Neuron, Vol. 44, 365–378, October 14, 2004, Copyright ©2004 by Cell Press

Activity in Posterior Parietal Cortex Is Correlated with the Relative Subjective Desirability of Action

Michael C. Dorris[1,2] and Paul W. Glimcher[1,*]
[1]Center for Neural Science
New York University
New York, New York 10003

Summary

Behavioral studies suggest that making a decision involves representing the overall desirability of all available actions and then selecting that action that is most desirable. Physiological studies have proposed that neurons in the parietal cortex play a role in selecting movements for execution. To test the hypothesis that these parietal neurons encode the subjective desirability of making particular movements, we exploited Nash's game theoretic equilibrium, during which the subjective desirability of multiple actions should be equal for human players. Behavior measured during a strategic game suggests that monkeys' choices, like those of humans, are guided by subjective desirability. Under these conditions, activity in the parietal cortex was correlated with the relative subjective desirability of actions irrespective of the specific combination of reward magnitude, reward probability, and response probability associated with each action. These observations may help place many recent findings regarding the posterior parietal cortex into a common conceptual framework.

Introduction

Physiological studies conducted during different types of visual-saccadic decision making indicate a correlation between neural activity in the lateral intraparietal area (area LIP) and several variables related to an animal's decision about what eye movement to produce. Among others, these variables include the intention to make a saccade (Gnadt and Andersen, 1988), the log likelihood that a given eye movement will result in a reward (Gold and Shadlen, 2000, 2001), the integral of perceptual signals indicating which saccade will be rewarded (Shadlen and Newsome, 1996, 2001), the probability and magnitude of reward associated with a saccade (Platt and Glimcher, 1999), the average rate at which choice of an option has been rewarded in the recent past (Sugrue et al., 2004), the instantaneous likelihood, or hazard, that a reinforced saccade will be instructed (P. Janssen and M.N. Shadlen, 2003, Representation of the hazard function of elapsed time by neurons in macaque area LIP, Soc. Neurosci., abstract), or some combination of these variables (Coe et al., 2002). Each of these variables, in turn, may influence saccade production via the connection of area LIP with other saccade-related areas (Andersen et al., 1985, 1990; Paré and Wurtz, 2001). By manipulating the probability of reward, the magnitude of reward, or the recent history of rewards, these previous studies were similar in that they altered the overall desirability of producing a given saccade.

In related work, behavioral and theoretical studies have begun to describe the algorithms by which rational human, animal, and artificial choosers actually compute the desirability of actions. Within the machine learning community, choosers use algorithms that estimate the value of each available course of action and then select for execution the course of action estimated to have the highest average value (cf. Sutton and Barto, 1998). Within economic circles, expected utility theory has provided a useful, if imperfect, estimate of the desirability of actions for human and animal choosers (Camerer, 2003). Expected utility is calculated as a function of the probability, magnitude, and delay to a reward (cf. Kreps, 1990). These studies indicate that decision making can, in principle, be modeled as a process in which many variables influence a common representation of desirability, and it is this common representation upon which the choice mechanism is hypothesized to operate when it selects an action for execution.

If decisions about what saccade to produce are generated in part by the neurons in area LIP, as a number of physiologists have proposed, and decision making involves a representation of the overall desirability of each available course of action, as decision scientists propose, then it would be logical to hypothesize that the subjective desirability of potential saccades may be represented in area LIP. A strategy for examining this hypothesis further is to systematically disassociate the subjective desirability of a saccade from other decision variables associated with a saccade. To accomplish this, we borrowed a behavioral task from game theory, a branch of economics developed to describe the complex behavior that arises during strategic interactions among agents whose choices dynamically influence the desirability of each other's actions. As stated by von Neumann and Morgenstern (1944) when they developed the first tools to tackle these problems: "If two or more persons exchange goods with each other, then the results for each one will depend in general not merely upon his own actions but on those of the others as well. Thus each participant attempts to maximize a function…of which he does not control all the variables. This is certainly no [classical] maximization problem, but a peculiar and disconcerting mixture of several conflicting maximization problems."

The task we used was the *inspection game* (Kreps, 1990), and it has two features that are important for our purposes. First, there is no single correct action. Free to choose, rational subjects adopt a mixed strategy in which they devote a certain proportion of responses to each action (Fudenberg and Tirole, 1994). Second, Nash (1950, 1951) defined these mixed strategy equilbria as situations in which the subjective desirability of each available action is, on average, equivalent. Essentially, he proposed that if one of the available actions was

*Correspondence: glimcher@cns.nyu.edu
[2]Present address: Department of Physiology, Queen's University, Kingston, Ontario, Canada (dorrism@biomed.queensu.ca).

perceived as more desirable than any other, then rational choosers would always select that action. Mixed strategies would only arise when the two or more actions being mixed were of precisely equal desirability. The critical concept underlying our experiments is that multiple decision-related variables trade off against each other with the end result being that the subjective desirability of each available action remains equal whenever mixed strategy equilibrium of the type Nash predicted is produced. Here we use these features of mixed strategy equilibria to test whether area LIP represents the subjective desirability of available actions even when individual decision variables or movement probabilities vary.

More specifically, we hypothesize that LIP activity is correlated with the *relative* subjective desirability of saccades, that is, the subjective estimate of desirability associated with the saccade in the neuron's response field divided by the sum of the subjective desirability associated with all available saccades (cf. Herrnstein, 1961). Although Nash's own work was initially based on calculations of *absolute* subjective desirability, subsequent behavioral experiments have suggested that humans employ a measure of relative desirability during most decision making (Kahneman and Tversky, 1979). Moreover, our current understanding of LIP (Platt and Glimcher, 1999; Gold and Shadlen, 2001) and of the general mechanisms of cortical processing (Heeger et al., 1996) argue that if LIP does contribute to decision making it should do so in a relativistic fashion. If this hypothesis could be further tested, then we might be able to conclude that the equilibrium computations that real organisms perform are computed in relative terms.

Although we did not measure the precise subjective desirability associated with the actions available to our monkeys, like other decision scientists, we assume that their patterns of choices must reflect these desirabilities. These presumptions are supported by our measurements of monkey behavior during the inspection game. Working from this starting point, we systematically varied the relative subjective desirability of two saccades during nonstrategic blocks of trials in which the responses of the monkeys were held constant. Under these conditions, LIP firing rates varied in a manner similar to relative subjective desirability. Next we varied the responses of the monkeys by manipulations made during a strategic game, a condition during which relative subjective desirability should remain paradoxically constant while behavior varies. Under these conditions, LIP firing rates were largely constant, fluctuating only weakly and in a manner that would be consistent with mechanistic predictions about how game theoretic equilibriums are maintained (Camerer, 2003). Finally, we explicitly tested the notion that LIP represents the subjective desirability of actions in a relative and not an absolute manner. Under these conditions, we found that LIP firing rates were more tightly correlated with relative estimates of subjective desirability than with absolute estimates of subjective desirability. Taken together, these data suggest that neurons in the posterior parietal cortex may encode response desirability. LIP firing rates are influenced by many decision variables specifically because they provide a general purpose encoding of movement desirabilities of the type that would be expected in a response selection system.

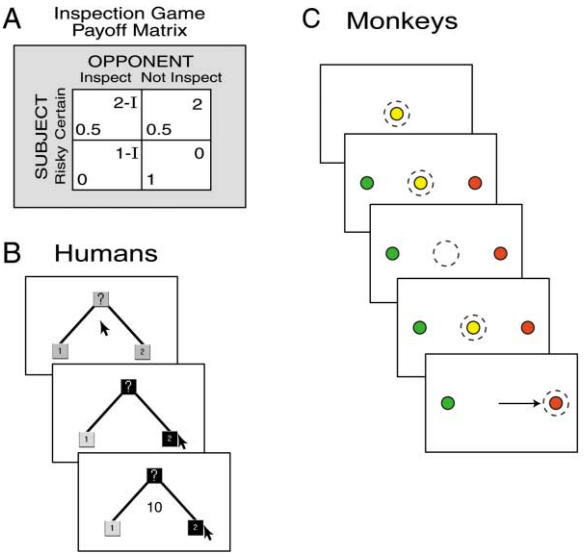

Figure 1. The Mixed Strategy Inspection Game

(A) General form of the payoff matrix. The variables in the bottom left of each cell determine the subject's payoffs, and the variables in the top right of each cell determine the opponent's payoffs for each combination of player's responses. I = cost of inspection to the opponent = 0.1–0.9 in 0.2 steps. One unit of payoff = 0.25 mL of water for monkey = $0.05 for human. See Experimental Procedures for more details.

(B) Manual inspection game played by human subjects. Humans were free to choose either the certain (button 1) or risky (button 2) option by clicking the appropriate button with a pointer controlled by a computer mouse (black arrow). At the end of each trial, their payoff in cents was presented in the center of the display.

(C) Oculomotor inspection game played by monkey subjects. Monkey subjects were free to choose either the certain target (green) or the risky target (red) when the central fixation point was extinguished for the second time. The dashed circles indicate the current direction of gaze. Instructed trials were identical except the fixation point reappeared as either green or red, indicating which target would be rewarded at the end of the trial (not shown).

Results

Behavior

Our first goal was to compare the behavior of humans and monkeys during a strategic game to determine whether monkeys showed the same mixed strategy behavior that the Nash equilibrium predicts will occur in human subjects (cf. Nash, 1950, 1951; Kreps, 1990; Fudenberg and Tirole, 1994). We chose the inspection game because the proportion of responses that subjects devote to each option can be experimentally manipulated while the payoffs associated with each option remain constant (Figure 1A). On each trial, subjects chose either the *certain option*, whose payoff was always one unit of reward, or the *risky option*, whose payoff was either two or zero units of reward (see Experimental Procedures for details). The subject received the risky reward only if the other player (the opponent) selected the "no inspect" option; otherwise, the subject received nothing. The likelihood that the opponent would choose the no inspect option was influenced, in large part, by an independent variable affecting the opponent's payoffs—the *cost of inspection* (Figure 1A; variable I). In-

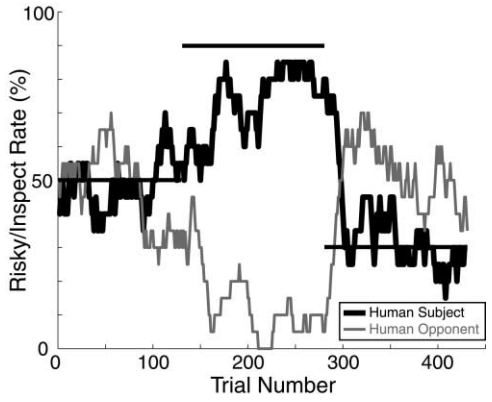

Figure 2. Human versus Human Choice Behavior during Three Blocks of Inspection Game Trials

The thick black line denotes the 20 trial running average of the percentage of the risky option chosen by the subject. The horizontal black lines denote the subject's rate of choosing the risky option predicted at the Nash equilibrium. The thin gray line denotes the corresponding 20 trial running average of the percentage of the inspect option chosen by the opponent. The opponent's predicted rate of choosing the inspect option at Nash equilibrium was 50% for all blocks of trials. The opponent's costs of inspection were stepped sequentially from 0.5 to 0.9 to 0.3 across the three blocks of trials; 150 trials/block.

creasing this variable encouraged the opponent to more aggressively prevent exploitation of the risky option by the subject. In contrast, when the cost of inspection variable was lowered, the opponent was less aggressive in response to exploitation of the risky option by the subject. *The Nash formulation for our game thus predicts that a subjects' response probability will be adjusted by the competitive interaction until the opponent counters the risky choice 50% of the time, regardless of the equilibrium level reached by the subjects*. Critical for the logic of this paper is that a mixed strategy equilibrium is a behavioral measure of an otherwise hidden parameter—equivalence of subjective desirability across available options.

Humans versus Humans

During the first experiment, human subjects competed against human opponents in a manual version of the inspection game (Figure 1B). Figure 2 shows a running average over the last 20 trials of the strategic choice behaviors of a human subject (black) and human opponent (gray). Across the three blocks of trials, the opponent's cost of inspection was stepped from 0.5 to 0.9 to 0.3, respectively. The percentage of risky choices predicted at the Nash equilibrium is represented by the horizontal black lines (see Experimental Procedures, Equations 4–6). The opponent was predicted to counter risky choices by choosing the inspect option 50% of the time across all blocks of trials (see Experimental Procedures, Equations 1–3). During the first and third blocks, the subject's behavior closely tracked the Nash equilibrium predictions as did the opponent's behavior, which approached a 50% inspect rate. Of course, the players' rates of responding are not fixed to those predicted by the Nash equations. This is exemplified in the second block of trials in which the subject failed to reach the predicted 90% rate of choosing the risky option. This had the effect of lowering the desirability of the inspect option for the opponent and, therefore, reduced his inspect rate.

To quantify the influence of changing the opponents' cost of inspection variable on the subjects' rates of choosing the risky option, we averaged the choice behavior over the last half of each block once subjects had presumably reached a stable equilibrium state. We found that the human subjects ($n = 5$) did indeed adopt mixed strategies and that the equilibrium rates of choosing the risky option were influenced by changes in the opponents' payoffs (Figure 3A, black circles). Choosing mixed strategies of a certain proportion of each response is consistent with the notion that the average subjective desirability of both available actions were equal for both players under these conditions (Nash, 1950).

Humans versus Computer

Having quantified behavior when humans competed against human opponents, we then employed a standardized computer "opponent" for the remainder of the experiments. In brief, the computer opponent entered the subject's history of choices into a reinforcement learning algorithm to estimate the probability that the subject would choose each of the two available options on the upcoming trial. Using this estimate of the subject's choice probability, the computer then chose either to *inspect* or *not inspect* so as to maximize its own virtual reward (see Experimental Procedures for details). Our goal was not to develop a computer opponent that perfectly mimicked the behavior of human opponents. Instead our goal was to develop a computer opponent that, like human opponents, elicited mixed strategy equilibrium behavior in subjects with associated response probabilities that were dependent on a free variable that affected the computer opponent's payoffs (Figure 1A; variable I, cost of inspection). Standardizing the opponent in this manner provided consistency from one experimental session to another and allowed the behavioral strategies to be compared across species.

Human subjects displayed mixed strategy behavior when competing against this computer algorithm, and manipulations of the algorithm payoffs with the cost of inspection variable changed the proportion of choices devoted to the risky option across blocks of trials (Figure 3A, red triangles) (ANOVA, d.f. $= 4$, $F = 177$, $p < 0.0001$, $n = 8$ subjects). These results are consistent with subjective desirability of both available actions being on average equal when subjects were at behavioral equilibrium with the computer algorithm (Nash, 1950).

Monkeys versus Computer

We then trained monkeys to play an oculomotor version of the inspection game and assessed whether their behavior was comparable to that of human subjects when competing against the same computer algorithm. In the monkey experiments, thirsty animals competed for water rewards rather than for money, and they indicated their choices on each trial with a saccadic eye movement directed to one of two eccentric visual targets rather

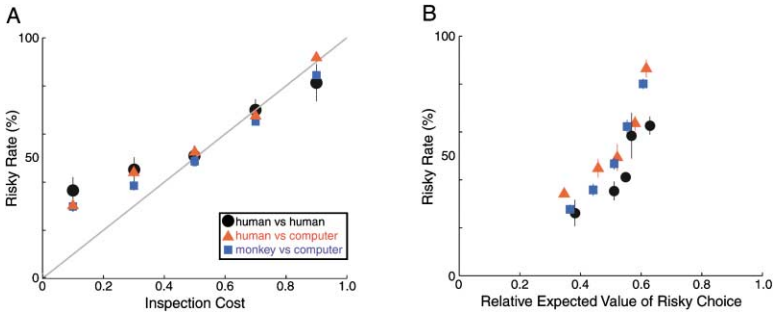

Figure 3. Influence of Changing the Opponent's Cost of Inspection Variable on the Subject's Response Rate and the Expected Value of Choices

(A) Average choice behavior (±SEM) of human subjects competing against human opponents (black) and human (red) and monkey (blue) subjects competing against the computer algorithm during inspection game trials. Human versus human data: n = 5, 150 trials/block. Human versus computer data: n = 8 subjects, 150 trials/block. Monkey data: n = 29 blocks of trials (13 blocks from monkey 1, 16 blocks from monkey 2), 100 ± 14.3 SD trials/block. The diagonal gray line denotes the predicted rates of responding at the Nash equilibrium.

(B) Relationship between expected value and choice behavior. The relative expected value of the risky choice plotted on the abscissa was calculated as the product of the probability of receiving the reward for choosing the risky option and the magnitude of reward associated with the risky option divided by the sum of the expected values of both the risky and certain options. Each datum point is calculated from blocks of trials with a different cost of inspection. Note: Same data sets as Figure 3A except n = 3 for human versus human data. For the first two experiments, the opponent's choices were not saved and therefore the subject's probability of receiving a reward could not be calculated.

than with a button press (Figure 1C; see Experimental Procedures for details). Animals were taught to perform two types of trials. During *inspection game trials*, the fixation point reappeared as yellow and subjects were free to choose either the risky red target or the certain green target. During *instructed trials*, if the fixation point reappeared as red, then they received reward only if they chose the red target, and if the fixation point reappeared as green, they received reward only if they chose the green target.

Figure 3A plots the average choice behavior of two monkey subjects as the cost of inspection variable was varied across blocks of inspection game trials (blue squares). The choice behavior of the monkeys varied with the cost of inspection (ANOVA, d.f. = 4, $F = 1095$, $p < 0.0001$, n = 29 blocks of trials from two monkeys). Therefore, monkeys displayed mixed strategy behavior during the inspection game, suggesting that the subjective desirability of both options were equal on average when at behavioral equilibrium.

Comparison of Monkey and Human Choice Behavior during Inspection Game

The choice behavior of humans and monkeys did not differ when competing against the same computer algorithm (Figure 3A; blue versus red; ANOVA, d.f. = 1, $F = 5.27$, $p > 0.01$). The choice behavior of both species varied in a similar way from that predicted by a simple form of the Nash equations (diagonal line in Figure 3A) in which subjective and objective desirability were presumed to be equivalent (see Experimental Procedures, Equations 4–6). Specifically, monkeys chose the certain option more often than predicted by this set of equations in blocks for which the cost of inspection variable was 0.7 and 0.9, and both species chose the risky option more often than predicted when this variable was 0.1 and 0.3 (Student-Newman-Keuls test, $p < 0.01$). Although, some of the observed deviations from the Nash predictions were doubtless due to imperfections in the computer algorithm (see Barraclough et al., 2004), it should be noted that humans playing this opponent did not differ significantly from humans playing other humans (ANOVA, d.f. = 1, $F = 0.22$, $p > 0.05$). More likely, much of the deviation resulted from errors in the quantitative Nash prediction that arise from our assumptions about the equivalence of subjective and objective desirability, the influence of block transitions, and the well established observation that empirical behavior often differs from normative behavior during strategic interactions (Camerer, 2003). In this case, deviations from the simple Nash predictions during the lowest cost of inspection blocks are likely the result of asymmetries in information regarding the opponent's behavior inherent in the inspection game. Subjects can only update their estimate of the probability that the opponent will inspect from payoffs associated with choosing the risky option. When subjects choose the certain option they gain no information about the behavior of their opponent. Thus, subjects may choose the risky option more often than predicted during blocks with low costs of inspection because there may be some advantage gained by obtaining a more accurate estimate of the opponent's behavior. This classic trade off between exploiting a resource and exploring the possibility of better resources elsewhere has been described previously (Sutton and Barto, 1998).

We were interested, however, in quantifying how much subjective desirability differed from objective desirability. Starting from the assumption that the subjective desirabilities of the two actions are equal at equilibrium, we then sought to compare this to an objectively measured desirability of each action. We computed the objective value of each action, a quantity known as *expected value*, across blocks of trials (Figure 3B). Expected value is defined as the product of the probability of receiving a reward and the magnitude of that reward. The probability of receiving a reward was always 100% for trials on which the subject selected the certain option but varied across blocks with the opponent's cost of inspection variable for risky trials (Table 1). Specifically, we calculated the relative expected value of the risky option as the expected value of the risky option divided by the sum of the expected values of both the risky and certain options. Across blocks of inspection game trials, the objective value of the risky option varied ($p < 0.01$) when humans competed against human opponents

Table 1. Reward Probability for Risky Responses across Different Cost of Inspection Blocks

Cost of inspection	0.1	0.3	0.5	0.7	0.9
Human versus human	31.85 ± 7.1%	52.6% ± 3.1%	61.0% ± 0.7%	67.9% ± 12.3%	85.1% ± 4.9%
Human versus computer	26.6% ± 1.3%	43.5% ± 3.9%	56.4% ± 5.6%	69.5% ± 3.0%	81.3% ± 3.7%
Monkey versus computer	30.2% ± 2.0%	41.8% ± 2.5%	54.0% ± 2.5%	63.9% ± 2.7%	78.2% ± 2.2%

Reward probability ± SEM.

(black) and when both humans (red) and monkeys (blue) competed against the computer algorithm ($p < 0.01$). This raises the question of whether neuronal activity in posterior parietal cortex tracked this objective measure of desirability—expected value—which varied across blocks of trials or subjective desirability, which is presumed to remain constant across blocks of trials.

Each data point in Figure 4A represents the average of the ratio of choices devoted to each option as a function of the product of the ratio of rewarded trials and the ratio of reward magnitudes. This is the classic matching law plot of Herrnstein (1997). Note that under these conditions behavior approached, but did not obtain, the idealized matching law solution when humans competed against human opponents (black circles) and when both humans (red triangles) and monkeys (blue squares) competed against the computer opponent. In aggregate, the subjects "overmatch" slightly; the slope of the least square regression was 1.32 (black line), which is greater than a slope of 1 predicted by the matching law (gray line).

The observation that there was overmatching in the aggregate behavioral strategy, however, should not be read to suggest that the subjects necessarily used a stationary matching-type strategy during this dynamic conflict. Unlike during traditional matching-type psychological tasks, during mixed strategy games the opponent's behavior is dynamic and responds actively to the subject's choices. Figure 4B demonstrates this property during inspection game trials with the cost of inspection variable set at 0.5. In response to the opponent's transient shift to a high inspect rate, the subject decreased his rate of choosing the risky option (first arrow). After a brief lag, the opponent reacted to this change by decreasing it's inspect rate, which was then followed by the subject again choosing the risky option more frequently (second arrow). This example emphasizes the strategic nature of game theoretic interactions. It can be true that, in aggregate, behavior during these games appears similar to behavior in nonstrategic environments, but the underlying mechanisms that generate this aggregate behavior on a trial-by-trial basis are critically dynamic in a game theoretic setting.

To quantify the dynamic nature of games further, we tested for unpredictability in subjects' choice-by-choice behavior during equilibrium play by performing a Markov chain analysis on the choice behavior on blocks with a 0.5 cost of inspection. These were blocks in which subjects chose the risky option approximately 50% of the time. This analysis determined, for a given rate of choosing an option, whether subjects repeated or alternated their choices more often than would be expected by a random process that did not depend on previous choices. The strategic nature of this interaction makes a perfectly random selection of action an optimal response in these blocks when a player faces an efficient opponent. When competing against the computer algorithm, three out of eight humans showed behavior that was significantly different from what would be predicted by a perfectly stochastic process when compared to the previous trial (χ^2 test, $p < 0.01$). Two of these three subjects as well as one further subject showed behavior that was significantly different from a purely stochastic strategy when compared to the previous two trials. The behavior of the two monkeys was slightly more stochastic than that of humans; behavior in 6 out of 47 blocks of trials were significantly different from pure stochasticity when compared to the previous trial, and behavior in 8 out of 47 blocks was significantly different from stochastic when compared to the previous two trials. In summary, the human and monkey behavior was imperfect with regard to unpredictability. This might, of course, reflect a limitation intrinsic to our players or might reflect an inability of the computer opponent we designed to sufficiently exploit nonrandom patterns in the behavior of our players (Glimcher, 2005). Existing evidence from other monkey studies, however, does suggest that had our computer opponent been better able to identify pat-

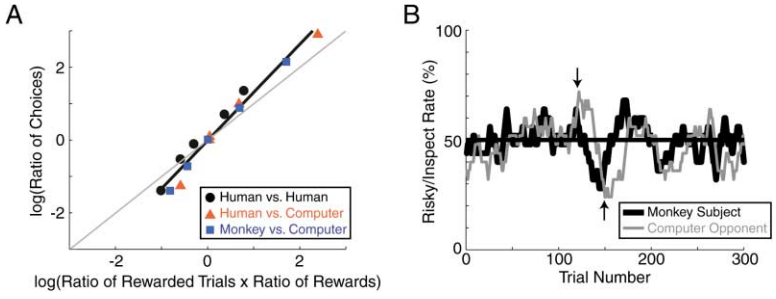

Figure 4. The Influence of Strategic Opponent on Choice Behavior

(A) Matching law plot. Ratio of choices plotted on the ordinate was calculated as the number of risky choices divided by the number of certain choices. On the abscissa, the ratio of rewarded trials was calculated as the number of rewarded risky trials divided by the number of rewarded certain trials, and the ratio of rewards was calculated as the magnitude of the risky reward divided by the magnitude of the certain reward. Both axes are plotted in log coordinates. Each datum point represents averages from blocks with the same cost of inspection. Same data sets as Figure 3B. The best fit linear regression (black line) has a slope of 1.32. The line of unity (gray) has a slope of 1.

(B) Dynamics of strategic interaction between monkey subject (black) and computer algorithm (gray). See text for details.

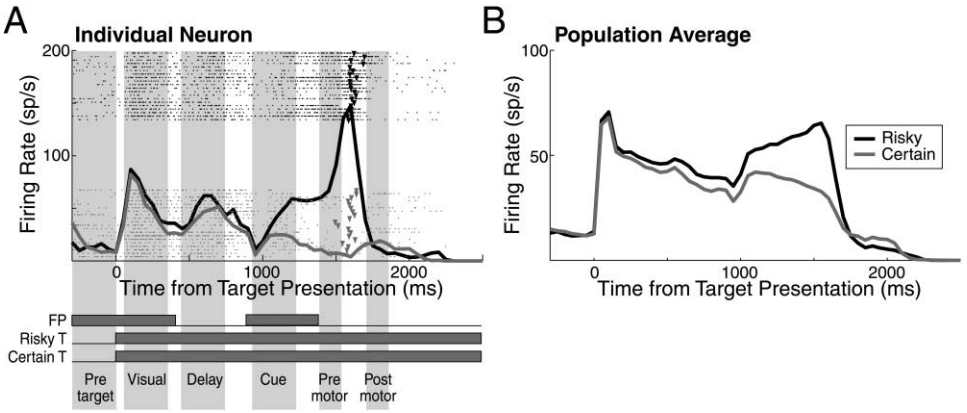

Figure 5. LIP Activity Related to Response Direction

Data are shown during blocks of trials in which the cost of inspection was 0.5 and the monkey chose either the target in (black, risky target) or opposite (gray, certain target) the neuron's response field. (A) Individual neuron. Each raster represents the time of an extracellularly recorded action potential, and each row of rasters represents action potentials from the last 20 trials of each response option. Triangles represent the time of initiation of the saccade for each trial. Average firing rate is displayed as a poststimulus time histogram with 50 ms bin widths for all trials during a block. The timing of the presentation of the fixation point (FP), risky target (Risky T), and certain target (Certain T) are shown at the bottom of the panel. The six epochs during which neuronal activity was calculated for subsequent analyses are depicted by the translucent gray bars. (B) Neuronal population activity ($n = 52$) of the same comparison as in (A). Note that the ordinate is half the range of that in (A).

terns in the behavior of the subjects this might have reduced the nonrandom component of the behavior we observed during these particular blocks (Barraclough et al., 2004).

Neuronal Activity

Basic Inspection Game

For these experiments, we studied 52 LIP neurons with a mixture of *instructed* and/or *inspection game trials*. Figure 5A plots the average response of a single LIP neuron during inspection game trials as a function of whether or not the monkey decided to look at the red risky target located in the neuronal response field (black) or the green certain target located opposite the neuronal response field (gray). After the fixation point reappeared as yellow, indicating an inspection game trial, neuronal activity distinguished between trials that ended within or opposite the neuronal response field. Overall, neuronal activity became selective for movement direction for the population of neurons before the fixation point indicated whether it was an instructed or inspection game trial (Figure 5B; $p < 0.01$ for visual, delay, cue, and premotor epochs, paired t test, $n = 52$). For the remainder of this paper, we restrict our analyses to trials directed into the response fields of our neurons (Figures 5A and 5B, black). Therefore, any differences in neuronal activity cannot be attributed to sensory stimuli or movement parameters that remain identical on every trial.

Below, we test three predictions of the hypothesis that the activity of neurons in area LIP reflects the relative subjective desirability of upcoming saccades. First, the firing rates of LIP neurons should vary with desirability as it varies across blocks of trials in a nonstrategic setting. We tested this by varying the magnitude of reward associated with each option across sequentially presented blocks composed entirely of instructed trials. This first prediction constitutes a replication of Platt and Glimcher (1999). Second, whenever the animals are at a mixed strategy equilibrium during the inspection game, the average firing rates of LIP neurons should be fixed. This is because the average subjective desirability of the responses are presumed to be equivalent (Nash, 1950), irrespective of the response probability, reward magnitude, or reward probability of the available options at equilibrium. On a trial-by-trial basis, however, the mixed strategy equilibrium is presumed to be maintained by small fluctuations in the subjective desirability of each option around this fixed level caused by dynamic interactions with the opponent. Therefore, the third prediction is that LIP activity will be correlated with these small fluctuations in subjective desirability on a trial-by-trial basis.

LIP Firing Rate Varies with Desirability during Instructed Trials

To test the first prediction, desirability was manipulated by changing the amount of liquid reward associated with the targets in two successive blocks composed entirely of instructed trials while the probability of responding to the targets was maintained at 50%. The first two blocks of trials in Figure 6A show a running average of choice behavior (jagged black line) and neuronal activity (gray dots) during the visual epoch under these conditions. During the first block, a movement to the red target in the neuron's response field yielded twice as much reward (0.5 ml of water) as a movement to the green target opposite the neuron's response field (0.25 ml of water). During the second block of trials, the payoffs associated with each target were reversed such that a movement to the red target in the neuron's response field yielded only half as much reward as the green target opposite the neuron's response field. Neuronal firing rate was high during the large reward condition and low during the small reward condition. This

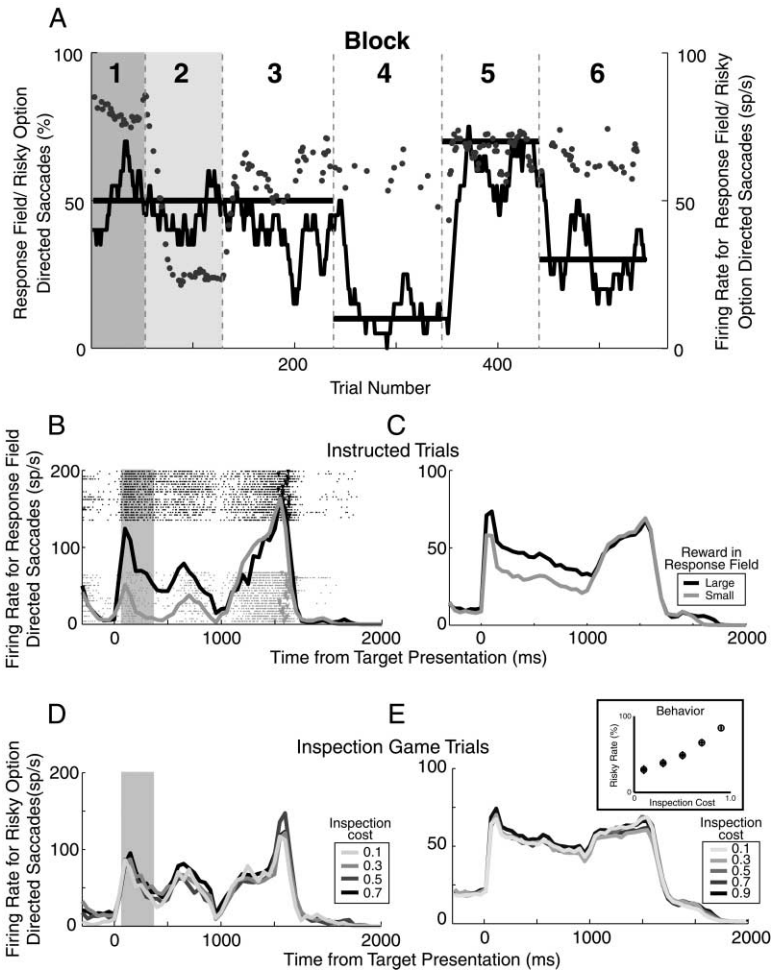

Figure 6. Monkey's Choice Behavior and Corresponding Activity of LIP Neurons during Instructed and Inspection Game Trials

Same neuron as in Figure 5A. (A) Two purely instructed trial blocks (block 1, large 0.5 ml reward into response field; block 2, small 0.25 ml reward into the response field) followed by four blocks of inspection game trials. The jagged black line denotes a 20 trial running average of the percentage of saccades directed to the target in the neuron's response field. The horizontal black lines denote the predicted percentage of trials directed to the target in the neuron's response field either as instructed by the color of the fixation point (instructed trials) or by the Nash equilibrium equations (inspection game trials). The gray dots denote a 20 trial running average of the firing rate during the visual epoch only for those trials in which the animal chose the target in the neuron's response field. Because only those trials in which the monkey looked into the response field were included, this firing rate average is composed of more observations during blocks when the monkey chooses this option more often (e.g., block 5) than less often (e.g., block 4). (B) Average activity of the cell in (A) throughout the duration of large (black) and small (gray) rewarded blocks of instructed trials. Same conventions as Figure 5A. (C) Population firing rate ($n = 30$) during instructed trials. (D) Average activity of the cell in (A) throughout the duration of four blocks of inspection game trials with different costs of inspection. The interleaved instructed trials were removed for this analysis. (E) Population firing rate ($n = 52$) during five blocks of inspection game trials in which the opponent's cost of inspection variable ranged from 0.1 to 0.9. (Inset) Average choice behavior associated with the five waveforms in (E). Translucent gray bar in (B) and (D) denotes the visual epoch sampled in (A).

difference in firing rate occurred early during the trial, when there was uncertainty as to which target would be rewarded, for both this individual neuron (Figure 6B) and our neuronal sample (Figure 6C; $p < 0.01$, paired t test for visual and delay epochs, $n = 30$). After the fixation point indicated that the target in the response field would be rewarded, the firing rate of this individual neuron (Figure 6B) and our neuronal sample (Figure 6C) was the same irrespective of the amount of reward associated with the movement. The early difference in firing rate could not be the result of visual stimuli, response probability, or movement direction because each of these variables remained the same across both blocks of trials. Thus, the first prediction that LIP firing rates vary with the desirability of saccades during the instructed task was fulfilled (see also Platt and Glimcher, 1999).

LIP Firing Rate Constant at Behavioral Equilibrium during Inspection Game Trials

To test the second prediction, we changed the algorithm's cost of inspection variable across blocks of inspection game trials while recording the activity of this same neuron (Figure 6A, blocks 3–6). At equilibrium, the relative subjective desirability of the two choices presumably remained constant even as response probability fluctuated across blocks of trials. Concomitant neuronal activity remained relatively constant at a level between that observed for the high and low reward conditions during instructed trials. This equivalence in firing rate remained throughout the duration of the trial as observed in the average firing rate of this neuron for each block of trials (Figure 6D). This was true even though the objective expected value of this option changed from block to block (Figure 3B). Unlike instructed trials, the reappearance of the fixation point did not resolve the uncertainty of whether the risky target would be rewarded. Consistent with the hypothesis that these neurons reflect the relative subjective desirability of the available actions, the firing rate of the neuron during the cue and premotor epochs was less for inspection game trials (Figure 6D) than for instructed trials (Figure 6B).

Next we examined 52 neurons (including 13 that were also studied in the instructed task) while animals played five blocks of the inspection game in which their response probability varied as the algorithm's cost of inspection variable ranged from 0.1 to 0.9 (Figure 6E, inset). LIP firing rates remained constant across the five

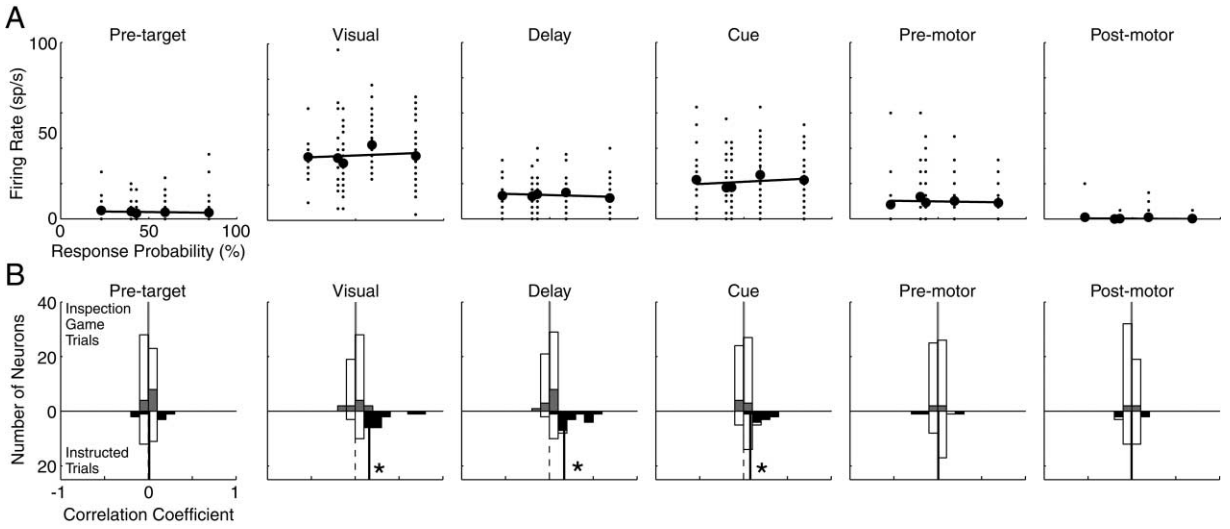

Figure 7. Correlations between Individual LIP Activity and Response Probability

(A) The firing rate of an individual LIP neuron and response probability during inspection game trials. The small data points represent firing rates on single trials, and the large data points represent the average firing rate (±SEM) for each block. The black lines represent the least squares linear regressions fit to these data. The same plots were made for the instructed task (data not shown) in which the abscissa was replaced with the relative magnitude of rewards for the large and small rewarded blocks.

(B) Histograms of correlation coefficients for the population obtained from the analyses in (A). Filled bars represent neurons with statistically significant correlations (Fischer's *r* to *z* test, $p < 0.05$) for inspection game (gray) and instructed (black) trials. Solid vertical lines represent the average correlation coefficients for inspection game (gray) and instructed (black) trials. Those average correlations that differ significantly from zero (dashed lines) are denoted with an asterisk (t test, $p < 0.01$).

blocks of trials (Figure 6E, $p > 0.01$, paired t test for all epochs, $n = 52$), thus fulfilling the second prediction that LIP activity, like relative subjective desirability, is equivalent at mixed strategy equilibriums.

At the Nash equilibrium, the relative subjective desirability of options should remain equivalent regardless of the response probability, reward magnitude, or reward probability associated with those options. If LIP activity reflects relative subjective desirability, then firing rates should also remain constant at equilibrium despite manipulations of these variables. The relationship between the firing rates of individual LIP neurons and response probability is quantified in Figure 7. The top panels represent the trial-by-trial firing rate of a single LIP neuron as a function of response probability for each epoch. Response probability is represented as the overall percentage of saccades directed toward the risky target during a block of trials in which the algorithm's cost of inspection was fixed. There is no significant correlation between the firing rate of this neuron and response probability for any epoch (Figure 7A). Although some individual neurons showed a significant correlation (Figure 7B, gray-filled histograms above the zero line, $p < 0.05$), the average correlation coefficients (gray lines) did not vary from zero for any of the epochs tested ($p > 0.01$, Fischer's *r* to *z* test). This same analysis was performed for LIP firing rates and the relative magnitude of rewards across two blocks of instructed trials (data not shown). As expected under these conditions, many neurons showed a significant positive correlation during the epochs before the fixation point indicated which target would be rewarded (Figure 7B, black-filled histograms below the zero line, $p < 0.05$) and the average of these correlation coefficients differed from zero (black lines) for the visual, delay, and cue epochs ($p < 0.01$).

For a subset of 24 neurons, we examined the effects of reversing the locations of risky and certain targets during a block of trials in which the algorithm's cost of inspection variable was fixed at 0.5. Switching the target in the response field from the risky to certain option changed both the probability of reward (i.e., from ~50% to 100%) and the magnitude of reward (i.e., from 2 objective units to 1 objective unit of reward), but according to game theory, the relative subjective desirability of the two options should have remained constant. Firing rates should differ across the blocks if they reflect either the probability or magnitude of reward but remain constant if they reflect the relative subjective desirability of the options. Firing rates remained constant, consistent with the hypothesis that these neurons encode the relative subjective desirability of choices (Figure 8, $p > 0.05$, paired t test, $n = 24$, for all six epochs).

LIP Firing Rate Correlated with Dynamic Estimate of Relative Subjective Desirability

The third prediction was that LIP firing rates should be correlated to the small trial-by-trial fluctuations in the relative subjective desirability of the choices resulting from the strategic interactions of the two opponents at behavioral equilibrium. Casual inspection of the activity pattern of the neuron shown in Figure 6A during inspection game trials suggests it may have this character. Of course, developing such a correlation is difficult because we do not know exactly how the subjective desirability of the two movements is assessed on a trial-by-trial basis by the animal. However, we attempted to

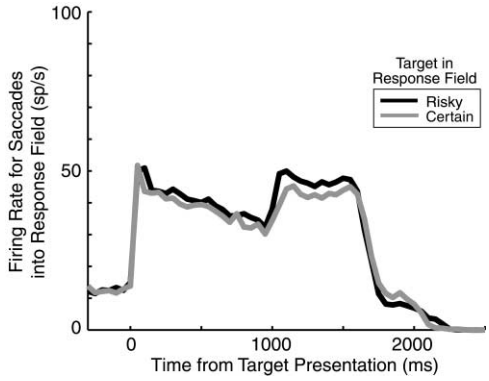

Figure 8. Influence of Changing Reward Magnitude and Reward Probability on LIP Firing Rate

The target in the response fields was changed from the risky (black) to certain (gray) option across two blocks of the inspection game when the opponent's cost of inspection variable was fixed at 0.5. Switching targets changed both the probability and magnitude of reward while the relative subjective desirability of the target in the response field remained unchanged. n = 24.

derive a crude estimate of how the relative subjective desirability of choosing each option might fluctuate from trial to trial and asked whether this estimate was correlated with the fluctuations in neuronal rate.

We therefore employed the algorithm developed for use as the computer opponent to estimate how subjective desirability might be fluctuating from trial to trial. The computer algorithm normally tracks the monkey's behavior and combines this with its own potential payoffs to determine the desirability of inspecting and not inspecting on the upcoming trial (see Experimental Procedures for details). We simply reversed the inputs to this algorithm, having it analyze offline the choice behavior of the computer and the payoffs received by the monkey throughout a block of trials to calculate the subjective desirability of choosing the risky option on the upcoming trial. Finally, using the monkey's behavior recorded during the same experimental session, we performed an optimization based on maximum likelihood methods on the variable α (see Experimental Procedures, Equation 7) that determined the learning rate of the reinforcement learning algorithm. This optimization successfully converged for 48 out of 52 neurons with a mean α of 0.27 $\pm$ 0.13. This trial-by-trial estimate of the relative subjective desirability of the two options was then correlated to the trial-by-trial measurement of LIP activity. Note that the presence of 20% interleaved instructed trials were not ideal for this analysis, and although far from a perfect solution, these trials were simply excluded.

To see how any such correlation evolved throughout the duration of a trial, we segregated each trial into six sequential epochs. For this neuron, there was a positive correlation between our estimate of relative subjective desirability and firing rate for two of the four epochs during which the targets were visible (Figure 9A, $p < 0.05$, Fisher's *r* to *z* test, during visual and delay epochs). Of our 48 analyzed neurons, 23 had a significant positive correlation between these two variables during at least one of the epochs ($p < 0.05$). Conversely, only 6 of the 48 neurons had a significant negative correlation between these variables during the same epochs ($p < 0.05$). To ensure that this was not the result of a spurious secondary correlation, we next performed a multiple regression analysis that correlated firing rate with saccade amplitude, peak velocity, latency, and the cost of inspection variable. We then repeated our correlation analysis on the residual variance that remained after this multiple regression had been performed (Figure 9B, relative desirability). During the inspection game trials, none of these other individual regressions reached the level of significance except the correlation between saccade amplitude and firing rate during the postmotor epoch ($p < 0.05$). The residual correlation between this estimate of relative subjective desirability and firing rate remained significant during the visual, delay, and cue epochs ($p < 0.05$). Performing the same multiple regression analysis on the instructed trial blocks showed that neuronal activity was also significantly correlated with a subjective desirability estimate under those conditions during the visual and cue epochs (data not shown, $p < 0.05$). While we do not yet know how our monkeys determine the subjective desirability of each available option, this crude estimate of that value on a trial-by-trial basis is correlated with the trial-by-trial fluctuations in neuronal rate that we observed. Once again, this is exactly the observation one would expect if area LIP neurons reflect the relative subjective desirability of saccades.

Encoding Relative versus Absolute Subjective Desirability

Although our results have supported the notion that neurons in area LIP represent the subjective desirability of saccades, they have not addressed whether this representation is in absolute or relative terms. Nash (1950) envisioned that the subjective desirability of each option was represented in absolute terms and it was that option whose subjective desirability was highest that was subsequently chosen. The neural instantiation of this in area LIP would correspond to the firing rates of single neurons being a function of the absolute subjective desirability of the option in their response fields. However, subsequent behavioral studies (Herrnstein, 1961; Kahneman and Tversky, 1979), coupled with our understanding of LIP physiology (Platt and Glimcher, 1999; Gold and Shadlen, 2001), suggest that the subjective desirability of options are represented in relative terms during decision making.

Therefore, we wished to explicitly test the hypothesis that LIP neurons encode the *relative* subjective desirability of actions rather than the *absolute* subjective desirability of actions. Monkeys performed two blocks of the inspection game in which the cost of inspection variable was fixed at 0.5 and, therefore, responses were typically divided equally between the risky and certain targets. Standard magnitudes of reward were used for one block of trials, whereas the magnitudes of reward were doubled for all targets in the other block. If LIP activity is sensitive to the absolute subjective desirability of the saccade in the neuron's response field, the neurons should fire more for the block of inspection game trials on which the rewards are doubled. If, however, LIP activity is sensitive only to the relative subjective

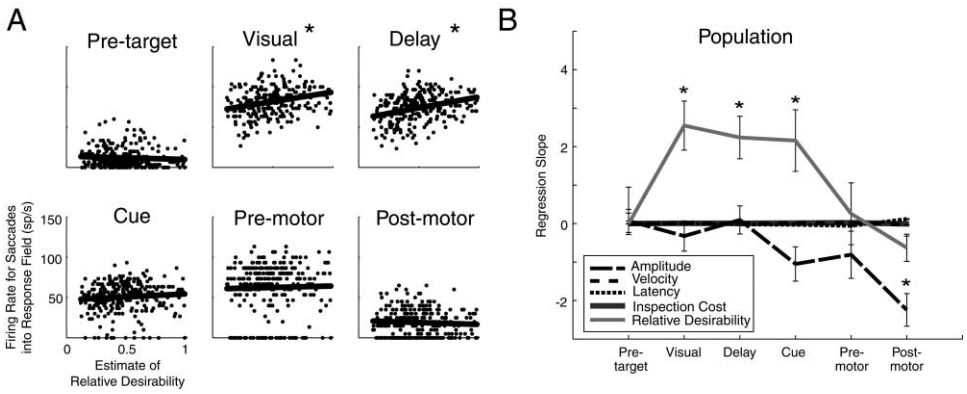

Figure 9. Correlation between LIP Firing Rates and an Estimate of the Relative Desirability of Choices

(A) Trial-by-trial correlations between firing rate and an estimate of relative desirability for each of the six epochs. Only inspection game trials in which the movement was directed into the response field were included in this analysis and, thus, the sensory inputs and motor responses were largely identical for each trial. Statistically significant least squares linear regressions (black lines) are denoted with an asterisk ($p < 0.05$, Fisher's r to z test).

(B) Regression analysis for the population of neurons ($n = 48$) for the six epochs during the inspection game trials. The average slopes of the regression lines of firing rates versus five behavioral parameters are shown. Because the range of the abscissa varies greatly between different parameters, a comparison of the absolute value of the slopes for each of these parameters can be misleading. However, those values that differed significantly from zero are denoted by an asterisk ($p < 0.01$), and this statistic is not influenced by the range of abscissa values.

desirability of choices inside compared to outside the neuron's response field, then firing rate should be roughly the same for both of these blocks of trials. Consistent with LIP encoding the relative subjective desirability of saccades, there was no significant change in the firing rate of these neurons during this manipulation (Figure 10, $p > 0.05$, paired t test, $n = 22$, for all six epochs).

Discussion

Over the course of the last decade, a number of researchers have begun to develop evidence suggesting that areas in the frontal cortex, posterior parietal cortex, and the basal ganglia participate in some classes of decision making (Schall and Thompson, 1999; Hikosaka et al., 2000; Glimcher, 2003a). Typically these studies have demonstrated that when the actual costs or benefits of an action are manipulated, the activity in brain areas associated with that action is also modulated. Both humans and animals are known, however, to make decisions based on a subjective internal representation of the costs and benefits of their actions rather than based upon the actual costs or benefits of those actions (Stephens and Krebs, 1986; Kreps, 1990; Kagel and Roth, 1995; Krebs and Davies, 1996; Glimcher, 2003b). Neural correlates of these subjective internal representations of costs and benefits during decision making, however, are rare.

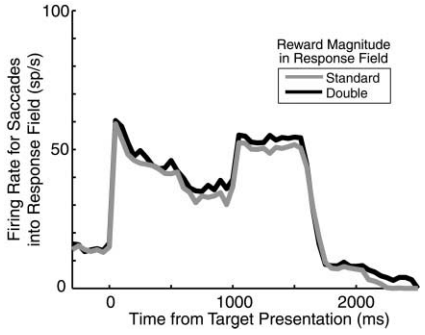

Figure 10. Relative versus Absolute Subjective Desirability

Data are shown only for trials in which saccades were directed into neuronal response fields during blocks in which the algorithm's cost of inspection variable was fixed at 0.5. For both blocks of trials, the risky target was located in the neuron's response field. In one block of trials, the standard reward magnitudes were used (gray), and, in the other block, reward magnitudes associated with both targets were doubled (black). This manipulation changed the absolute subjective desirability of the option in the response field while the relative subjective desirability of the target in the response field remained unchanged. $n = 22$.

In the Breiter et al. (2001) study, for example, human subjects passively viewed a lottery. Although activity in the sublenticular extended amygdala was correlated with subjective evaluations of the desirability of outcomes, or *prospects*, these evaluations did not influence the outcome of the lottery. Similarly, Platt and Glimcher (1999) had monkeys perform a task involving early uncertainty about what eye movement would be reinforced, followed by an instructional cue that resolved this uncertainty. Although parietal activity during this early period was correlated with the actual values of the two actions, the later instruction ultimately indicated which saccade would be rewarded. Therefore, it was equivocal whether this parietal activity would be involved in decision making under free choice conditions. Other neurophysiological studies have employed dynamic reward contingencies to elicit more deliberate, free choice behavior (Coe et al., 2002; Shima and Tanji, 1998); however, there was no effort to gauge the subjective desirability during these tasks.

In the experiment presented here, we examined a form of free choice behavior that has been well studied in humans, decision making during strategic conflict (Camerer, 2003). Our behavioral experiments indicate that

rhesus monkeys produce behavior that is empirically very similar to the behavior produced by humans engaged in the same strategic conflict. To gain access to the subjective desirability of the two options, we exploited the Nash equilibrium concept that contends that subjective desirability is rendered equivalent by the strategic play of the two opponents.

We tested the hypothesis that LIP activity was correlated with subjective desirability under a number of experimental manipulations. First, we verified that neuronal activity was modulated when the desirability of the options was manipulated with reward magnitude during an instructed task (Platt and Glimcher, 1999). When animal subjects were free to select either of two options during the inspection game, we found that the average neuronal activity remained constant like the average subjective desirability at a Nash equilibrium. Furthermore, trial-by-trial fluctuations in neuronal activity were correlated with our trial-by-trial estimate of subjective desirability. Our results are not consistent with the view that LIP activity encodes strictly the probability or intention to make a saccade (Andersen and Buneo, 2002) because activity remained constant across blocks of trials while response probability varied. Moreover, LIP activity cannot encode strictly the probability of receiving a reward or magnitude of reward associated with a certain saccade (Platt and Glimcher, 1999) because neuronal activity remained unchanged whether a risky large reward or smaller certain reward was in the neuron's response field. Finally, LIP cannot encode the objective expected value of saccades because we showed that the expected value of the risky target varied across blocks of trials even though neuronal firing rates did not. Instead our findings are consistent with the hypothesis that LIP neurons encode relative subjective desirability in a manner appropriate for selecting a course of action.

Importantly, we also found that LIP activity was more tightly correlated with relative rather than absolute subjective desirability. This study (Figure 6) and previous work out of our lab (Platt and Glimcher, 1999) has reported that the firing rates of neurons in area LIP are correlated with the magnitude of reward yielded by the movement they encoded divided by the sum of all available rewards; the neurons encoded the relative values (in this limited sense) of the saccades with which they were associated. However, many previous investigators had informally varied the magnitude of reward delivered during single target tasks and seen little or no effect of that variation on LIP firing rates. These seemingly incompatible results are, in fact, exactly what would be expected if LIP activity encodes the relative subjective desirability of saccades. Varying the magnitude of reward in a single target task changes the absolute subjective desirability for the option in the response field but the relative subjective desirability remains unchanged because it is compared to the only available option, itself. In fact, LIP displays differential firing with different reward magnitudes early during instructed trials, but the activity becomes equivalent after the instruction indicates which target will be rewarded, essentially changing a two target task into a one target task (Figures 6B and 6C) (see Basso and Wurtz, 1998, for similar result in the superior colliculus). Also consistent with this view is the finding that doubling the reward magnitude of both options during the inspection game had no effect on LIP firing rates given that this manipulation did not change the relative subjective desirability of those options. If LIP activity encodes relative subjective desirability, the prediction follows that firing rates should scale as a function of the number of available options. For example, when faced with only one option (e.g., a single target task) firing rates should approach the maximum of a neuron's dynamic range, when faced with two equally desirable options (e.g., our inspection game) firing rates should be roughly one-half of that maximum, and when faced with three equally desirable options (e.g., rock-paper-scissors) firing rates should be roughly one-third of that maximum. In this context it is interesting to note that the average firing rate of the posterior parietal neurons we examined in the inspection game was about 50 Hz. This evidence that the neuronal architecture encodes the value of an action relative to the value of other actions also accords well with studies which suggest that human decision making is almost always based on relative values (cf. Kahneman and Tversky, 1979).

Function of Posterior Parietal Cortex

Over the course of the last two decades, a number of hypotheses have been advanced as to the principle functions of the posterior parietal cortex. Studies in monkey suggest that the cluster of cortical subareas lining the intraparietal sulcus may participate in the allocation of attentional resources (Colby and Goldberg, 1999), the transformation of sensory data into coordinate frameworks appropriate for movement generation (Andersen et al., 1997), and the selection of movements for execution (Andersen and Buneo, 2002). Importantly, these hypotheses are not necessarily mutually exclusive; indeed they are computationally and behaviorally interrelated. As a result, evidence examining the relationship between neuronal firing rates in these areas and these psychological processes might be expected to proceed in parallel. The results presented here suggest a correlation between activity in area LIP and the psychological processes involved in decision making. Over the course of the last several decades there have been several attempts to demonstrate that neural activity in a particular region is uniquely associated with a single psychological process, even if that psychological process shares many features with other closely related processes. While there is little doubt that many psychological processes will prove to have separable neural bases, simply on the grounds of efficient neural coding (Barlow, 1961) we might expect to see an incomplete neural segregation of psychological processes that share many informational features. Importantly, the results here do not test the hypothesis that this pattern of neural activations may also be related to the allocation of attention. They merely indicate that activity in area LIP, which is anatomically tied to the saccade-generating circuitry itself (Platt et al., 2003), carries signals appropriate for saccade selection.

Conclusion

Neurobiologists have been increasingly interested in how the primate neural architecture produces voluntary

decisions but have been limited in their ability to bring these behaviors into a laboratory setting. Previous neurophysiological tasks have been unsatisfying in this regard because they required animal subjects to react in a stereotyped manner to fixed stimuli or the animals were free to make any response but received rewards for only one action. To study voluntary decision making in nonhuman primates, therefore, requires tasks specialized for this class of decision making.

Within the social sciences, game theory has become a popular tool for the study of voluntary decision making. Economists have argued that when humans make voluntary decisions they do so by weighing the relative subjective desirability of the actions available to them. During mixed strategy games, this results in unpredictable behavior from choice to choice and subjects self-report that they have behaved in a volitional manner. Indeed, our human subjects reported this conviction when they played the inspection game in our laboratory. The similarity in behavioral data across species presented here suggests that nonhuman primates provide a good model for human decision makers under these conditions of strategic conflict.

While we cannot ask nonhuman primates to report their subjective impressions of the decision making process during these experiments, the data described here may still shed light on the mechanisms of human voluntary decision making. The neural data presented here suggest that the neurons of area LIP encode the relative subjective desirability of saccadic eye movements. If human choices are guided by circuits involving neurons like those in monkey area LIP, then it is tempting to speculate that the average firing rates of these neurons may also encode the subjective desirability of actions in humans. Furthermore, variations in these spike trains may be the source of unpredictability in decisions from choice to choice which would be consistent with the highly variable spike trains of cortical neurons (Dean, 1981; Glimcher, 2005; Glimcher and Dorris, 2004; Tolhurst et al., 1981). Completely unanswered, however is a question central to game theory: how does this stochastically maintained equilibrium of neuronal activities give rise to game theoretic behavioral equilibriums in which subjects mix two or more responses asymmetrically? We can only speculate that future studies of neuronal and behavioral dynamics during game play may provide insight into the mechanistic nature of the volitional process.

Experimental Procedures

Subjects played a repeated, mixed strategy, game theoretic task known as the inspection game for either monetary (humans) or water (monkeys) reward. All human procedures were approved by the New York University Committee on Activities Involving Human Subjects. All monkey procedures were approved by the New York University Animal Care and Use Committee and were in compliance with the Public Health Service's *Guide for the Care and Use of Laboratory Animals*.

General Behavioral Task

Subjects played the inspection game, whose general payoff matrix is shown in Figure 1A. On each trial, the subject's payoff was determined by their own action and that of their opponent. It is assumed that rational decision makers choose the option that they perceive as being most desirable on each trial. If both the subject and opponent act rationally in this sense, then a behavioral equilibrium will be reached in which the average subjective desirability for each option is rendered equivalent for both players through their dynamic interaction. The available options and their associated payoffs will be described first for the experimental subjects (or "employees" in inspection game terminology) followed by those of the opponent (or "employers" in inspection game terminology).

On each trial, the experimental subject was simply required to choose either the *certain option* or the *risky option* (Figure 1A). The *certain option* was guaranteed to yield the same small reward regardless of the opponent's choice. The *risky option* yielded twice the magnitude of reward if the opponent did not inspect but yielded zero reward if the opponent did inspect. Unless otherwise stated (i.e., Figure 10), the options available and their associated payoffs were the same for the subjects across all blocks of trials.

Similarly, the subject's opponent had to choose either the *inspect* option or the *no inspect* option on each trial. The payoffs associated with inspecting were varied across blocks of trials by manipulating the *cost of inspection* variable (Figure 1A, variable I).

The normative rates of choosing each option at the Nash equilibrium are outlined below. The subject's rate of choosing the risky option increases linearly with the opponent's cost of inspection variable given the simplifying assumptions that the subjective desirability of an option is the normative expected value of that option (expected value = probability of receiving a reward × magnitude of reward), that block boundaries can be neglected, and that each player can assume his opponent to be perfectly rational. Of course, perfect linearity should not be expected in subjects' actual behavior (e.g., Figure 3) for two principle reasons: (1) true subjective desirability is known to vary from objective expected value, and (2) there are asymmetries in information associated with the inspection game (see Results).

For the subject, at Nash equilibrium the subjective desirability for choosing the certain option is equal to the subjective desirability for choosing the risky option:

$$SD(certain) = SD(risky) \quad (1)$$

which given the payoff matrix (Figure 1A) expands to

$$p(inspect) \times 0.5 + (1 - p(inspect)) \times 0.5 = p(inspect) \times 0 + (1 - p(inspect)) \times 1 \quad (2)$$

solving for p(inspect)

$$p(inspect) = 0.5 \quad (3)$$

where SD(certain) is the subjective desirability for choosing the certain option, SD(risky) is the subjective desirability for choosing the risky option, p(inspect) is the probability of the opponent choosing to inspect, and 1 − p(inspect) is the probability of the opponent choosing to not inspect when at equilibrium.

Similarly, for the opponent at Nash equilibrium the subjective desirability for choosing the inspect option is equal to the subjective desirability for choosing the not inspect option.

$$SD(inspect) = SD(not\ inspect) \quad (4)$$

which given the payoff matrix (Figure 1A) expands to

$$p(risky) \times (1 - I) + (1 - p(risky)) \times (2 - I) = p(risky) \times 0 + (1 - p(risky)) \times 2 \quad (5)$$

solving for p(risky)

$$p(risky) = I \quad (6)$$

where p(risky) is the probability of the subject choosing the risky option and 1 − p(risky) is the probability of the subject choosing the certain option when at equilibrium. Across blocks of trials, I, the opponent's cost of inspection variable, was experimentally manipulated between 0.1 and 0.9 in steps of 0.2 with the intended effect of changing the subject's rate of choosing the risky option (Equation 6).

Human Behavioral Task

Human subjects were required to choose with a computer mouse one of two buttons on a monitor that corresponded to either the *certain* or the *risky* option (Figure 1B). Subjects were naive to the nature of the payoff matrix and were simply instructed to "make as much money as possible," learning through trial and error. The payoff for each trial was presented in the center of the screen at the end of each trial along with a cumulative total of earnings over the last ten trials. The first block of 50 trials was a practice session. Afterward, five separate blocks of 150 trials, each associated with a different cost of inspection, were played in a randomized order. At the end of the session, subjects were paid their cumulative earnings that depended on performance and were typically about $35 US. Subjects were not aware of the nature of their opponent, which could be either another human in a different room or a standardized computer algorithm (see below).

Monkey Behavioral Task

Monkeys played the same computer algorithm as the human subjects in an oculomotor version of the inspection game for water reward (Figure 1C). On each trial, the monkey began by looking at a central yellow fixation point. Two targets were presented 300 ms later: a red *risky* target in the center of the neuron's response field and a green *certain* target at a position of equal eccentricity but in the opposite direction from the fixation point. The fixation point was extinguished for 500 ms and could reappear as one of three colors: yellow, red, or green. After 500 ms, the fixation point disappeared, which was the cue for the monkey to indicate its choice with an eye movement to one of the targets within 70–500 ms to obtain reward. On 80% of the trials, the fixation point reappeared as yellow. On these *inspection game trials*, the monkey could voluntarily choose either target. On each of 10% of the remaining trials, the fixation point reappeared either as red or green. On these *instructed trials*, the monkey was required to choose the target that matched the color of the fixation point; 2 units of reward were earned for red instructed trials and 1 unit of reward for green instructed trials. Although not entirely successful (i.e., Figure 5B), the intended purpose of these interleaved instructed trials was to prevent the monkey from fully deciding on a saccadic target until each trial was underway. Otherwise, these instructed trials that were interleaved with inspection trials were not analyzed further. During some experimental sessions, however, monkeys performed blocks composed solely of instructed trials (e.g., Figures 6B and 6C). These instructed trials were analyzed to determine how the desirability of options affected LIP firing rates as it was experimentally manipulated by changing the magnitude of reward associated with the options across blocks of trials.

Computer Opponent

For the majority of experiments, both human and monkey subjects competed against a standardized computer algorithm which played the role of the opponent (see http://www.cns.nyu.edu/~glimcher/inspection_game/ for MATLAB code of the complete algorithm). In brief, the computer algorithm worked by tracking two variables of the subject's behavior: (1) the history of the subject's choices to give an estimate of the overall p(risky), and (2) the subject's repetition rate (rep_{actual}), that is, how often a subject repeated the response of the previous trial. The expected repetition rate ($rep_{expected}$) was calculated for a given p(risky) assuming the probability of a response on each trial was controlled by a random process independent of previous choices.

The computer updated its estimate of the probability of the subject choosing the risky option on each trial using the following reinforcement learning algorithm:

$$p(risky)_{t+1} = p(risky)_t + \alpha(C - p(risky)_t) \quad (7)$$

where t is the current trial. If the subject chose the risky option on the current trial, $C = 1$. If the subject chose the certain option on the current trial, $C = 0$. α, which determined the rate of learning, was set to 0.1 during game play. The iterative nature of this reinforcement learning algorithm results in an estimate of $p(risky)_{t+1}$ that is derived from all the subject's previous choices with the most recent choices being weighted most heavily.

$$rep_{expected} = (p(risky) \times p(risky)) + ((1 - p(risky)) \times (1 - p(risky))) \quad (8)$$

the difference in the rep_{actual} from $rep_{expected}$ was used to bias the computer's estimate of p(risky) for the upcoming trial

$$p(risky)_{corrected} = p(risky) + \lambda(rep_{expected} - rep_{actual}) \quad (9)$$

in which λ was set to 0.1.

The variable $p(risky)_{corrected}$ represents an estimate of the probability of the subject choosing the risky option given his past history of doing so and allows the algorithm to exploit dependencies of upcoming behavior on actions taken during the previous trial. The variable $p(risky)_{corrected}$ was substituted for p(risky) in calculating the relative subjective desirability of inspecting and not inspecting on the upcoming trial which, in turn, was used to guide the opponent's choice. In addition, an *exploration bonus* was added, which gradually increased as the algorithm continued to produce a single response. Because p(risky) was only updated after inspection trials—and by extension so was the estimate of relative subjective desirability as well—the exploration bonus was necessary so the computer did not get stuck always choosing no inspect. Variations of this exploration bonus are used for similar purposes by other reinforcement learning algorithms in an effort to strike a balance between exploring for potentially more desirable resources and exploiting resources already available (Sutton and Barto, 1998).

The computer opponent would be deterministic if it always chose the option with the highest desirability on every trial. If a subject had sufficient precision in a trial-by-trial estimate of their own p(risky), they could accurately predict the actions of the algorithm. In order to incorporate stochasticity into the actions of the algorithm, we employed a decision rule that converted relative subjective desirability into a response probability. When inspecting and not inspecting had equal subjective desirability, the decision rule randomly selected the inspect and no inspect options with equal probability. As subjective desirability increased for one option over another, the probability that the more desirable response would be selected increased gradually.

Neurophysiology

In two monkeys, we recorded the activity of single neurons that were located in the lateral bank of the intraparietal sulci using standard electrophysiological techniques. The location of the neurons was established using neuronal response properties and neurosonography (Glimcher et al., 2001; Platt and Glimcher, 1997). After isolation of a single neuron, animals were required to make a series of 50 to 100 eye movements beginning from a central fixation point and directed to eccentric targets selected randomly from among several hundred peripheral locations. The monkeys were trained to withhold the eye movement to the eccentric target until the fixation point was extinguished (500–800 ms after target presentation). We used this *delayed saccade task* to identify the center of the neuronal response field, defined as the position of the eccentric target for which the neuron was maximally active during the delay period.

To study the evolution of neuronal activity, we divided each trial into six epochs (Figure 5A): *pre-target* (300 ms before presentation of targets), *visual* (50–350 ms after target presentation), *delay* (50–350 ms after initial fixation point offset), *cue* (50–350 ms after fixation point reappearance), *pre-motor* (0–150 ms after second fixation point offset), and *post-motor* (350–500 ms after second fixation point offset). Neurons were included in this study if, during inspection game trials with a 0.5 cost of inspection, their activity: (1) was choice selective (activity was higher for trials into the response field than for movements opposite the response field during at least one epoch while the choice targets were present at $p < 0.01$ by paired t test during visual, delay, cue, or pre-motor epochs), (2) increased with the presentation of a target in the neuronal response field (there was a significant increase in activity 50–150 ms after the targets were presented compared to the 150 ms before the targets were presented at $p < 0.01$ by paired t test).

Acknowledgments

We thank P. Zak for advice on analysis and critical feedback; D. Heeger and S. Greenstein for useful feedback on the manuscript;

S. Zhu and D. Ballard who provided us with much of the algorithm for the computer opponent; B. Lau for assistance optimizing the reinforcement learning algorithm to the monkey's behavior; A. Gindes and E. Ryklin for technical assistance; H. Bayer, M. Brown, V. Ciaramitaro, M. Jazayeri, and B. Lau for comments on the manuscript. This project was funded by the Human Frontiers Science Program (M.C.D.) and the McKnight Foundation and Klingenstein Fund (P.W.G.).

Received: February 2, 2004
Revised: July 22, 2004
Accepted: September 2, 2004
Published: October 13, 2004

References

Andersen, R.A., and Buneo, C.A. (2002). Intentional maps in the posterior parietal cortex. Annu. Rev. Neurosci. *25*, 189–220.

Andersen, R.A., Asanuma, C., and Cowan, W.M. (1985). Callosal and prefrontal associational projecting cell populations in area 7A of the macaque monkey: a study using retrogradely transported fluorescent dyes. J. Comp. Neurol. *22*, 443–455.

Andersen, R.A., Asanuma, C., Essick, G., and Siegel, R.M. (1990). Corticocortical connections of anatomically and physiologically defined subdivisions within the inferior parietal lobule. J. Comp. Neurol. *296*, 65–113.

Andersen, R.A., Snyder, L.H., Bradley, D.C., and Xing, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. Annu. Rev. Neurosci. *20*, 303–330.

Barlow, H.B. (1961). The coding of sensory messages. In Current Problems in Animal Behavior, W.H. Thorpe and O.L. Zangwill, eds. (Cambridge: Cambridge University Press), pp. 331–360.

Barraclough, D.J., Conroy, M.L., and Lee, D. (2004). Prefrontal cortex and decision making in a mixed strategy game. Nat. Neurosci. *7*, 404–410.

Basso, M.A., and Wurtz, R.H. (1998). Modulation of neuronal activity in superior colliculus by changes in target probability. J. Neurosci. *18*, 7519–7534.

Breiter, H.C., Aharon, I., Kahneman, D., Dale, A., and Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. Neuron *30*, 619–639.

Camerer, C.F. (2003). Behavioral Game Theory (Princeton, NJ: Princeton University Press).

Coe, B., Tomihara, K., Matsuzawa, M., and Hikosaka, O. (2002). Visual and anticipatory bias in three cortical eye fields of the monkey during an adaptive decision-making task. J. Neurosci. *22*, 5081–5090.

Colby, C.L., and Goldberg, M.E. (1999). Space and attention in parietal cortex. Annu. Rev. Neurosci. *22*, 319–349.

Dean, A.F. (1981). The variability of discharge of simple cells in the cat striate cortex. Exp. Brain Res. *44*, 437–440.

Fudenberg, D., and Tirole, J. (1994). Game Theory (Cambridge, MA: MIT Press).

Glimcher, P.W. (2003a). The neurobiology of visual saccadic decision making. Annu. Rev. Neurosci. *26*, 133–179.

Glimcher, P.W. (2003b) Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics (Cambridge, MA: MIT Press).

Glimcher, P.W. (2005). Indeterminacy in brain and behavior. Annu. Rev. Psychol. *56*, in press.

Glimcher, P.W., and Dorris, M.C. (2004). Neuronal studies of decision making in the visual-saccadic system. In The Cognitive Neurosciences III, M.S. Gazzaniga, ed. (Cambridge, MA: MIT Press), in press.

Glimcher, P.W., Ciaramitaro, V.M., Platt, M.L., Bayer, H.M., Brown, M.A., and Handel, A. (2001). Application of neurosonography to experimental physiology. J. Neurosci. Methods *108*, 131–144.

Gnadt, J.W., and Andersen, R.A. (1988). Memory related motor planning activity in posterior parietal cortex of macaque. Exp. Brain Res. *70*, 216–220.

Gold, J.I., and Shadlen, M.N. (2000). Representation of a perceptual decision in developing oculomotor commands. Nature *404*, 390–394.

Gold, J.I., and Shadlen, M.N. (2001). Neural computations that underlie decisions about sensory stimuli. Trends Cogn. Sci. *5*, 10–16.

Heeger, D.J., Simoncelli, E.P., and Movshon, J.A. (1996). Computational models of cortical visual processing. Proc. Natl. Acad. Sci. USA *93*, 623–627.

Herrnstein, R.J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. J. Exp. Anal. Behav. *7*, 179–182.

Herrnstein, R.J. (1997). The Matching Law: Papers in Psychology and Economics (Cambridge, MA: Harvard University Press).

Hikosaka, O., Takikawa, Y., and Kawagoe, R. (2000). Role of the basal ganglia in the control of purposive saccadic eye movements. Physiol. Rev. *80*, 953–978.

Kagel, J.H., and Roth, A.E. (eds.) (1995). The Handbook of Experimental Economics (Princeton, NJ: Princeton University Press).

Kahneman, D., and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. Econometrica *47*, 263–291.

Krebs, J.R., and Davies, N.B. (1996). Behavioural Ecology, Fourth Edition (Cambridge: Blackwell Publishers).

Kreps, D.M. (1990). A Course in Microeconomic Theory—First Edition (Princeton, NJ: Princeton University Press).

Nash, J.F. (1950). Equilibrium points in *n*-person games. Proc. Natl. Acad. Sci. USA *36*, 48–49.

Nash, J.F. (1951). Non-cooperative games. Annals of Mathematics *54*, 286–295.

Paré, M., and Wurtz, R.H. (2001). Progression in neuronal processing for saccadic eye movements from parietal cortex area lip to superior colliculus. J. Neurophysiol. *85*, 2545–2562.

Platt, M.L., and Glimcher, P.W. (1997). Response of intraparietal neurons to saccadic targets and visual distractors. J. Neurophysiol. *78*, 1574–1589.

Platt, M.L., and Glimcher, P.W. (1999). Neural correlates of decision variables in parietal cortex. Nature *400*, 233–238.

Platt, M.L., Lau, B., and Glimcher, P.W. (2003). Situating the superior colliculus within the gaze control network. In The Oculomotor System: New Approaches for Studying Sensorimotor Integration, W.C. Hall and A.K. Moschovakis, eds. (Boca Raton, FL: CRC Press), pp. 1–34.

Schall, J.D., and Thompson, K.G. (1999). Neural selection and control of visually guided eye movements. Annu. Rev. Neurosci. *22*, 241–259.

Shadlen, M.N., and Newsome, W.T. (1996). Motion perception: seeing and deciding. Proc. Natl. Acad. Sci. USA *93*, 628–633.

Shadlen, M.N., and Newsome, W.T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. J. Neurophysiol. *86*, 1916–1936.

Shima, K., and Tanji, J. (1998). Role for cingulated motor area cells in voluntary movement selection based on reward. Science *282*, 1335–1338.

Stephens, D.W., and Krebs, J.R. (1986). Foraging Theory (Princeton, NJ: Princeton University Press).

Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2004). Matching behavior and the representation of value in the parietal cortex. Science *304*, 1782–1787.

Sutton, R.S., and Barto, A.G. (1998). Reinforcement Learning (Cambridge, MA: MIT Press).

Tolhurst, D.J., Movshon, J.A., and Thompson, I.D. (1981). The dependence of response amplitude and variance of cat visual cortical neurones on stimulus contrast. Exp. Brain Res. *41*, 414–419.

Von Neumann, J.V., and Morgenstern, O. (1944). Theory of Games and Economic Behaviour (Princeton, NJ: Princeton University Press).