onto the right fusiform gyrus – and the consistency of the associated ROI-averaged timecourses with effects reported above for the lingual gyrus – suggested that this cluster might be most appropriately interpreted as mapping onto the fusiform. The second of the two clusters was found in the left middle frontal gyrus.

As was mentioned in the main text, the absence of these regions from the voxelwise old preview analyses, in addition to the evidence suggesting that the values of the novel preview stimuli were not readily inferred, discourages the interpretation of these regions as playing a value-signaling role. To inform speculation of what function these regions might be serving, the two ROIs were submitted to a similar sequence of analyses as was applied to the old-preview ROIs. That is, for each cluster, we extracted the ROI-averaged timecourses for the novel-preview, old-preview, and Trial 1 task components, and examined the resulting data with trend

analysis (linear and quadratic relationships with value) and with 2-way ANOVAs (valence *x* magnitude). These results are presented in Tables D2-D3 and Figures D2-D7; possible interpretations are discussed in the section that follows.

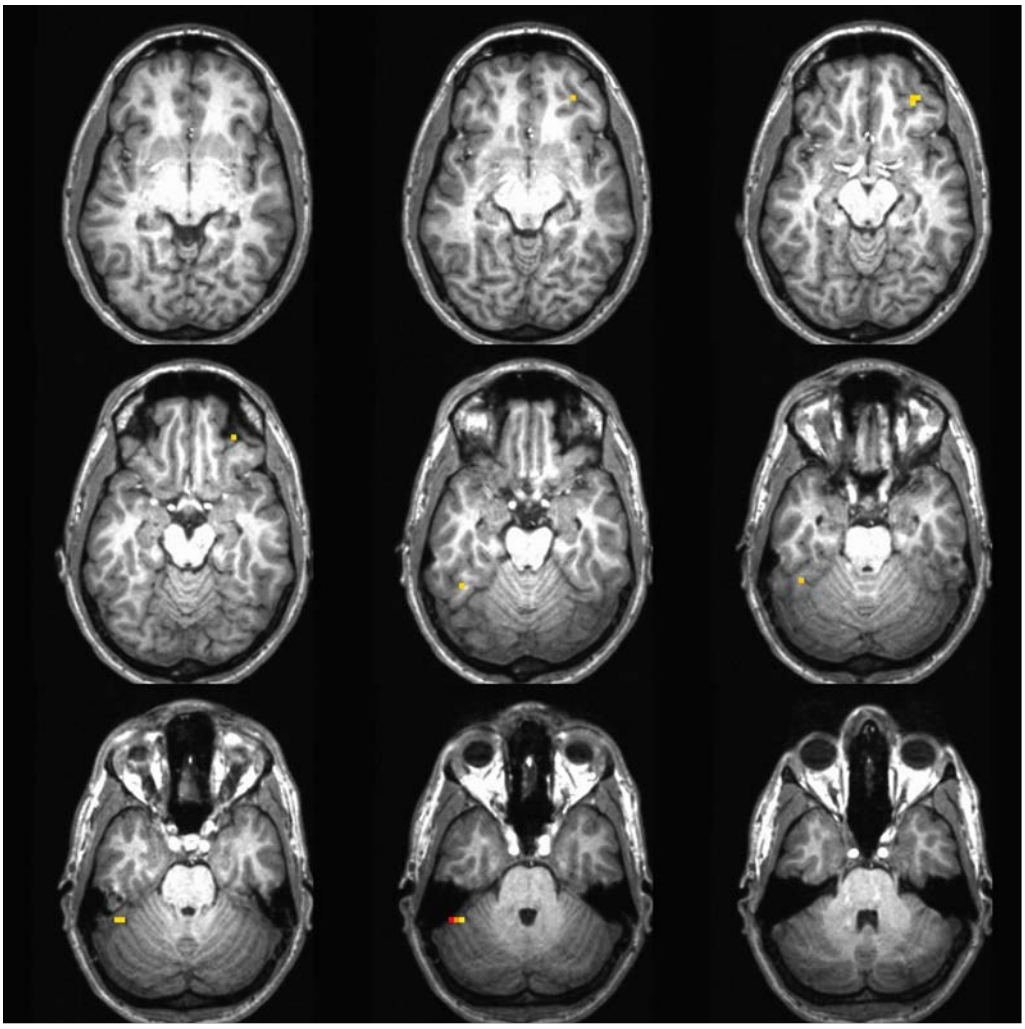


Figure D1. Regions in which activity exhibited a direct linear relationship with the values of the "novel" preview stimuli ($p < .001$, contiguity threshold = 5 voxels). Images are presented according to radiological convention (right = left).

Table D1. Linear trends related to the values of the novel, unsampled preview stimuli, p < .001, contiguity threshold = 5 voxels. BA, Brodmann area, TT, Talairach and Tournoux atlas.

ROI Location	Laterality	BA	Peak *t* value	Peak TT coords.	*n* Voxels
Fusiform gyrus / Culmen	R	37	7.27	[49,-43,-23]	7
Middle Frontal Gyrus	L	11 / 47	4.36	[-29,40,-6]	5

D.1 ROI ANALYSIS SUMMARY

Table D2. ROI-Averaged Timecourse Analyses: *p* values, Right Fusiform Gyrus. See text for an explanation of the statistical analyses. M.E. = Main effect. Main effect of magnitude is omitted as these results are identical to those reported for the quadratic value-related trend. Bold font indicates an effect that is significant at $\alpha = .05$.

Task Component	linear trend	quad. Trend	M.E. of valence	val. *x* mag. intxn.
Novel Previews	**2.86×10^{-5}**	.83	**8.37×10^{-5}**	.19
Old Previews	.29	.79	**.048**	.45
Trial 1 Outcomes	.31	.31	.16	.39

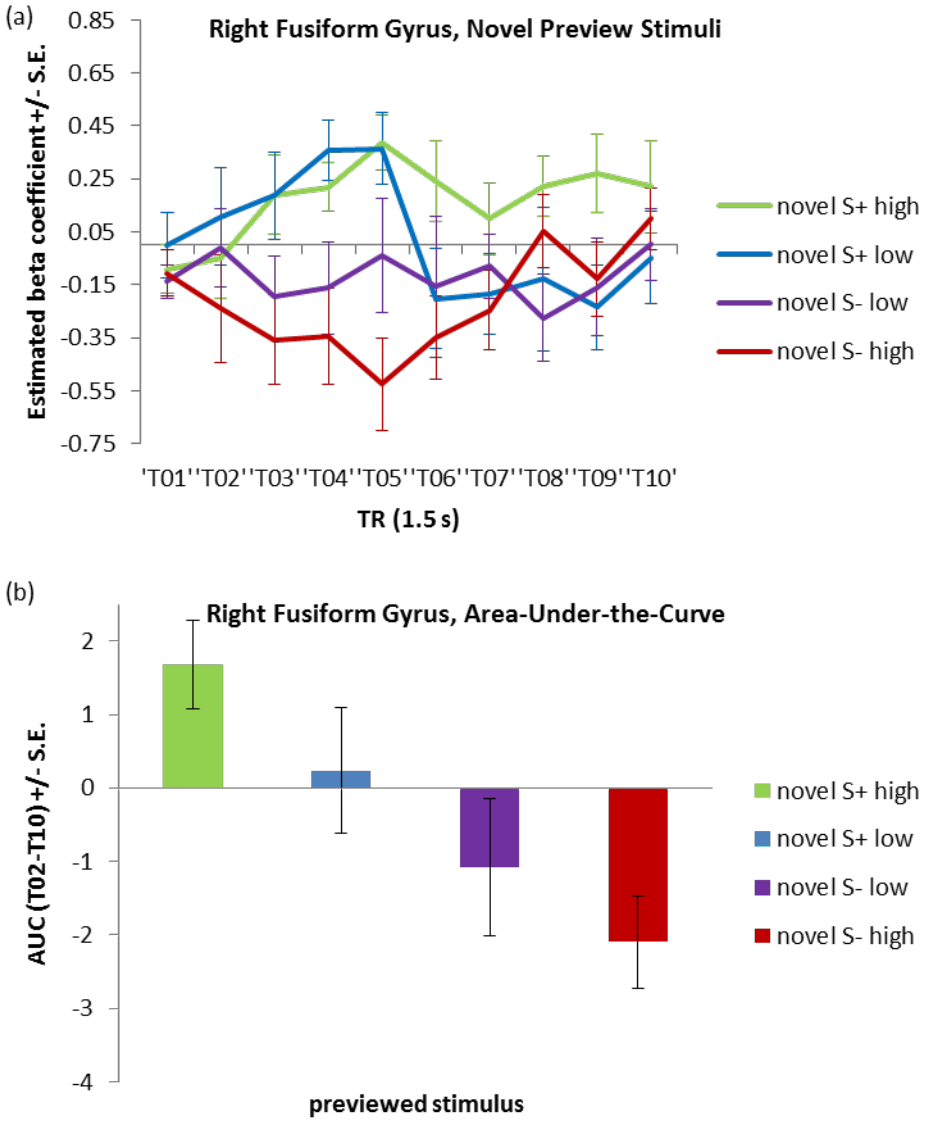


Figure D2. (a) Mean estimated hemodynamic responses extracted from the fusiform gyrus ROI that demonstrated a significant linear effect of novel preview stimulus value. (b) Mean area-under-the-curve estimates for the T02-T10 segment of the curves in Figure D2a.

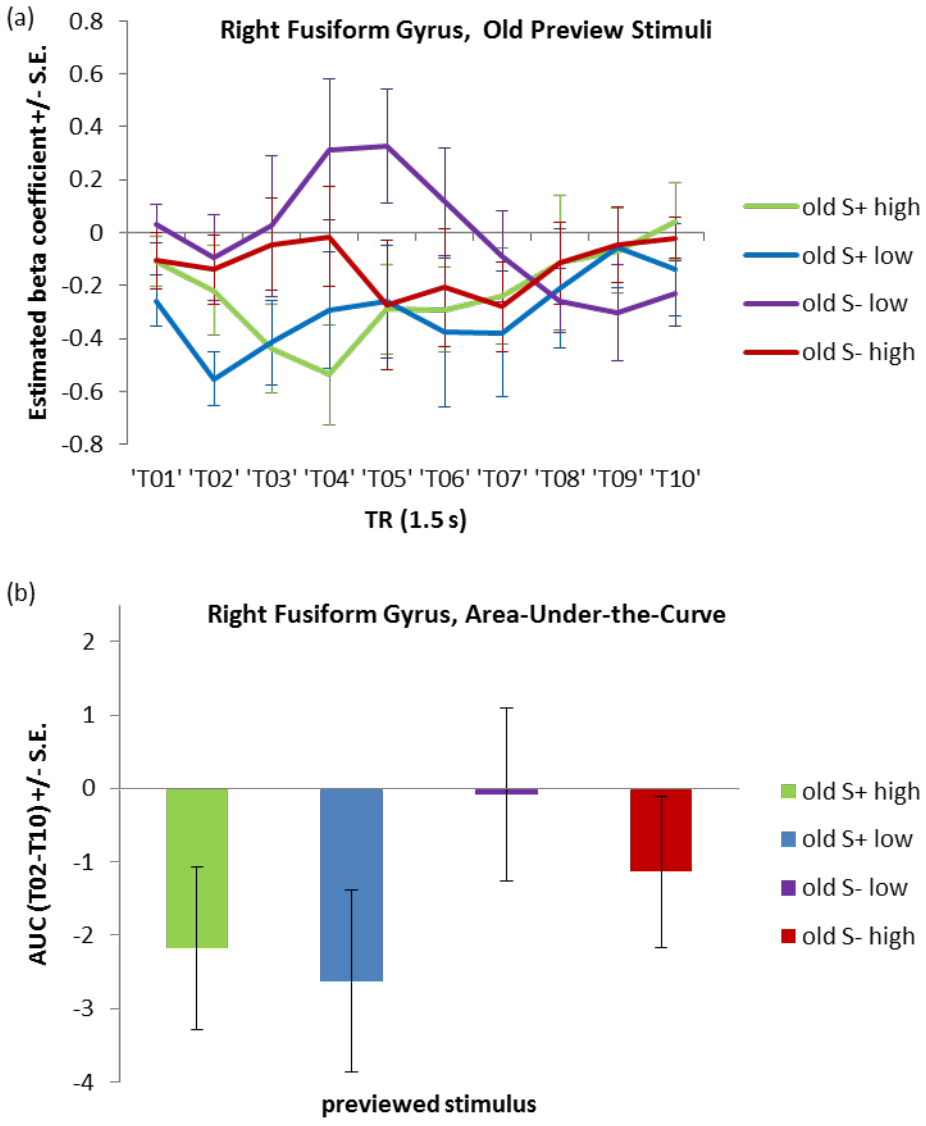


Figure D3. (a) Mean estimated hemodynamic responses to the old preview stimuli, extracted from the fusiform gyrus ROI defined through the linear contrast of novel preview stimulus value. (b) Mean area-under-the-curve estimates for the T02-T10 segment of the curves in Figure D3a.

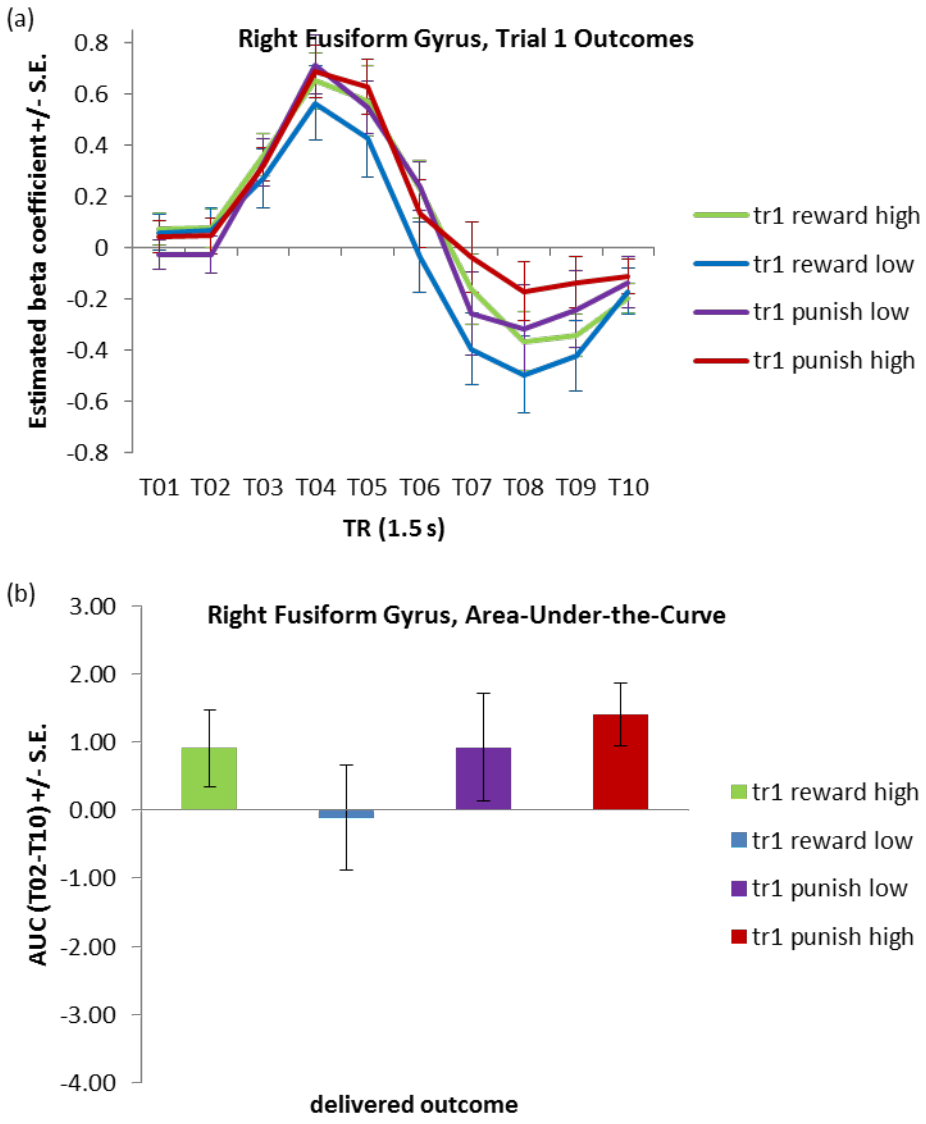


Figure D4. (a) Mean estimated hemodynamic responses to the Trial 1 choice and outcome displays, extracted from the fusiform gyrus ROI defined through the linear contrast of novel preview stimulus value. (b) Mean area-under-the-curve estimates for the T02-T10 segment of the curves in Figure D4a.

Table D3. ROI-Averaged Timecourse Analyses: *p* values, Left Middle Frontal Gyrus. Table conventions are the same as those described for Table D2.

Task Component	linear trend	quad. Trend	M.E. of valence	val. *x* mag. intxn.
Novel Previews	**1.77×10^{-4}**	.39	**.0044**	.39
Old Previews	.30	.84	.19	.66
Trial 1 Outcomes	.080	.91	.16	.92

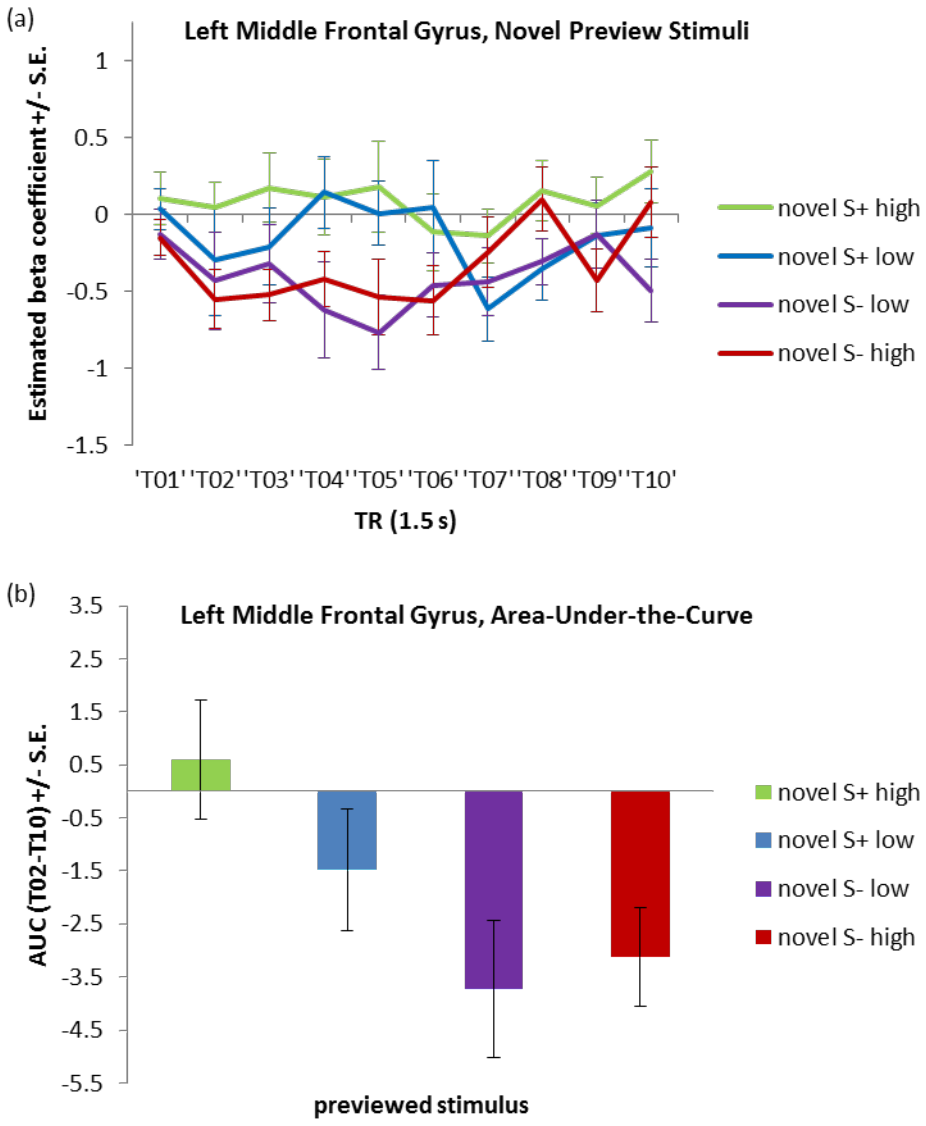


Figure D5. (a) Mean estimated hemodynamic responses extracted from the middle frontal gyrus ROI that demonstrated a significant linear effect of novel preview stimulus value. (b) Mean area-under-the-curve estimates for the T02-T10 segment of the curves in Figure D5a.

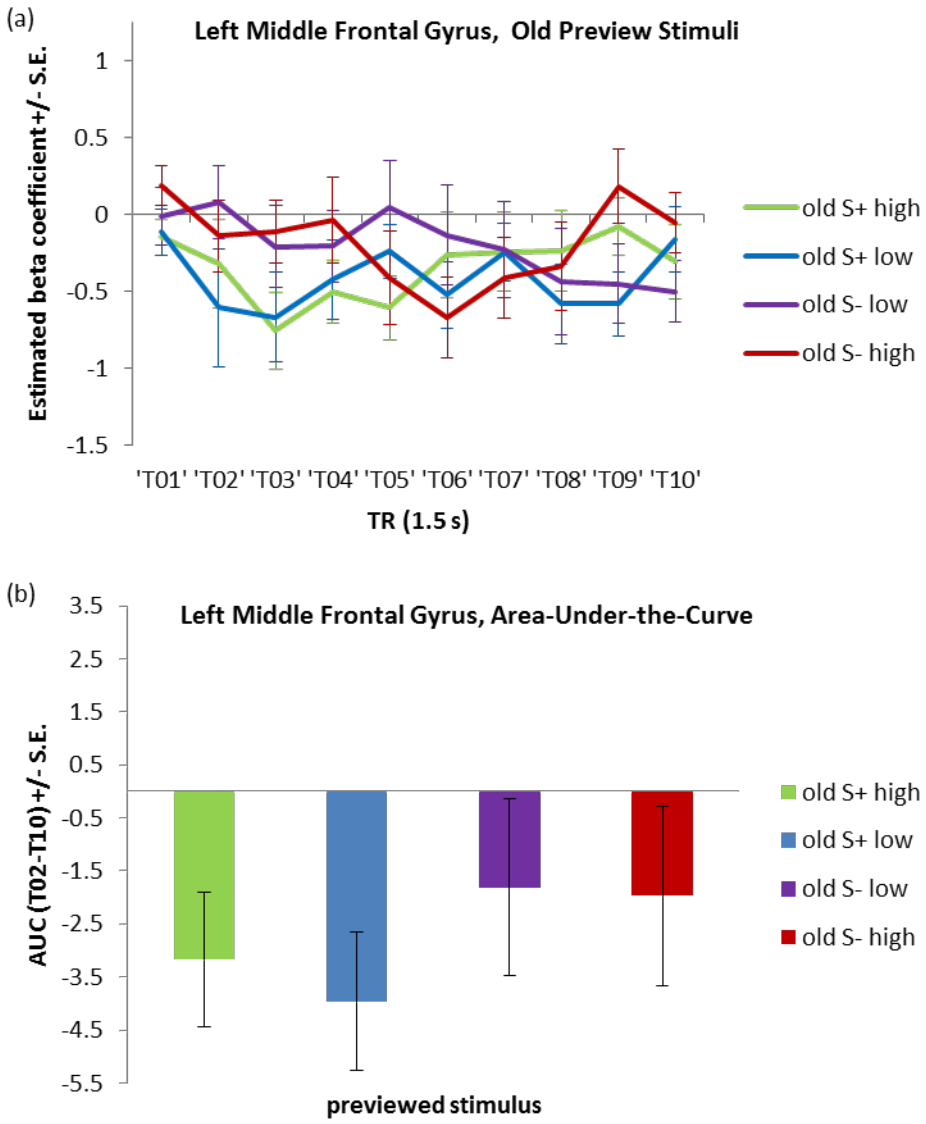


Figure D6. (a) Mean estimated hemodynamic responses to the old preview stimuli, extracted from the middle frontal gyrus ROI defined through the linear contrast of novel preview stimulus value. (b) Mean area-under-the-curve estimates for the T02-T10 segment of the curves in Figure D6a.

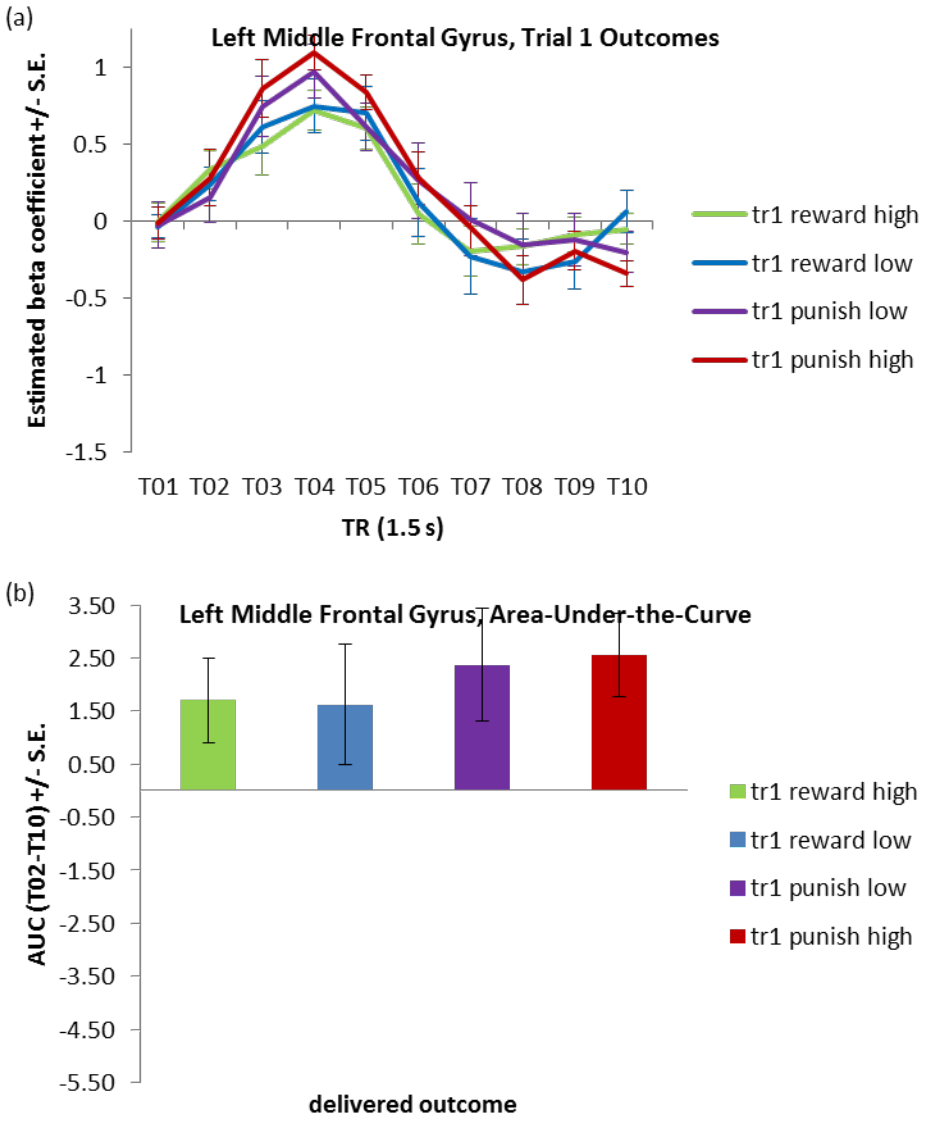


Figure D7. (a) Mean estimated hemodynamic responses to the Trial 1 choice and outcome displays, extracted from the middle frontal gyrus ROI defined through the linear contrast of novel preview stimulus value. (b) Mean area-under-the-curve estimates for the T02-T10 segment of the curves in Figure D7a.

D.2 ROI ANALYSIS DISCUSSION

Follow up analysis of the ROI-averaged data from the right fusiform ROI suggested that the linear relationship between activity in this region and the values of the novel preview stimuli was primarily driven by the main effect of valence (which, unlike the valence *x* magnitude interaction, was statistically significant in the 2-way ANOVA analysis). The other informative aspect of the follow-up ROI analyses came from the old-preview analyses, which similarly revealed a main effect of valence, albeit one which occurred in the opposite direction (old S- > old S+).

How should this pattern of effects be interpreted? Two aspects of the data suggest a similar interpretation might apply to this ROI as was described in the Discussion to explain the results in the lingual gyrus (Section 1.3.1.4). First, the fusiform gyrus shares the general status of the lingual gyrus as a region which was likely to be important for processing the surface properties of the card patterns; previous research using the FMR-adaptation approach showed this region to be responsive to changes in both stimulus color and stimulus shape (Cavina-Pratesi, et al., 2010). Second, the S- > S+ ordering of the old-preview valence effect in the fusiform is the same as the ranking of the valence effects in the lingual gyrus. Therefore, one possible interpretation of the fusiform effects, at least within the old-preview data *per se*, would appeal to the occurrence of something akin to a "tip-of-the-tongue" effect in which extra visual imagery is employed in the attempt to retrieve a forgotten stimulus label (i.e., its positive or negative valence). Although the MEG study that supported this TOT explanation for the lingual gyrus (Lindin, et al., 2010) did not report the same TOT effect in the fusiform, other areas of the literature support the general notion that the fusiform gyrus might be recruited during prolonged attempts to determine the identity of an ambiguous stimulus (e.g., Ploran et al., 2007).

A somewhat more difficult challenge involves the explanation of the novel S+ > S- effect in the fusiform: If the TOT state applies to stimuli which are known but for which a label cannot be recollected, then how could such a state be generated by a subset of the stimuli which are, at least by objective definition, "new"? Since considerable evidence throughout the text suggests that participants may have more reliably attended to and/or encoded the initially-unsampled stimulus when this was an S+ as opposed to an S-, one possibility is that the novel S- stimuli were more likely to have been appropriately judged as "new" than the novel S+ stimuli were (since participants would be more likely to recognize an "N-" as not identical to the successfully-encoded, initially-unsampled S+ stimulus). In other words, when an N+ was viewed, it may have been relatively difficult for the participants to judge whether this was truly a "new" stimulus or simply the appearance of an initially-unsampled S- that they failed to recognize as familiar. If the N+ stimuli were occasionally, inappropriately judged to be "known" stimuli, then this may have created a situation similar to the TOT state, and triggered similar invocation of visual imagery processes in an attempt to retrieve the appropriate valence label for the stimulus.

Follow-up analyses of the left middle frontal gyrus (lMFG) ROI did not reveal any significant effects beyond the expected linear value-related trend and main effect of valence for the novel preview stimuli. The qualitative activity pattern that this region exhibited during the old preview stimuli was suggestive of a main effect of stimulus valence (old S- > old S+, as seen with the fusiform), but this effect was not significant. A marginally-significant, inverse linear relationship with the values of the Trial 1 outcomes was found.

Overall, the lMFG responses are consistent with the general memory-related interpretations applied to the fusiform and lingual gyri, particularly if one assumes that a reliable old-preview valence effect would have emerged under conditions of greater statistical power.

Left MFG has been implicated in the retrieval of visual information (e.g., Wheeler, Petersen, & Buckner, 2000) and has been associated with greater activation in the presence of increased retrieval effort (e.g., Maril, Wagner, & Schacter, 2001). Therefore, a similar interpretation of the N+ > N- effect as was suggested for the fusiform might also be appropriate in the case of the MFG; that is, the relative sensitivity of the lMFG to the N+ stimuli might be explained in terms of participants' effortful (and mistaken) attempts to retrieve past information regarding these stimuli. Previous evidence documenting the involvement of the lMFG during the encoding of visual images (e.g., Wheeler, et al., 2000) may also provide insight into the marginally-significant, inverse association of lMFG activity and Trial 1 outcome value, if one assumes that the tendency for greater accuracy following Trial 1 S- selections is mediated by more successful encoding of the initially-unsampled S+ stimulus.

APPENDIX E

VALUE-RELATED ACTIVITY DURING TRIAL 1 OUTCOME DELIVERY: RESULTS

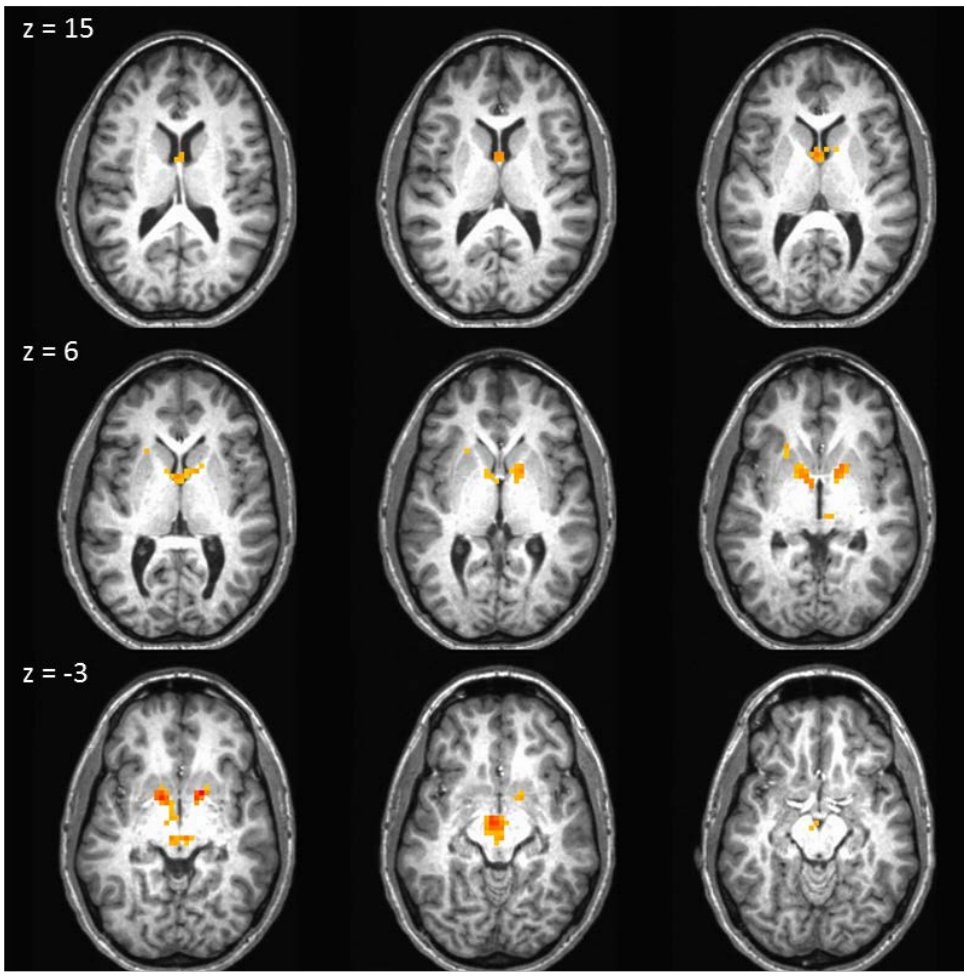


Figure E1. Striatum, globus pallidus, and midbrain regions in which activity exhibited a direct linear relationship with the values of the delivered Trial 1 outcomes ($p < .001$, contiguity threshold = 5 voxels). Images are presented according to radiological convention (right = left).

Table E1. Linear trends related to the values of the Trial 1 outcomes, $p < .001$, contiguity threshold = 5 voxels. BA, Brodmann area, TT, Talairach and Tournoux atlas.

ROI Location	Laterality	BA	Peak *t* value	Peak TT coords.	*n* Voxels
Striatum, Globus Pallidus, Midbrain	L, R	n/a	6.58	[-15,4,3]	114
Putamen	R	n/a	4.51	[22,20,0]	5

REFERENCES

Badre, D. (2008). Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends in cognitive sciences, 12*(5), 193-200. doi: 10.1016/j.tics.2008.02.004

Badre, D., Kayser, A. S., & D'Esposito, M. (2010). Frontal cortex and the discovery of abstract action rules. *Neuron, 66*(2), 315-326. doi: 10.1016/j.neuron.2010.03.025

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral cortex, 10*(3), 295-307.

Bhanji, J. P., Beer, J. S., & Bunge, S. A. (2010). Taking a gamble or playing by the rules: dissociable prefrontal systems implicated in probabilistic versus deterministic rule-based decisions. *NeuroImage, 49*(2), 1810-1819. doi: 10.1016/j.neuroimage.2009.09.030

Boksem, M. A., Meijman, T. F., & Lorist, M. M. (2006). Mental fatigue, motivation and action monitoring. *Biological psychology, 72*(2), 123-132. doi:10.1016/j.biopsycho.2005.08.007

Boorman, E. D., Behrens, T. E., & Rushworth, M. F. (2011). Counterfactual choice and learning in a neural network centered on human lateral frontopolar cortex. *PLoS biology, 9*(6), e1001093. doi: 10.1371/journal.pbio.1001093

Boorman, E. D., Behrens, T. E., Woolrich, M. W., & Rushworth, M. F. (2009). How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron, 62*(5), 733-743. doi: 10.1016/j.neuron.2009.05.014

Bouvier, S. E., & Engel, S. A. (2006). Behavioral deficits and cortical damage loci in cerebral achromatopsia. *Cerebral cortex, 16*(2), 183-191. doi: 10.1093/cercor/bhi096

Browning, P. G., Easton, A., & Gaffan, D. (2007). Frontal-temporal disconnection abolishes object discrimination learning set in macaque monkeys. *Cerebral cortex, 17*(4), 859-864. doi: 10.1093/cercor/bhk039

Browning, P. G., & Gaffan, D. (2008). Prefrontal cortex function in the representation of temporally complex events. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 28*(15), 3934-3940. doi: 10.1523/JNEUROSCI.0633-08.2008

Bunge, S. A., Kahn, I., Wallis, J. D., Miller, E. K., & Wagner, A. D. (2003). Neural circuits subserving the retrieval and maintenance of abstract rules. *Journal of neurophysiology, 90*(5), 3419-3428. doi: 10.1152/jn.00910.2002

Burgess, P. W., Dumontheil, I., & Gilbert, S. J. (2007). The gateway hypothesis of rostral prefrontal cortex (area 10) function. *Trends in cognitive sciences, 11*(7), 290-298. doi: 10.1016/j.tics.2007.05.004

Cavina-Pratesi, C., Kentridge, R. W., Heywood, C. A., & Milner, A. D. (2010). Separate channels for processing form, texture, and color: evidence from FMRI adaptation and visual object agnosia. *Cerebral cortex, 20*(10), 2319-2332. doi: 10.1093/cercor/bhp298

Charron, S., & Koechlin, E. (2010). Divided representation of concurrent goals in the human frontal lobes. *Science, 328*(5976), 360-363. doi: 10.1126/science.1183614

Chib, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 29*(39), 12315-12320. doi: 10.1523/JNEUROSCI.2575-09.2009

Cho, S., Moody, T. D., Fernandino, L., Mumford, J. A., Poldrack, R. A., Cannon, T. D., . . . Holyoak, K. J. (2010). Common and dissociable prefrontal loci associated with component mechanisms of analogical reasoning. *Cerebral cortex, 20*(3), 524-533. doi: 10.1093/cercor/bhp121

Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J. K., Holyoak, K. J., & Gabrieli, J. D. (2001). Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *NeuroImage, 14*(5), 1136-1149. doi: 10.1006/nimg.2001.0922

Cox, K. M., & Fiez, J. A. (2010). *Decision-making guided by the inferred value of unsampled stimuli*. Paper presented at the Annual Meeting of the Cognitive Neuroscience Society, Montreal, Canada.

Cox, R. W. (2010). AFNI deconvolution, from http://afni.nimh.nih.gov/pub/dist/edu/latest/afni06_decon/

Daw, N. D., & Doya, K. (2006). The computational neurobiology of learning and reward. *Current opinion in neurobiology, 16*(2), 199-204. doi: 10.1016/j.conb.2006.03.006

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron, 69*(6), 1204-1215. doi: 10.1016/j.neuron.2011.02.027

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature, 441*(7095), 876-879. doi: 10.1038/nature04766

Deets, A. C., Harlow, H. F., & Blomquist, A. J. (1970). Effects of intertrial interval and trial 1 reward during acquisition of an object-discrimination learning set in monkeys. *Journal of Comparative and Physiological Psychology, 73*(3), 501-505.

Deichmann, R., Gottfried, J. A., Hutton, C., & Turner, R. (2003). Optimized EPI for fMRI studies of the orbitofrontal cortex. *NeuroImage, 19*(2 Pt 1), 430-441.

Delgado, M. R., Locke, H. M., Stenger, V. A., & Fiez, J. A. (2003). Dorsal striatum responses to reward and punishment: effects of valence and magnitude manipulations. *Cognitive, affective & behavioral neuroscience, 3*(1), 27-38.

Delgado, M. R., Miller, M. M., Inati, S., & Phelps, E. A. (2005). An fMRI study of reward-related probability learning. *NeuroImage, 24*(3), 862-873. doi: 10.1016/j.neuroimage.2004.10.002

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine, 33*(5), 636-647.

Fuhs, M.C., & Touretzky, D.S. (2007). Context learning in the rodent hippocampus. *Neural computation, 19*(12), 3173-3215. doi:10.1162/neco.2007.19.12.3173

Gaffan, D. (1985). Hippocampus: memory, habit and voluntary movement. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 308*(1135), 87-99.

Gaffan, D., & Harrison, S. (1988). Inferotemporal-frontal disconnection and fornix transection in visuomotor conditional learning by monkeys. *Behavioural brain research, 31*(2), 149-163.

Genovesio, A., & Wise, S. P. (2008). The neurophysiology of abstract response strategies. In S. A. Bunge & J. D. Wallis (Eds.), *Neuroscience of rule-guided behavior* (pp. 81-106). New York: Oxford University Press.

Gilbert, A. M., & Fiez, J. A. (2004). Integrating rewards and cognition in the frontal cortex. *Cognitive, affective & behavioral neuroscience, 4*(4), 540-552.

Glascher, J., Hampton, A. N., & O'Doherty, J. P. (2009). Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cerebral cortex, 19*(2), 483-495. doi: 10.1093/cercor/bhn098

Glimcher, P. W., Camerer, C. F., Fehr, E., & Poldrack, R. A. (Eds.). (2009). *Neuroeconomics: Decision making and the brain*. London, UK: Elsevier.

Green, A. E., Kraemer, D. J., Fugelsang, J. A., Gray, J. R., & Dunbar, K. N. (2010). Connecting long distance: semantic distance in analogical reasoning modulates frontopolar cortex activity. *Cerebral cortex, 20*(1), 70-76. doi: 10.1093/cercor/bhp081

Grefkes, C., & Fink, G. R. (2005). The functional organization of the intraparietal sulcus in humans and monkeys. *Journal of anatomy, 207*(1), 3-17. doi: 10.1111/j.1469-7580.2005.00426.x

Hampton, A. N., Bossaerts, P., & O'Doherty, J. P. (2006). The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 26*(32), 8360-8367. doi: 10.1523/JNEUROSCI.1010-06.2006

Harlow, H. F. (1949). The formation of learning sets. *Psychological review, 56*(1), 51-65.

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical image analysis, 5*(2), 143-156.

Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature neuroscience, 10*(12), 1625-1633. doi: 10.1038/nn2007

Koechlin, E., & Hyafil, A. (2007). Anterior prefrontal function and the limits of human decision-making. *Science, 318*(5850), 594-598. doi: 10.1126/science.1142995

Kumaran, D., Summerfield, J. J., Hassabis, D., & Maguire, E. A. (2009). Tracking the emergence of conceptual knowledge during human decision making. *Neuron, 63*(6), 889-901. doi: 10.1016/j.neuron.2009.07.030

Levine, M. (1959). A model of hypothesis behavior in discrimination learning set. *Psychological review, 66*, 353-366.

Levine, M., Levinson, B., & Harlow, H. F. (1959). Trials per problem as a variable in the acquisition of discrimination learning set. *Journal of comparative and physiological psychology, 52*, 396-398.

Lindin, M., Diaz, F., Capilla, A., Ortiz, T., & Maestu, F. (2010). On the characterization of the spatio-temporal profiles of brain activity associated with face naming and the tip-of-the-tongue state: a magnetoencephalographic (MEG) study. *Neuropsychologia, 48*(6), 1757-1766. doi: 10.1016/j.neuropsychologia.2010.02.025

Lockhart, J. M., Parks, T. E., & Davenport, J. W. (1963). Information Acquired in One Trial by Learning-Set Experienced Monkeys. *Journal of comparative and physiological psychology, 56*, 1035-1037.

Maril, A., Wagner, A. D., & Schacter, D. L. (2001). On the tip of the tongue: an event-related fMRI study of semantic retrieval failure and cognitive conflict. *Neuron, 31*(4), 653-660.

Murray, E. A., & Gaffan, D. (2006). Prospective memory in the formation of learning sets by rhesus monkeys (Macaca mulatta). *Journal of experimental psychology. Animal behavior processes, 32*(1), 87-90. doi: 10.1037/0097-7403.32.1.87

Nieuwenhuis, S., Heslenfeld, D. J., von Geusau, N. J., Mars, R. B., Holroyd, C. B., & Yeung, N. (2005). Activity in human reward-sensitive brain areas is strongly context dependent. *NeuroImage, 25*(4), 1302-1309.

Ollinger, J. M., Corbetta, M., & Shulman, G. L. (2001). Separating processes within a trial in event-related functional MRI. *NeuroImage, 13*(1), 218-229. doi: 10.1006/nimg.2000.0711

Palminteri, S., Boraud, T., Lafargue, G., Dubois, B., & Pessiglione, M. (2009). Brain hemispheres selectively track the expected value of contralateral options. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 29*(43), 13465-13472. doi: 10.1523/JNEUROSCI.1500-09.2009

Pan, X., Sawa, K., Tsuda, I., Tsukada, M., & Sakagami, M. (2008). Reward prediction based on stimulus categorization in primate lateral prefrontal cortex. *Nature neuroscience, 11*(6), 703-712. doi: 10.1038/nn.2128

Ploran, E. J., Nelson, S. M., Velanova, K., Donaldson, D. I., Petersen, S. E., & Wheeler, M. E. (2007). Evidence accumulation and the moment of recognition: dissociating perceptual recognition processes using fMRI. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 27*(44), 11912-11924. doi: 10.1523/JNEUROSCI.3522-07.2007

Ramnani, N., & Owen, A. M. (2004). Anterior prefrontal cortex: insights into function from anatomy and neuroimaging. *Nature reviews. Neuroscience, 5*(3), 184-194. doi: 10.1038/nrn1343

Rygula, R., Walker, S. C., Clarke, H. F., Robbins, T. W., & Roberts, A. C. (2010). Differential contributions of the primate ventrolateral prefrontal and orbitofrontal cortex to serial reversal learning. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 30*(43), 14552-14559. doi: 10.1523/JNEUROSCI.2631-10.2010

Simmons, W. K., Ramjee, V., Beauchamp, M. S., McRae, K., Martin, A., & Barsalou, L. W. (2007). A common neural substrate for perceiving and knowing about color. *Neuropsychologia, 45*(12), 2802-2810. doi: 10.1016/j.neuropsychologia.2007.05.002

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.

Thompson, R. K. R., & Oden, D. L. (2000). Categorical perception and conceptual judgments by nonhuman primates: The paleological monkey and the analogical ape. *Cognitive Science, 24*(3), 363-396.

Tsujimoto, S., Genovesio, A., & Wise, S. P. (2011). Frontal pole cortex: encoding ends at the end of the endbrain. *Trends in cognitive sciences, 15*(4), 169-176. doi: 10.1016/j.tics.2011.02.001

Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., MIshkin, F. S., de Menezes Santos, M., . . . Miller, B. L. (1999). A system for relational reasoning in human prefrontal cortex. *Psychological science, 10*(2), 119-125.

Ward, B. D. (1998). Deconvolution analysis of fMRI time series data.

Ward, B. D. (2006). Analysis of variance for fMRI data. Milwaukee, WI: Medical College of Wisconsin, Biophysics Research Institute.

Wheeler, M. E., Petersen, S. E., & Buckner, R. L. (2000). Memory's echo: vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences of the United States of America, 97*(20), 11125-11129.

Woods, R. P., Grafton, S. T., Holmes, C. J., Cherry, S. R., & Mazziotta, J. C. (1998). Automated image registration: I. General methods and intrasubject, intramodality validation. *Journal of computer assisted tomography, 22*(1), 139-152.

Woods, R. P., Grafton, S. T., Watson, J. D., Sicotte, N. L., & Mazziotta, J. C. (1998). Automated image registration: II. Intersubject validation of linear and nonlinear models. *Journal of computer assisted tomography, 22*(1), 153-165.

Wunderlich, K., Rangel, A., & O'Doherty, J. P. (2009). Neural computations underlying action-based decision making in the human brain. *Proceedings of the National Academy of Sciences of the United States of America, 106*(40), 17199-17204. doi: 10.1073/pnas.0901077106

Wunderlich, K., Rangel, A., & O'Doherty, J. P. (2010). Economic choices can be made using only stimulus values. *Proceedings of the National Academy of Sciences of the United States of America, 107*(34), 15005-15010. doi: 10.1073/pnas.1002258107

Zaghloul, K. A., Blanco, J. A., Weidemann, C. T., McGill, K., Jaggi, J. L., Baltuch, G. H., & Kahana, M. J. (2009). Human substantia nigra neurons encode unexpected financial rewards. *Science, 323*(5920), 1496-1499. doi: 10.1126/science.1167342