JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 2005, **84,** 537–554 NUMBER 3 (NOVEMBER)

PREFRONTAL BRAIN ACTIVITY PREDICTS TEMPORALLY EXTENDED DECISION-MAKING BEHAVIOR

Tal Yarkoni, Todd S. Braver, Jeremy R. Gray, and Leonard Green

WASHINGTON UNIVERSITY AND YALE UNIVERSITY

Although functional neuroimaging studies of human decision-making processes are increasingly common, most of the research in this area has relied on passive tasks that generate little individual variability. Relatively little attention has been paid to the ability of brain activity to predict overt behavior. Using functional magnetic resonance imaging (fMRI), we investigated the neural mechanisms underlying behavior during a dynamic decision task that required subjects to select smaller, short-term monetary payoffs in order to receive larger, long-term gains. The number of trials over which the long-term gains accrued was manipulated experimentally (2 versus 12). Event-related neural activity in right lateral prefrontal cortex, a region associated with high-level cognitive processing, selectively predicted choice behavior in both conditions, whereas insular cortex responded to fluctuations in amount of reward but did not predict choice behavior. These results demonstrate the utility of a functional neuroimaging approach in behavioral psychology, showing that (a) highly circumscribed brain regions are capable of predicting complex choice behavior, and (b) fMRI has the ability to dissociate the contributions of different neural mechanisms to particular behavioral tasks.

Key words: brain–behaviour, decision making, functional magnetic resonance imaging, maximizing, temporal integration, choice, humans

The neural mechanisms governing learning and reward-seeking behavior in animals have been investigated extensively (for reviews, see Berridge, 2003; Schultz, 2000, 2002). Aided by neuroimaging technologies such as functional MRI, neuroscientists now are able to overcome many of the ethical and technical barriers that once prevented such research in humans, resulting in a dramatic rise in the number of studies employing experimental paradigms that examine behavioral choice processes and the economic factors that influence them. By associating gains or losses with the occurrence of specific events during experimental scanning sessions, researchers have examined the neural mechanisms underlying several aspects of decision making (for reviews, see Krawczyk, 2002; Montague & Berns, 2002; Schultz, 2000, 2002). Recent research indicates, for example, that dissociable neural mechanisms are involved in the experience versus the anticipation of reward (Pagnoni, Zink, Montague, & Berns, 2002); that the scale of a neural response tends to corresond to the magnitude of a gain or loss (Elliott, Newman, Longe, & Deakin, 2003); that the predictability of a reward influences the neural response (Berns, McClure, Pagnoni, & Montague, 2001); and that generalized, conditioned reinforcers (e.g., money) recruit many of the same brain regions that represent the value of primary reinforcers (O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001).

Such results are informative to neuroscientists, whose goal is to develop models of brain function. From the perspective of researchers primarily interested in predicting and modifying overt behavior, however, the heavy reliance of these studies on passive tasks (e.g., observing a sequence of predetermined rewards without the subject's behavioral interaction) or on tasks with little individual variability (e.g., simple gambling tasks or cognitive tasks where performance is at ceiling) limits their usefulness in clarifying the relationship between brain activity and behavior. It is common for cognitive neuroscientists to report significant differences in brain activity without corresponding behavioral changes (Wilkinson & Halligan, 2004). In contrast, what behavioral psychologists are likely to desire from a neuroscientific approach is evidence that brain activity can predict behavior at the individual level, rather than simply displaying aggregated sensitivity to a variety of experimental manipulations or reward payoffs.

This research was supported by a grant from the McDonnell Center for Higher Brain Function.

Tal Yarkoni, Todd Braver, and Leonard Green are at Washington University; Jeremy Gray is at Yale University.

Address correspondence to: Tal Yarkoni, Washington University, Department of Psychology, Campus Box 1125, St. Louis, Missouri 63130 (e-mail: tyarkoni@wustl.edu).

doi: 10.1901/jeab.2005.121-04

Increasingly, cognitive neuroscientists have been rising to the challenge, complementing group-based analysis with analysis of individual-difference variables. The clearest sign that neuroscience and behavioral psychology are converging is the growing number of neuroimaging studies aimed at investigating the extent to which the activity of circumscribed brain regions directly predict behavior. For example, Sanfey, Rilling, Aronson, Nystrom, and Cohen (2003) scanned subjects while they were playing an Ultimatum Game in which one player, the *proposer*, is allotted an amount of money, any portion of which can be offered to another person, the *responder*, who can either accept or reject the offer. If the responder accepts, then she receives the amount offered, and the proposer gets the remainder. If she rejects the offer, however, then neither receives anything. Sanfey et al. found that, in the responder, both emotional and cognitive regions were activated in response to what the responder considered to be an unfair offer. Notably, individual differences in the activation of the insula—an area associated with the experience and recognition of disgust—were correlated with individual differences in the rejection of such offers. Rilling et al. (2002) found that activation of striatal regions associated with reward predicted cooperation in an iterated Prisoner's Dilemma game. In our own work, we have shown that lateral prefrontal and parietal brain regions associated with higher cognitive processing (Cabeza & Nyberg, 2000; Smith & Jonides, 1999) selectively mediate the relationship between fluid intelligence (a trait measure of reasoning and problem-solving ability) and cognitive-task performance under high interference conditions (Gray, Chabris, & Braver, 2003).

The purpose of the present paper is to extend recent work at the boundary of cognitive neuroscience and behavioral psychology by applying functional neuroimaging to a complex decision-making paradigm. We emphasize the relationship between brain activation and overt behavior, and so primarily report analyses that link neural activity directly to performance. We also wish to move beyond simple (correlational) prediction, however, and use our analyses to illustrate general principles and issues that arise when trying to relate brain and behavior. The promise of fMRI for helping to understand behavior, we argue, lies not merely in the ability to identify parts of the brain that support certain kinds of behaviors, but also in its ability to provide more detailed insights regarding the temporal dynamics and mechanisms involved in producing those behaviors.

To examine the relation between brain activity and behavior, we adapted a behavioral decision-making task (Gray, 1999; Herrnstein, Loewenstein, Prelec, & Vaughan, 1993; Kudadjie-Gyamfi & Rachlin, 1996) in which subjects repeatedly sample from two decks of cards, one of which maximizes reward in the long run, the other of which minimizes reward in the long run. Crucially, these maximal–minimal effects of choosing from the decks are delayed, and sampling from the good deck results in a noticeable immediate *decrease* in reward, whereas sampling from the bad deck leads to an immediate increase in reward. Thus, subjects have to overcome the immediate, short-term gain associated with choice of the bad deck in favor of the delayed, long-term gains associated with choice of the good deck. The choice of an immediate, smaller reward over a delayed, larger reward is said to exemplify impulsivity, whereas choice of the delayed, larger reward is said to exemplify self-control (Ainslie, 2001; Green & Myerson, 1993; Logue, 1988; Rachlin & Green, 1972). The dynamic nature of this task (i.e., the fact that the reward function interacts with a subject's choice history) sets it apart from most other neuroimaging decision-making studies, which usually employ static reward contingencies that do not reflect the dynamic nature of most real-world environments.

The difficulty of the task was varied by changing the number of trials (the *window size*) over which the long-term effects accrue. When the window size is small, incremental changes in reward are large and thus relatively easy to detect; however, as the window size grows, the changes in reward associated with the delayed effect become smaller and, as a consequence, are more difficult to detect. In the present study, we used two window-size conditions: an easy condition, in which the delayed effects accrued quickly (over 2 trials; termed the short-window condition), and a more difficult condition, in which the delayed effects accrued slowly (over 12 trials; termed the long-window condition). Previous work (including our own unpublished pilot data) using similar parameters indicated that these values were likely to induce substantial differences in behavior, with subjects performing relatively well (i.e., choosing mostly from the immediately smaller, but long-term higher payoff deck) in the short-window condition and performing relatively poorly (i.e., choosing more often from the immediately larger, but long-term lower payoff deck) in the long-window condition (Herrnstein et al., 1993).

The present experiment had three goals: First, we wanted to determine whether activation in relatively circumscribed brain regions could predict complex decision-making behavior. Second, we predicted that manipulation of the window size over which the delayed effects accrue would modulate the brain-behavior relationship, but that some commonalities should remain (i.e., the same regions should predict performance in both the short- and long-window conditions, although perhaps to differing degrees). Finally, we sought to dissociate the influence of cognitive from affective or motivational neural activations on behavior. Cognitive factors are expected to drive behavior by influencing the ease with which relevant perceptual information can be extracted from stimuli and by modulating the amount or complexity of information that needs to be integrated in order to generate the optimal course of action. In contrast, affective factors influence the relative saliency of specific representations of rewards or punishers, and operate by biasing behavior in a more transient fashion. The relation between these systems is complex and interactive, but substantial evidence suggests that at least some components of the neural systems subserving cognitive and affective processing are separable (Bush, Luu, & Posner, 2000; Phan, Wager, Taylor, & Liberzon, 2002). However, in many cases it is difficult to assess the relative contribution of cognitive and affective factors to behavioral performance in complex task situations. Our goal was to illustrate that functional neuroimaging may be well suited to the decomposition of behavior into these component systems.

METHOD

Subjects

Subjects were 28 right-handed undergraduates at Washington University in St. Louis and individuals from the surrounding community (mean age = 22.4 years, SD = 3.6) recruited using flyers placed around the Washington University campus. All subjects consented to participate in return for financial remuneration ($25 per hr plus task earnings). Task earnings were redeemed at the rate of 1 cent for each point earned, and subjects were subsequently mailed a check for the total amount. Potential subjects with a neurological or psychiatric history were excluded from the experiment, which was approved by the Washington University IRB. Due to technical malfunctions during neuroimaging, 2 subjects did not complete both conditions. Results, therefore, are reported for the 26 subjects for whom we had complete data.

Behavioral Procedures

Prior to scanning, each subject was familiarized with the behavioral paradigm alone in a small room, facing a computer monitor. The computerized decision task was adapted from previous behavioral experiments (e.g., Herrnstein et al., 1993) and presented using the PsyScope software package (Cohen, MacWhinney, Flatt, & Provost, 1993). Subjects read the following instructions:

> The task is to draw cards from one of two decks to earn money. Each time you see a question mark, you should choose a card from the left or right deck. Some cards pay more money and others less money; money is never lost. The two decks are different, and there is a way to earn significantly more or less points. You will be paid real money for each point. Try to earn as much as possible.

On each trial a question mark was shown midway between two decks of cards. When one of two response keys was pressed, the corresponding deck was highlighted and the top card of that deck was ''turned over'' to reveal a number, representing the amount earned on that trial. The response time, deck chosen, and amount earned were recorded. Cumulative earnings were updated on each trial. Subjects were given practice trials until they indicated they understood and were comfortable with the procedure. This typically required approximately five trials and no more than 15 in any case. Note that the purpose of the practice was solely to familiarize subjects with the computer program; they were not exposed to the actual

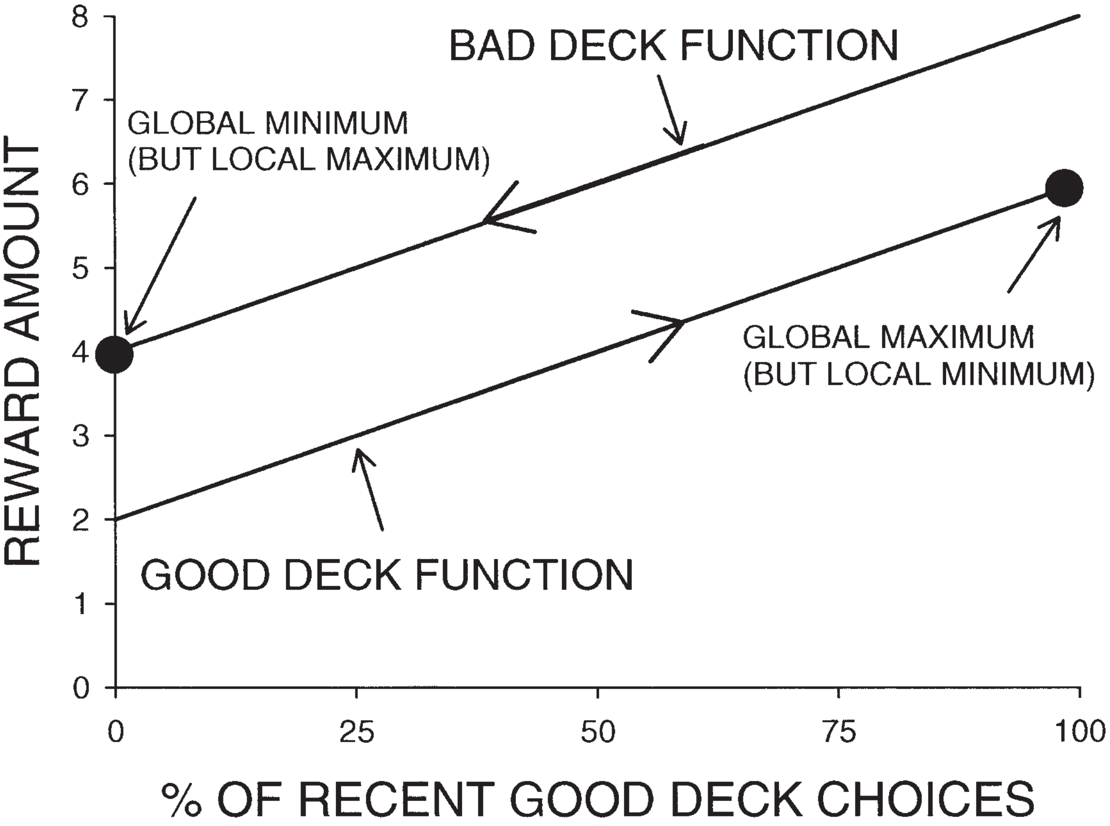

Fig. 1. Points earned as a function of the percentage of choices of the good deck in the previous 2 (short condition) or 12 (long condition) trials. Notice that the global and local maxima were in direct opposition: Choice of the bad deck produced an immediate larger but long-term lower payoff, whereas choice of the good deck led to an immediate smaller but long-term greater payoff (see text for full details).

contingencies used during the task (i.e., during the practice, all deck selections resulted in five points). Once practice was over, subjects were led into the scanner room and placed in the scanner. No further instructions were given.

The reward contingencies were programmed such that choice of one deck always paid out a larger reward than the other deck on that current trial, but gradually decreased the amount of the reward to be received from both decks over the next several trials. In contrast, choice of the other deck paid a smaller amount on that trial, but gradually increased the amount to be received from both decks over the next several trials. Thus, choice of the locally optimal deck, called the *bad* deck (i.e., the one with the higher immediate reward), led to the global minimum in terms of total earnings. Conversely, the globally larger reward was earned by choosing the deck that paid off less on any given trial, called the *good* deck because it led to the global maximum.

The payoff functions of the good and bad decks are shown as solid lines in Figure 1. Both lines have the same slope, but the line for the bad deck is two points higher than the line for the good deck, reflecting a switch-associated increase or decrease in points earned. That is, switching from the bad deck to the good deck moved the payoff from the higher line to the lower one, and resulted in an immediate two-point decrease in earnings on the switch trial. Conversely, switching from the good to the bad deck moved the payoff from the lower to the higher line and resulted in an immediate two-point increase. Repeated selection from the bad deck led to a decrease in earnings on each subsequent trial (i.e., the payoff moved leftward along the x axis), whereas repeated selection from the good deck led to an

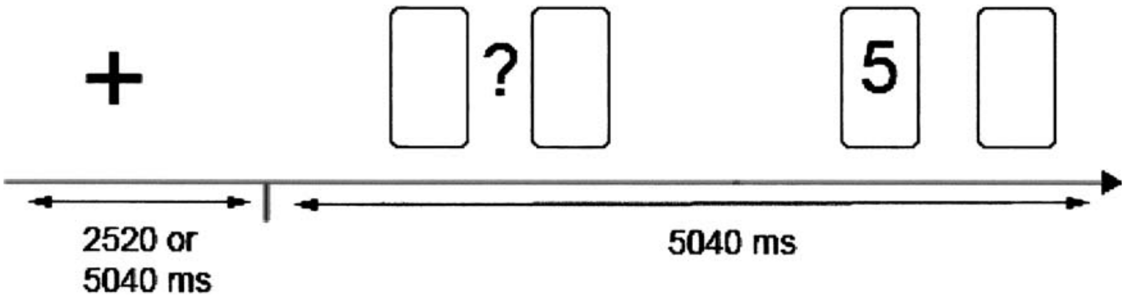

Fig. 2. Time course of an individual trial. Subjects viewed a fixation cross for either 2520 or 5040 ms (depending on random jittering), following which two decks appeared on-screen, and subjects made their selection. Each trial lasted 5040 ms. Subjects had up to 2520 ms within which to make their response, following which the deck they chose was immediately uncovered and the reward amount was displayed for the remainder of the trial.

increase in earnings on subsequent trials (a move rightward along the x axis). The long-term effect (i.e., one's position on the x axis) was computed as a function of the number of *good* deck choices a subject made over the last few trials. For example, if a subject had chosen from the good deck on 75% of the most recent trials, then she would be three-quarters of the way up on the reward function; if the subject had selected from the good deck on 25% of the most recent trials, then she would be a quarter of the way up on the reward function.

The ease with which subjects were expected to discern the underlying pattern and choose the good deck depended critically on the temporal-window parameter, *W*. This parameter was the number of trials over which amount of reward was computed. In Figure 1, the x axis refers to the percentage of good-deck choices made within the last *W* trials; thus, the subjects' locations along the abscissa were determined by dividing the number of trials on which they had selected the good deck (out of the last *W* trials) by *W*. The globally optimal strategy would be to select exclusively from the good deck. However, responding optimally required that the subject temporally integrate responses and their payoffs across multiple trials. The larger the value of *W*, the greater the difficulty in performing the task successfully since each move along the x axis was smaller and more difficult to discern as compared to the immediate effect of switching (e.g., when *W* is 2, it would take only 2 trials to move the entire length of the x axis, but when *W* is 12, it would take 12 trials).

W was set at 2 in one condition and 12 in the other. We refer to these as the short and long conditions, respectively. To ensure that the task would not be too easy, random fluctuations were introduced: On 40% of all trials, one point was randomly added to or subtracted from the amount earned. Performance was defined as the percentage of choices from the good deck (range: 0–100%; no consistent preference or chance performance = 50%). Because trials on which subjects failed to respond resulted in no payoff, response omissions (which totaled only 1.4% of trials across all subjects) were grouped with *bad* selections for the purpose of computing choice performance.

To accommodate fMRI scanning, subjects performed the task as eight runs of 40 card-choice trials each, with 160 trials at each temporal-window parameter. The sequence of events during a trial is shown in Figure 2. Following 2520 or 5040 ms of passively viewing a fixation cross (a variable inter-trial interval necessary for the fMRI design), two decks appeared on the screen. The decks remained on screen for 2520 ms or until the subject responded by pushing one of two buttons (mapped to the index and middle fingers of the subject's right hand) on a response apparatus to indicate his or her choice of a deck. Immediately following a choice, subjects saw how much they had earned on that trial; the outcome remained on the screen for the remainder of the trial, which totaled 5040 ms. All eight scanning runs were performed consecutively on the same day. After the first four runs (160 trials), subjects were told that they now were doing a task that looked the same but was, in fact, a different task and that they should approach it ''starting from scratch.'' The order in which subjects experienced the two conditions was counterbalanced.

Neuroimaging procedures

Imaging occurred while participants were lying prone within the bore of a head-only 3 Tesla Allegra System (Siemens, Erlangen, Germany) MRI scanner. Task stimuli were presented visually on a rear-projection display that was viewed with a mirror positioned directly above the eyes. Head movement during scanning (which can cause significant image artifacts) was minimized using pillows and tape. Functional images were acquired using an asymmetric spin-echo echo-planar sequence (TR = 2,520 ms, TE = 25 ms, flip = 90°), sensitive to blood-oxygen-level-dependent (BOLD) magnetic susceptibility. Each scanning run gave 165 sets of brain volumes (40 slices of 3.75 mm thickness), which allowed whole-brain coverage. Due to a change in acquisition software, the last 12 subjects were erroneously scanned with 1.875 mm spacing between slices. However, it is unlikely that this error substantially affected the results, since considerable smoothing of the fMRI data was employed, rendering minor gaps relatively inconsequential. Moreover, any effect would consist in a relative reduction in power rather than a systematic bias, and thus would not influence interpretation of the results (though it might artificially weaken them). Anatomical images were acquired using an MP-RAGE T1-weighted sequence, which enabled registration and localization of the functional data to brain anatomy. After movement and artifact correction, functional images were normalized within each scanning run and temporally aligned within each brain volume. Functional images were resampled into 3 mm isotropic voxels, transformed into atlas space, and smoothed with a Gaussian filter (9 mm FWHM).

A general linear model (GLM) approach was used to estimate brain activity during the decision-making task. We used a mixed blocked/event-related design in order to decompose brain activity into *state* and *item* components (Donaldson, Petersen, Ollinger, & Buckner, 2001; Visscher et al., 2003). The main conceptual point to note about this design is that it enables one to separate sustained brain activity related to demands imposed throughout the task (state-related activity) from transient brain activity associated with demands imposed specifically during individual trials (event-related activity). Previous research suggests that this design measures the operation of distinct functions (Braver, Reynolds, & Donaldson, 2003; Donaldson et al., 2001). In the present context, state-related activity was more likely to indicate persistent task demands such as maintenance of instructions or strategies, whereas event-related activity indexed the response to the transient demands of specific trials and thus likely indicated neural activity associated with rapid decision making, choice of deck, and the subject's neural response to individual reward payoff. A more detailed description of the neuroimaging analysis procedures used for this dataset can be found in a separate report (Yarkoni et al., 2005). More general treatments of neuroimaging methodology can be found elsewhere (e.g., Cabeza & Kingstone, 2001; Huettel, Song, & McCarthy, 2004).

RESULTS

We first verified that the manipulation of window size had the intended effect on deck selectivity, after which we analyzed subjects' acquisition of deck preferences and deck-switching behavior as a function of different task factors. Neuroimaging results are presented following the behavioral performance data. All results refer to data from the scanning sessions only.

Behavioral results

At both the group and individual levels, the window-size manipulation was successful. Mean percentage choice of the good deck in the short (window size = 2) and long (window size = 12) conditions was 71% and 23% across all trials, respectively, and each differed significantly from chance, $t(25) = 6.4$ and -12.5, respectively, both $ps < .001$, as well as from each other, $t(25) = 10.41$, $p < .001$. Figure 3 shows mean percentage choice in each window-size condition for each subject. Every subject performed better in the short condition than in the long condition (i.e., made more choices of the good deck, and earned more money). Additionally, there was an unanticipated negative correlation between

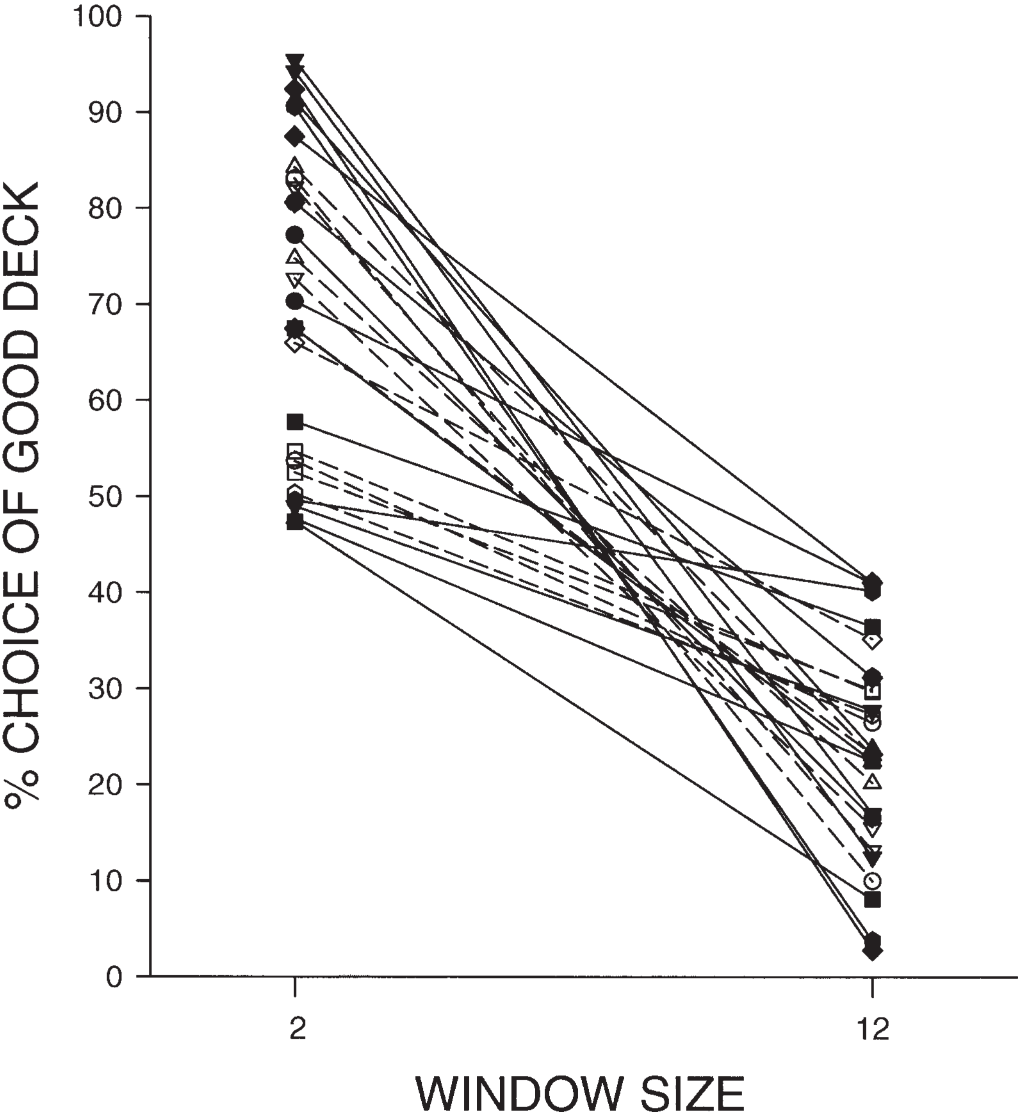

Fig. 3. Percentage choice of the good deck as a function of window size. Each line represents a single subject. Solid lines represent subjects who first were studied in the long ($W = 12$) condition; dashed lines represent those subjects first studied in the short ($W = 2$) condition.

performance in the two conditions, $r = -.41$, $p < .04$. As may be seen in Figure 3, subjects who chose the good deck more frequently in the short condition tended to choose the bad deck more frequently in the long condition.

Although all subjects performed better in the short condition than in the long condition, there was a marked asymmetry in the degree of susceptibility to the experimental manipulation across the two window-size conditions. Applying the binomial threshold for a departure from chance significant at the $p < .01$ level over the 160 trials in the short condition (i.e., choice of the good deck > 60%) resulted in a group of 17 subjects who performed significantly above chance and 9 who did not. In contrast, application of the same threshold to the long condition (choice of the good deck < 40%) indicated that 23 of 26 subjects performed significantly more poorly than chance. Eight of the 9 subjects who failed to show an effect in the short condition (i.e., choose from the good deck at greater-than-chance levels) nonetheless showed an effect in the long condition (i.e., choose from the bad deck at greater-than-chance levels), and only 1 subject failed to show an effect in both conditions. Thus, the negative correlation between the two conditions likely was not attributable to a failure of the manipulation to affect choices for some subjects. An alternative possibility was that differences in susceptibility to the manipulation across subjects might have been due to the fact that the order in which subjects performed the two conditions was counterbalanced. To evaluate this possibility, we regressed performance in each condition on an order variable. Results were nonsignificant in both cases, $p > .10$ (see also Figure 3; compare the solid lines, which represent choice of the good deck by subjects first studied in the long condition, with the dashed lines, which represent choice of the good deck by subjects first studied in the short condition). Thus, the negative correlation in performance across the two conditions appears to be due to individual differences and not to a failure of the manipulation or to order effects.

To determine whether individuals differed in the rate at which they acquired a preference for a deck in each condition, we examined choice of the good deck across the 160 trials of each condition. As Figure 4A shows, overall, subjects acquired their preferences quickly: in the short condition, mean percentage choice of the good deck differed significantly from chance after 20 trials, $t(25) = 2.43$, $p < .03$; in the long condition, only 10 trials were required before choice of the good deck was significantly below chance, $t(25) = -4.03$, $p < .001$. Once acquired, preference remained significantly different from chance throughout the remainder of the condition. There were notable differences, however, across subjects. Figure 4B shows mean percentage choice of the good deck under the short and long condition by those subjects who performed significantly above chance in the short condition. These subjects showed a statistically significant preference for the good deck within the first 20 trials in the short condition, $t(16) = 2.68$, $p < .02$, and a statistically significant preference for the bad deck within 10 trials in the long condition, $t(16) = 3.16$, p

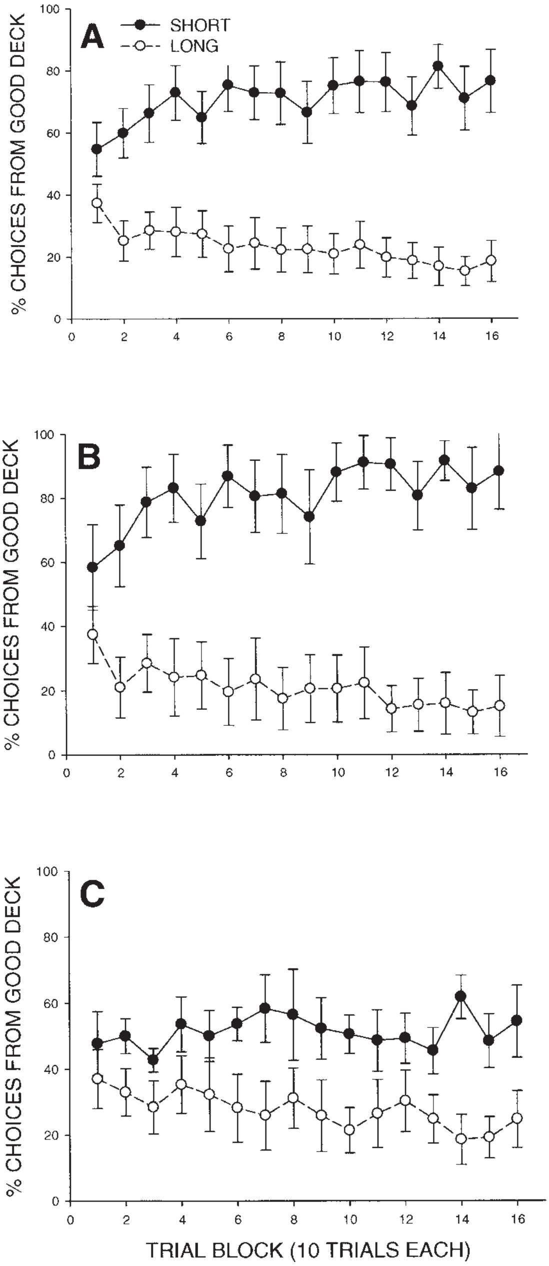

Fig. 4. (A) Mean percentage choice of the good deck in the short ($W = 2$) and long ($W = 12$) condition as a function of trial block. (B) Mean percentage choice of the good deck for the group of subjects whose performance differed significantly from chance in the short condition. (C) Mean percentage choice of the good deck for the group of subjects whose performance did not differ significantly from chance in the short condition. Error bars represent 95% confidence intervals.

Table 1

Proportion of trials (and standard deviations) on which subjects switched deck as a function of window size, change in reward amount, and last deck selected from. For example, the top-left cell (with a value of .31) indicates that, in the short condition, subjects who had just made a selection from the Good deck and received a smaller reward than on the previous trial selected from the Bad deck on the next trial 31% of the time. Values are derived from the 21 subjects who had a complete factorial table.

	Short Condition			Long Condition		
	Reward			Reward		
Deck	Decrease	No change	Increase	Decrease	No change	Increase
Good	.31 (.21)	.25 (.20)	.26 (.19)	.61 (.23)	.59 (.21)	.47 (.26)
Bad	.54 (.17)	.46 (.24)	.36 (.18)	.33 (.17)	.16 (.1)	.15 (.09)

$< .006$. In contrast, as can be seen in Figure 4C, the subjects who performed at chance level in the short condition by definition never developed a preference for the good deck in the short condition, $t(8) = 1.25$, $p > .10$, yet still showed a statistically significant preference for the bad deck within the first 10 trials in the long condition, $t(8) = 2.36$, $p < .05$. Regressing good-deck choice on the 160 trials in each condition (i.e., estimating the learning slope for each subject) yielded similar results. For high performers (i.e., those subjects who chose from the good deck at a rate significantly above chance level in the short condition), the regression coefficient differed significantly from chance in both conditions (short condition: mean = .19, $t(16) = 4.37$, $p < .001$; long condition: mean = −.15, $t(16) = -6.65$, $p < .001$), whereas for the low performers (i.e., those subjects who did not choose from the good deck at a rate above chance level in the short condition), the regression coefficient differed in the long condition (mean = −.1, $t(8) = -3.5$, $p < .008$), but not in the short condition (mean = .02, $t(8) = .78$, $p > .10$).

The results show that choice was influenced by the window-size manipulation. However, the rate of subjects' selections from the good deck was expected to vary systematically not just as a function of window size but also in response to changes in received reward amount. To evaluate whether subjects would be more likely to switch decks following a decrease in reward than after receiving an increase in reward (i.e., a win-stay/lose-shift [WS/LS] strategy), we conducted a 2 (window size: short vs. long) × 2 (last deck: good vs. bad) × 3 (reward change: increase, decrease, or no change) ANOVA with proportion of trials on which subjects switched decks as the dependent variable. Because 5 of the 26 subjects were missing trials in at least one cell in the design, we included only the 21 subjects who had a complete table of results. (Including the additional subjects where possible for individual main effects or two-way interactions did not change the pattern of results.)

Results of the analysis indicated significant main effects of deck ($F(1) = 7.17$, $p < .02$) and reward change ($F(2) = 12.53$, $p < .001$), which were qualified by a two-way window size × deck interaction ($F(1) = 119.32$, $p < .001$) and a three-way window size × deck × reward change interaction ($F(2) = 3.79$, $p < .04$). Table 1 displays the pattern of results. Subjects were more likely to switch decks following a decrease in reward (mean = .45) than following no change (mean = .37; $t(20) = 2.75$, $p < .02$) or an increase (mean = .31; $t(20) = 5.08$, $p < .001$) in reward, consistent with the use of a WS/LS strategy. Moreover, the probability of switching from a particular deck was modulated by the subjects' deck preference in each condition. In the short-window condition, subjects overall were more likely to switch from the bad deck (.45) than from the good deck (.27; $t(20) = 7.94$, $p < .001$), whereas in the long-window condition, subjects overall were more likely to switch from the good deck (.56) than from the bad deck (.21; $t(20) = 5.80$, $p < .001$). Finally, the three-way interaction suggests less differential susceptibility to changes in reward amount in the short-window condition when selecting from the good deck as compared to the other deck × window size conditions. That is, when selecting from the good deck in the short condition, there was a relatively small difference between the proportion of trials on which subjects switched following a decrease (.31)

Table 2

Talairach coordinates of all brain regions that either predicted behavioral choice or were sensitive to changes in reward amount. Coordinates indicate center of mass for each region. BA = Brodmann Area; mm^3 = volume of region in millimeters cubed.

		Coordinates				
Region	Hemisphere	BA	mm^3	x	y	Z
Regions predicting good deck choice in short condition						
Brainstem			2241	−2	−21	−3
Putamen	L		1026	−16	3	9
Pulvinar	L		837	−14	−24	12
Dorsal lateral PFC	R	9/46	2160	46	30	30
Superior lateral PFC	L	6/9	1512	−44	3	33
Superior lateral PFC	R	6/9	648	40	3	39
Inferior parietal cortex	L	39	567	−34	−60	36
Inferior parietal cortex	L	40	1323	−52	−42	−45
Medial prefrontal cortex	L	6/8	1566	−2	15	51
Regions predicting good deck choice in both conditions						
Temporal pole	L	36	1026	−16	3	−30
Superior lateral PFC	R	6/8	1161	44	18	45
Inferior parietal cortex	R	7	594	28	−72	48
Inferior parietal cortex	R	7	324	22	−78	51
Inferior parietal cortex	L	7	405	−32	−63	57
Regions sensitive to changes in reward amount						
Lateral cerebellum	R		351	28	−75	−42
Ventral temporal lobe	R	20	486	50	−33	−27
Medial orbitofrontal cortex	L	11	216	−14	45	−21
Temporal pole	L	38/47	432	32	12	−18
Anterior insula	L	13	648	−38	15	9
Inferior parietal cortex	R	40	648	46	−39	57

versus following an increase (.26) in reward. The corresponding difference was significantly larger in all three of the other conditions (all comparisons $p < .05$). Thus, the basic tendency of subjects to use a WS/LS strategy appears to be modulated by learning, such that the acquisition of a preference for the good deck in the short condition renders subjects less susceptible to trial-to-trial fluctuations in reward amount.

Neuroimaging results

Neuroimaging studies typically proceed by identifying brain regions that are sensitive to the main effects of an experimental manipulation and subsequently drawing inferences about the role of those regions. A standard approach given the present data, therefore, would be to contrast brain activity during the short condition with brain activity during the long condition. Such analyses are important in their own right and form the basis of a separate paper on this data set (Yarkoni et al., 2005). However, they are not the optimal analyses to conduct when the goal is to understand which brain regions contribute directly to behavioral differences between individuals. Instead, regression analyses that employ an index of overt behavior (e.g., choice of the good deck in the long condition) as a variable in an equation predicting brain activity are more appropriate. Thus, most of the analyses we report below predominantly rely on regression rather than subtractive contrasts. Although desirable, genuine single-subject analyses are not possible for the neuroimaging data because estimates of brain activity with an fMRI design involving rapid events require averaging across many trials. Thus, we could not analyze patterns of neural change across single trials within individual subjects.

To determine the degree to which event-related (i.e., transient) and state-related (i.e., sustained) brain activity predicted choice behavior in the two different conditions, we regressed performance separately on state- and event-related brain activity. To reduce the probability of false positives due to multiple comparisons (since our images each contained over 65,000 voxels), we employed the common

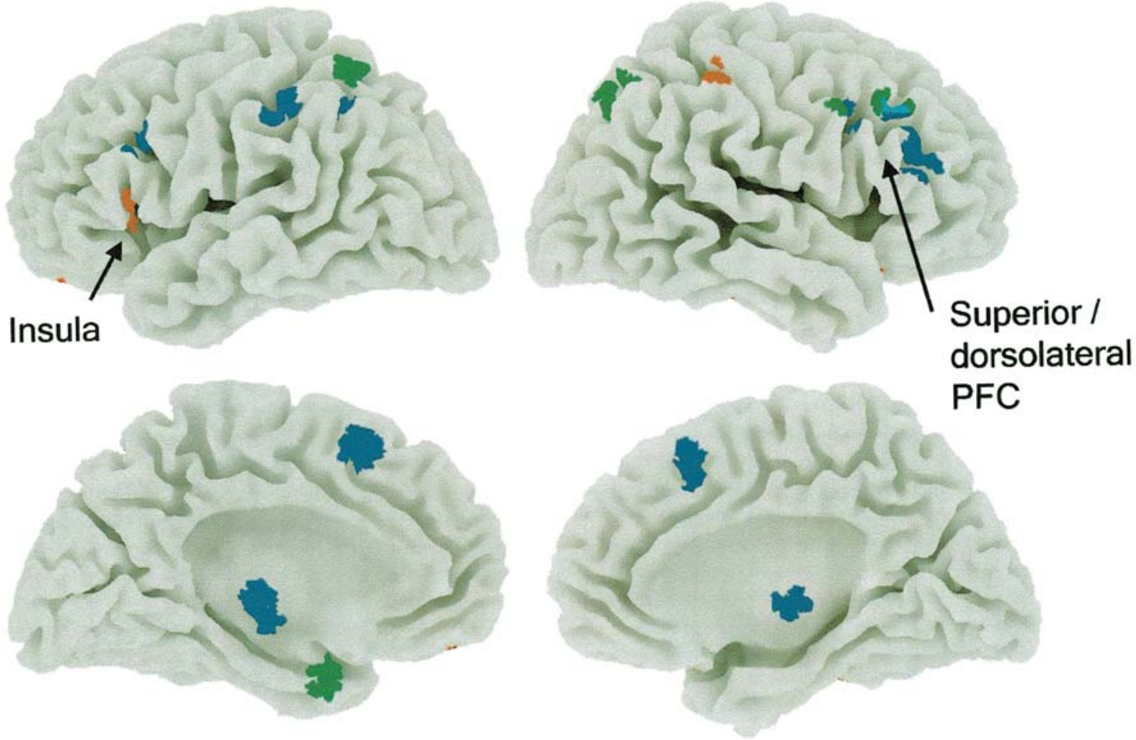

Fig. 5. Brain regions associated with either choice performance or reward sensitivity. Blue: regions in which variation in event-related neural activity in the short condition predicted individual differences in deck choice during the short condition. Green: regions in which variation in event-related neural activity in the short condition predicted individual differences in deck choice in both conditions. Red: regions that displayed sensitivity to variation in trial-to-trial changes in reward amount, aggregated across subjects. Top panels: lateral surface of the brain; bottom panels: medial surface; left panels: left hemisphere; right panels: right hemisphere. Note that all activations are mapped to a composite, standardized brain, and do not represent any single subject. For the regression analyses (blue and green), the figure shows regions in which between-subjects differences in level of activation predicted between-subject differences in behavioral choice. For the group-based analysis (red), the figure shows regions that were significant in a within-subjects ANOVA with change in reward amount as the only factor.

threshold of $p < .001$ uncorrected, and thresholded clusters for extent at 8 voxels (i.e., only clusters comprising 8 or more contiguous voxels, each significant at $p < .001$, were considered reliable).

When correlating choice of the good deck with state-related brain activity, no region attained significance in either the short or long condition. However, in the short condition, event-related brain activity in several regions was significantly correlated with choice. This set of regions included large clusters in right dorsolateral prefrontal cortex (DLPFC), left inferior parietal cortex, and medial frontal cortex (Table 2; Figure 5, blue). Good-deck choice correlated negatively with activity in all regions (i.e., those subjects who made more selections from the good deck showed less activation), and the association was strongest in the DLPFC cluster, $r = -.75$, $p < .001$.

To ensure that we did not miss any brain-behavior relationships of a more moderate effect size that would not attain significance at the threshold used above, we conducted a complementary analysis using a conjunction approach to identify regions that predicted performance in *both* conditions. That is, we sought to identify brain activation in either the short condition or the long condition that predicted choice of the good deck in both. The presence of regions that met these criteria was expected on the basis of behavioral data indicating a significant negative correlation between performance in the short and long conditions. The advantage of conducting a conjunction analysis rather than a simple correlational approach is that the probability of the same region correlating with two separate behavioral indices purely by chance is substantially lowered, thereby affording the use of more liberal thresholds for each index separately (in this case, $p < .05$, two-tailed; for similar approaches, see Braver et al., 2003; Yarkoni et al., 2005).

Thresholding for extent at 8 voxels, no pattern of state-related brain activation in either condition was significantly associated with behavioral performance in both conditions. The same held true for event-related brain activation in the long condition. In the short condition, however, event-related activation in several regions significantly predicted choice of the good deck in both the short and long conditions (Table 2; Figure 5, green). A strong effect (correlation with choice in short condition: $r = -.59$, $p < .001$; long condition: $r = .53$, $p < .001$) was observed in a cluster located in right superior PFC contiguous with (though more superior than) the one extracted in the previous analysis. Figure 6 shows scatter plots of performance in each condition regressed on event-related brain activity in this area during the short condition. When controlling for activation level in this area, the relation between behavioral performance in the short and long conditions became insignificant ($r = -.15$, $p > .10$). Statistical mediation analysis indicated this reduction in effect size to be marginally significant (Sobel test, $z = 1.85$, $p < .07$), tentatively suggesting that right superior activity partially mediated the behavioral relationship between performance in the two conditions. Note that this mediation analysis allows us to infer only that the data are *consistent* with a path model in which causality flows through the mediating variable (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002; Shrout & Bolger, 2002). Although the results are suggestive, they do not allow us to conclude with certainty that neural activity in right PFC *causally* mediated the behavioral relationship.

An interesting feature of the above results was that all of the analyses correlating brain activity with behavioral performance failed to identify brain areas commonly associated with processing of reward-related or affective information, including frontal cortex, basal ganglia, amygdala, and insular cortex (for reviews, see Montague & Berns, 2002; Phan et al., 2002; Schultz, 2000). Four possibilities might account for this observation. First, the failure to locate reward-sensitive regions could be technical in nature, as it is notoriously difficult to image accurately parts of ventromedial cortex (Ojemann et al., 1997). Second, it is possible that the experimental manipulation of reward amount was not strong enough for subjects to detect consistently, although this seems unlikely given the clear-cut behavioral results. Third, it is possible that the neural response was differentially sensitive to magnitude of gain, but that this sensitivity itself did not play a role in modulating choice performance. Finally, it may be that the regions sensitive to reward in the present task were identical to or overlapped with those that predicted choice performance.

To disambiguate these possibilities, we divided trial-to-trial changes in reward amount into 3 bins (viz., increase, decrease, and no change) and conducted an ANOVA with reward-amount change as the only factor in order to identify any brain regions sensitive to changes in reward at a threshold of $p < .001$ uncorrected. We restricted this analysis to the short condition since brain activity in the long condition was not predictive of behavior. Several clusters survived threshold, notably including a cluster in left anterior insula (Table 2; Figure 5, red). Previous fMRI studies have observed insular activation in response to both negative and positive outcomes (Elliott, Friston, & Dolan, 2000; Knutson, Fong, Adams, Varner, & Hommer, 2001). The insula also is associated with emotional responses during tasks involving cognitive demand (Phan et al., 2002), as well as with negative responses to unfair offers in an Ultimatum Game (Sanfey et al., 2003). In the present study, the insula responded more strongly to a decrease in reward payoff than to an increase or no change in reward (see Figure 7). Other regions displayed different patterns, for example, selectivity for increases in reward. Of note, there was no overlap between performance-sensitive and reward-sensitive regions (Figure 5). Activation in the reward-sensitive regions did not predict performance level, and conversely, performance-sensitive regions showed no sensitivity to changes in

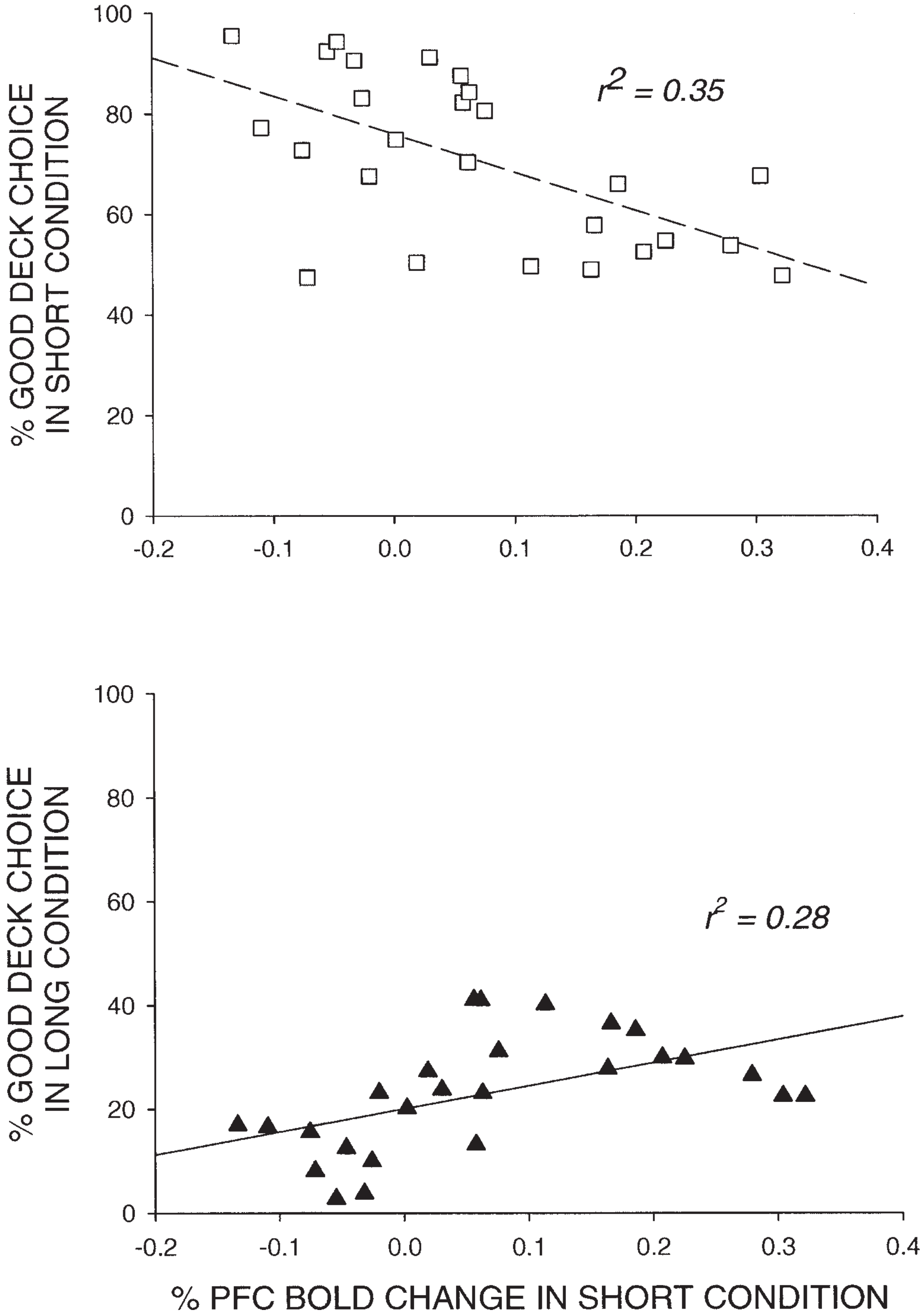

Fig. 6. Percentage choice of the good deck in the short ($W = 2$) condition (top panel) and long ($W = 12$) condition (bottom panel) as a function of right superior PFC event-related brain activity in the short condition.

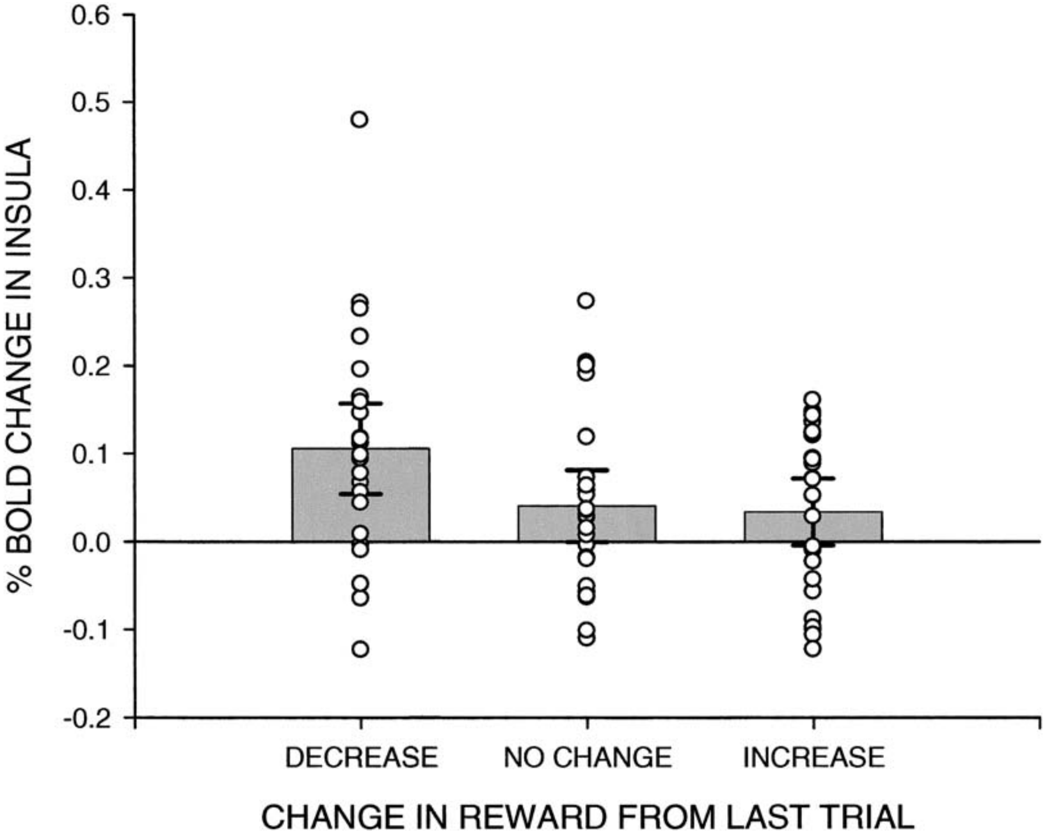

Fig. 7. Mean percentage change in BOLD signal in left anterior insula as a function of trial-to-trial changes (increase, no change, or decrease) in reward amount. Error bars represent 95% confidence intervals. Each circle represents a single subject.

reward magnitude. Thus, there appear to be dissociable neural mechanisms for detection versus integration of reward information. Whereas regions such as the insula signaled the presence of salient reward-related information on a trial-by-trial basis, the neural response in right PFC and other brain regions tracked the behavioral preferences that subjects acquired over multiple trials.

DISCUSSION

A key finding of the present study was that activation in right lateral PFC during the short-window condition predicted subjects' overall choice performance in both conditions. Previous neuroimaging studies have implicated lateral PFC activation in a variety of cognitive control processes, including inhibition of interference as well as maintenance, integration, and manipulation of information (Cabeza & Nyberg, 2000; Smith & Jonides, 1999). Indeed, the great majority of tasks involving effortful cognition have resulted in activation in this region (Duncan & Owen, 2000). Although the complexity of the decision-making task does not allow us to draw fine-grained conclusions about the cognitive processes employed by subjects, the results are consistent with the suggestion that variation in a neural network associated with higher cognitive processing drove variation in behavioral choice during the task. Subjects who performed well in the short condition and poorly in the long condition showed less event-related activation in lateral PFC, suggesting individual differences in the task-related need for higher cognitive processing. Because a defining feature of the decision-making task is a tension between immediate and delayed rewards, one possibility is that these individual differences reflect variation in subjects' impulse control, namely, the ability to inhibit choice of the bad deck, which is associated with an immediate gain, so as to receive the larger, long-term gains associated with choice of the good deck.

An unexpected finding was the negative correlation between choice performance in the short and long conditions. It would seem reasonable to suppose that subjects who show greater impulse control in the short condition also would show greater control in the long condition. This was not the case, however. One possibility that can account for the failure to find such a pattern is that individual differences in choice reflected degree of engagement in the task rather than an ability to control one's impulsive choices. Perhaps some subjects were actively engaged in the task whereas others chose to respond more randomly rather than make a genuine effort. The former group might be expected to succeed in the short condition (i.e., choose from the good deck), which has a relatively tractable solution, but to perform more poorly in the long condition, in which the most salient aspect is the immediate reward increase associated with choice of the bad deck. In contrast, subjects who respond relatively unsystematically likely would bring their performance closer to chance levels in both conditions, resulting in relatively poorer performance in the short condition but relatively better performance in the long condition. However, the fact that the majority of subjects were sensitive to the experimental manipulation in the long condition, performing below chance, argues against such an explanation. Nine subjects performed at chance level in the short condition but only 3 performed at chance level in the long condition. Thus, random deck selection by some subjects does not appear to be a plausible explanation for the negative correlation in choice responding between conditions.

Consistent with the negative correlation between choice in the two conditions, event-related activity in right PFC during the short condition correlated negatively with good deck choice in the short condition but correlated positively with good deck choice in the long condition. Although it may seem counterintuitive for the same brain region to predict opposite performance in the two conditions, it is important to realize that the event-related neural response is not a measure of *total* task-related neural activity but rather is a measure specifically of transient activation linked to individual decision-making trials (as opposed to entire task blocks). A larger event-related neural response in PFC most likely reflects a greater need for controlled processing during individual trials in subjects who did not develop a strong preference for one of the decks. In contrast, subjects who repeatedly chose from one deck may have required less controlled processing during trials since their selection process was relatively automated (i.e., almost always selecting from the same deck). In the short condition, in which most subjects rapidly acquired a preference for the good deck, such automation would lead to a negative correlation between PFC activity and good-deck choice, since those subjects who failed to acquire a preference (and consequently performed more poorly) would persist in exercising greater trial-related control processes. Conversely, in the long condition, in which most subjects rapidly developed an automated preference for the bad deck, a failure to develop a preference would lead to a positive correlation between PFC activity and good deck choice, since those subjects who persisted in exercising trial-related decision-making processes were likely to perform closer to chance (i.e., better).

Note that while this account explains the general role of event-related PFC activity in relation to the decision-making task, it sheds little light on the nature of the processes that led subjects to acquire deck preferences in the first place. Because most subjects developed preferences rapidly in both conditions (typically within 10–20 trials), the rapid event-related fMRI design we used did not allow for sufficiently fine-grained analyses to identify the neural mechanisms that drove preference acquisition. Based on the behavioral data indicating that, overall, subjects were more likely to switch decks more often following a decrease in reward than following an increase, one possibility is that at least some of the subjects were using a relatively simple win–stay/lose–shift (WS/LS) strategy. Alternatively, subjects may have used a more deliberative process. In a previous report on this data set we argued that a WS/LS account could not fully explain other aspects of the neuroimaging results (not reported here), and we provided evidence suggesting that the PFC may support the integration of information (i.e., trial outcomes) over time (Yarkoni et al., 2005). A more detailed understanding of the preference-acquisition process awaits further behavioral manipulations or alternative fMRI designs, but the interplay between behavioral and fMRI results in this experiment illustrates the mutually constraining roles that these different levels of analysis can play.

An important caveat to the interpretation of the event-related right PFC activity as an index of controlled processing is that state-related PFC activation did not significantly predict behavior in either the short or the long condition. The absence of a relation between sustained neural activity and task behavior may seem counterintuitive since one might expect that differences in strategy across subjects would be more likely to show up in sustained rather than transient patterns of brain activation. For example, a strategy of repeatedly choosing from the same deck should involve little cognitive effort on individual trials, but might require that the strategy be kept actively in mind throughout the task, thereby influencing sustained neural activity. One possible explanation for the failure to find a relation between sustained neural activity and choice is a lack of statistical power. Because the magnitude and variance of the state-related hemodynamic response in the short condition was much larger, on average, than the corresponding event-related response (state activity mean = 0.16% BOLD increase, SD = 0.34%, vs. event-related activity mean = 0.07% BOLD increase, SD = .13%), it may be that the strategic differences in task performance that were observed in the event-related response were obscured by other sources of variance in the state-related response. Regardless of the final explanation, this null result serves as a reminder that neuroimaging data are highly complex and are to be interpreted cautiously.

Results from functional neuroimaging may offer insight into the neural systems that underlie differences in behavioral performance. Given that the present task involves temporally extended decision-making involving explicit rewards, we postulated that two neural systems might be involved: (a) a higher cognitive system capable of integrating response-outcome relations in order to detect temporal contingencies and trends (e.g., amount of reward tends to increase following repeated selections of the good deck), and (b) an affective/reward system sensitive to fluctuations in trial-to-trial reward. From a purely behavioral standpoint, it is difficult to determine which mechanism predominates under the different conditions in the present task. The fMRI data, however, offer some insight in this regard. Whereas lateral PFC, a region closely tied to higher cognitive processing, predicted subjects' level of choice performance, no area associated with reward processing showed a similar relation. This was not due to a null result in the latter case since the neural response in the insula, a region associated with emotion and reward processing, was sensitive to changes in reward amount. Rather, it appears that choice depended primarily on variability in the degree to which subjects engaged higher cognitive processes, and not on the degree to which subjects manifested an affective response to trial-to-trial changes in reward. This result illustrates the ability of fMRI to dissociate the putative mechanisms underlying a behavioral task and offers evidence that neuroimaging can serve as an important complement to behavioral analyses.

A general principle exemplified in the present data is that activity in relatively circumscribed areas of the brain may accurately predict individual differences in complex behaviors. For example, the maximal correlation between right superior PFC activation and choice performance in the short condition was $r = -.75$. The magnitude of this correlation may surprise those unfamiliar with recent imaging studies that find similarly strong correlations between isolated brain activity and overt judgment or behavior across a wide variety of contexts (e.g., Canli et al., 2001; Lieberman et al., 2004). One might expect that successful performance on complex tasks would require the involvement of many parts of the brain acting in concert, thereby minimizing the likelihood of finding strong correlations with individual regions. Some critics have formalized this notion, going so far as to argue on principled grounds that higher cognition cannot be localized to circumscribed brain regions (e.g., Uttal, 2001).

Our use of cluster-size thresholding and complementary conjunction analyses on regions identified at lower thresholds reduced the probability of false positives in the present study. In addition, we showed that controlling for right DLPFC activation rendered the correlation between choice performance in the short and long conditions insignificant. We would suggest, therefore, that given the magnitude of the correlations obtained, it is not unreasonable to entertain the possibility that complex cognitive functions are localized to specific regions. Of course, the fact that most fMRI studies focus on a small number of brain regions for analysis does not mean that other parts of the brain do not play an important role in the behavior. In our study, the brain–behavior correlation was strongest in right lateral PFC, but activation in many other brain regions also predicted performance (Figure 5, blue and green). Moreover, lowering the threshold of significance would have led to the identification of still other brain areas that may be veridically involved in the task. Likewise, in another report in which we conducted standard group analyses on this data set, we observed that task performance produced changes in the activity of a widespread network of brain regions (Yarkoni et al., 2005). Thus, circumscribed brain activity can be predictive of behavior without implying that activity in a single region constitutes a necessary or sufficient condition for behavior.

Overall, our results demonstrate the utility of an approach that jointly employs functional neuroimaging and traditional behavioral analysis. The present study, although limited in several respects, represents one of the first uses of fMRI to investigate the relationship between neural activity and complex, overt decision-making behavior. Future studies could build on our results in several ways. One potential line of research would be to employ the same basic paradigm we used but modify aspects of the design in order to address some of the limitations of the current study. For example, the inability to estimate brain activity locked to individual trials could be ameliorated by using slower event-related fMRI designs with greater spacing between trials. Similarly, titrating task parameters such as window size and reward magnitude in order to slow down preference acquisition could increase the power to detect neural activity related to learning itself as distinct from its subsequent effects on choice.

Alternatively, the approach we used could be extended to other areas of research involving higher-order decision-making processes. For example, the hypothesis that PFC is critically involved in self control and impulsivity, and underlies choice in tasks involving long- versus short-term reward tradeoffs, predicts that individuals with poorer impulse control should show both poorer choice patterns and reduced lateral PFC activity on tasks similar to the present one. Although neuroimaging studies of this type have, to our knowledge, not yet been conducted, such a prediction is consistent with prior behavioral work (e.g., Hinson, Jameson, & Whitney, 2003).

REFERENCES

Ainslie, G. (2001). *Breakdown of will.* Cambridge, England: Cambridge University Press.

Berns, G. S., McClure, S. M., Pagnoni, G., & Montague, P. R. (2001). Predictability modulates human brain response to reward. *Journal of Neuroscience, 21,* 2793–2798.

Berridge, K. C. (2003). Pleasures of the brain. *Brain & Cognition, 52,* 106–128.

Braver, T. S., Reynolds, J. R., & Donaldson, D. I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron, 39,* 713–726.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences, 4,* 215–222.

Cabeza, R., & Kingstone, A. (Eds.) (2001), *Handbook of functional neuroimaging of cognition.* Cambridge, MA: MIT Press.

Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience, 12,* 1–47.

Canli, T., Zhao, Z., Desmond, J. E., Kang, E., Gross, J., & Gabrieli, J. D. E. (2001). An fMRI study of personality influences on brain reactivity to emotional stimuli. *Behavioral Neuroscience, 115,* 33–42.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25,* 257–271.

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

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences, 23,* 475–483.

Elliott, R., Friston, K. J., & Dolan, R. J. (2000). Dissociable neural responses in human reward systems. *Journal of Neuroscience, 20,* 6159–6165.

Elliott, R., Newman, J. L., Longe, O. A., & Deakin, J. F. (2003). Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: A parametric functional magnetic resonance imaging study. *Journal of Neuroscience, 23*, 303–307.

Gray, J. R. (1999). A bias toward short-term thinking in threat-related negative emotional states. *Personality and Social Psychology Bulletin, 25*, 65–75.

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience, 6*, 316–322.

Green, L., & Myerson, J. (1993). Alternative frameworks for the analysis of self control. *Behavior and Philosophy, 21*, 37–47.

Herrnstein, R. J., Loewenstein, G. E., Prelec, D., & Vaughan, W., Jr. (1993). Utility maximization and melioration: Internalities in individual choice. *Journal of Behavioral Decision Making, 6*, 149–185.

Hinson, J. M., Jameson, T. L., & Whitney, P. (2003). Impulsive decision making and working memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 29*, 298–306.

Huettel, S. A., Song, A. W., & McCarthy, G. (2004). *Functional magnetic resonance imaging*. Sunderland, MA: Sinauer Associates.

Knutson, B., Fong, G. W., Adams, C. W., Varner, J. L., & Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport, 12*, 3683–3687.

Krawczyk, D. C. (2002). Contributions of the prefrontal cortex to the neural basis of human decision making. *Neuroscience & Biobehavioral Reviews, 26*, 631–664.

Kudadjie-Gyamfi, E., & Rachlin, R. (1996). Temporal patterning in choice among delayed outcomes. *Organizational Behavior and Human Decision Processes, 65*, 61–67.

Lieberman, M. D., Jarcho, J. M., Berman, S., Naliboff, B., Suyenobu, B. Y., Mandelkern, M., & Mayer, E. (2004). The neural correlates of placebo effects: A disruption account. *NeuroImage, 22*, 447–455.

Logue, A. W. (1988). Research on self-control: An integrating framework. *Behavioral and Brain Sciences, 11*, 665–679.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7*, 83–104.

Montague, P. R., & Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron, 36*, 265–284.

O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience, 4*, 95–102.

Ojemann, J. G., Akbudak, E., Snyder, A. Z., McKinstry, R. C., Raichle, M. E., & Conturo, T. E. (1997). Anatomic localization and quantitative analysis of gradient refocused echo-planar fMRI susceptibility artifacts. *NeuroImage, 6*, 156–167.

Pagnoni, G., Zink, C. F., Montague, P. R., & Berns, G. S. (2002). Activity in human ventral striatum locked to errors of reward prediction. *Nature Neuroscience, 5*, 97–98.

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage, 16*, 331–348.

Rachlin, H., & Green, L. (1972). Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior, 17*, 15–22.

Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S., & Kiets, C. D. (2002). A neural basis for social cooperation. *Neuron, 35*, 395–405.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (13 June, 2003). The neural basis of economic decision making in the ultimatum game. *Science, 300*, 1755–1757.

Schultz, W. (2000). Multiple reward systems in the brain. *Nature Reviews Neuroscience, 1*, 199–207.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron, 36*, 241–263.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods, 7*, 422–445.

Smith, E. E., & Jonides, J. (12 March, 1999). Storage and executive processes in the frontal lobes, *Science, 283*, 1657–1661.

Uttal, W. (2001). *The new phrenology: The limits of localizing cognitive functions in the brain.* Cambridge, MA: MIT Press.

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

Wilkinson, D. T., & Halligan, P. W. (2004). The relevance of behavioural measures for functional imaging studies of cognition. *Nature Reviews Neuroscience, 5*, 67–73.

Yarkoni, T., Gray, J. R., Chrastil, E. R., Barch, D. M., Green, L., & Braver, T. S. (2005). Sustained neural activity associated with cognitive control during temporally extended decision making. *Cognitive Brain Research, 23*, 71–84.

Received: November 7, 2004
Final acceptance: September 16, 2005