Parietal neurons encode expected gains in instrumental information

Nicholas C. Foley[a,1], Simon P. Kelly[a,1], Himanshu Mhatre[a], Manuel Lopes[b], and Jacqueline Gottlieb[a,c,d,2]

[a]Department of Neuroscience, Columbia University, New York, NY 10032; [b]INRIA, 33405 Talence, France; [c]The Kavli Institute for Brain Science, Columbia University, New York, NY 10032; and [d]The Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY 10027

Edited by Thomas D. Albright, The Salk Institute for Biological Studies, La Jolla, CA, and approved March 7, 2017 (received for review August 26, 2016)

In natural behavior, animals have access to multiple sources of information, but only a few of these sources are relevant for learning and actions. Beyond choosing an appropriate action, making good decisions entails the ability to choose the relevant information, but fundamental questions remain about the brain's information sampling policies. Recent studies described the neural correlates of seeking information about a reward, but it remains unknown whether, and how, neurons encode choices of instrumental information, in contexts in which the information guides subsequent actions. Here we show that parietal cortical neurons involved in oculomotor decisions encode, before an information sampling saccade, the reduction in uncertainty that the saccade is expected to bring for a subsequent action. These responses were distinct from the neurons' visual and saccadic modulations and from signals of expected reward or reward prediction errors. Therefore, even in an instrumental context when information and reward gains are closely correlated, individual cells encode decision variables that are based on informational factors and can guide the active sampling of action-relevant cues.

information sampling | saccades | attention | decisions | reward

In natural behavior, animals have access to multiple sources of information, but few of these sources are relevant for learning or action. Making good decisions therefore entails not only the selection of the ultimate action but, more primarily, the decision of which source of information to sample. Decisions about information sampling are central for tasks as diverse as making a medical diagnosis (which is the best test to prescribe?), making categorization decisions (which is the most informative feature?) (1, 2), and guiding skilled actions (what should I keep my eyes on while driving?). Despite the ubiquity and significance of active sampling mechanisms, few studies have been devoted to understanding these mechanisms and their importance for decision theories. Evidence accumulation has been extensively examined in decision research (3) but has been portrayed as a passive process, in the sense that decision makers rely on predetermined (experimenter selected) sources of information but cannot themselves determine which source to consult to guide a future action.

Recent studies begin to shed light on this question by showing that animals (including pigeons, monkeys, and humans) prefer to observe cues that are predictive rather than nonpredictive about a future reward, and that the value of informative cues is encoded in the orbital frontal cortex and midbrain dopamine (DA) cells (4, 5). These investigations, however, have been limited to noninstrumental contexts in which animals seek to obtain information about a reward merely in order to know, but cannot act based on the information. Very little is known about the much more common scenario in which animals sample instrumental information to make decisions and guide future actions (6).

Understanding instrumental sampling poses two important challenges beyond those arising in noninstrumental contexts. First, because instrumental sampling is, by definition, coordinated with actions, understanding it requires developing sequential paradigms in which an individual first decides which information to sample, and then decides which action to take based on that information. Sequential tasks of this kind have been used in human observers (e.g., refs. 2 and 7–10) but are largely eschewed in work with experimental animals where decision-making is studied with single-step tasks (11) (see ref. 12 for a notable exception).

Second, in an instrumental context, a more reliable cue, by definition, supports a better choice of actions, strongly correlating gains in reward with gains in information. Behavioral and computational studies suggest that a full explanation of instrumental sampling behaviors requires considering not only rewards but also bona fide informational factors (2, 7–9, 13). However, we lack behavioral paradigms that can be used with animals and can clearly dissociate neuronal signals of instrumental information from those of expected reward gains, or furnish an understanding of neural signals related to these factors.

Here we address this question by examining the activity of target selective cells in the monkey lateral intraparietal area (LIP), a cortical area implicated in the top-down control of attention and gaze (3, 14). LIP neurons are particularly appropriate for this investigation, because they integrate multiple factors relevant for deciding where to allocate attention and gaze, and are sensitive to the expected reward values of alternative gaze locations (11, 15–22).

To determine whether LIP target-selective cells are also sensitive to expected gains in information, we used a two-step decision task in which the monkeys made a first rapid eye movement (saccade) to sample information from a visual cue, and a second saccade to select an action based on that information. We focus on the responses before the first information sampling saccade, and show that they encode the gains in information that the saccade was expected to bring, in a manner that is independent of the cumulative future rewards and reward prediction errors (RPEs) associated with the informative cues. Thus, parietal cortical neurons involved in active sensing decisions encode bona fide information-based decision variables that can guide the judicious sampling of action-relevant cues.

Significance

We examine how the brain guides active sensing in awake, behaving primates using a paradigm in which information sampling is dissociated from reinforcement variables, such as cumulative future reward or reward prediction errors. We show that target selective cells in lateral intraparietal cortex encode decision variables based on expected gains in instrumental information—the extent to which a visual cue, when discriminated, is expected to reduce the uncertainty of a subsequent action.

Author contributions: S.P.K. and J.G. designed research; N.C.F., S.P.K., H.M., and M.L. performed research; N.C.F. and S.P.K. analyzed data; and J.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]N.C.F. and S.P.K. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: jg2141@columbia.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1613844114/-/DCSupplemental.

Results

Two monkeys (*Macaca mulatta*) were trained on a task in which they made two saccades on each trial—a first saccade to obtain information from a visual cue, and a second saccade to report a decision based on that information (Fig. 1*A*). In each trial, after achieving central fixation, the monkeys had a 500-ms delay period during which they viewed a visual display containing two cues (round apertures containing small dots) and two targets (white squares; Fig. 1*A*, first and second panels). At the end of this period, the fixation point disappeared, and the monkeys had to make a saccade to one of the cues. When the monkey's gaze arrived on the cue, the dots inside the chosen cue began to move with 100% coherence toward one of the targets, indicating which target was most likely to deliver a reward (Fig. 1*A*, fourth panel, black arrows). After motion onset, the monkeys were free to indicate their final decision by making a second saccade to a target, and the trial ended with a reward or lack of reward (see below).

Our interest was in the neuronal activity related to the decision for the first saccade, which was made during the initial period of central fixation, and expressed the monkeys' choice of which cue to sample (Fig. 1*A*, delay, second panel). Note that, during this delay period, the cues did not yet deliver motion information, as the dots inside both apertures were stationary. However, the monkeys were informed about cue validity by means of a colored border surrounding each cue. Validity was defined as the probability that the motion delivered by a cue would correctly specify the rewarded final action, and could take values of 100%, 80%, or 55% (with the color-validity mapping held constant for each monkey but randomized across monkeys; *Methods*). Therefore, validity is a measure of how informative each cue was likely to be for the choice of the subsequent action, and is mathematically equivalent to two widely used measures of expected information gains (EIG)—Shannon information and the maximum a posteriori (MAP) estimate (Eqs. **S1** and **S2**)—and we use "validity" and "EIG" interchangeably in the paper. The task design, therefore, separated neural activity related to the choice of an informative item (based on peripheral validity information before the first saccade) from activity associated with the sensory discrimination (based on foveal motion information after the saccade).

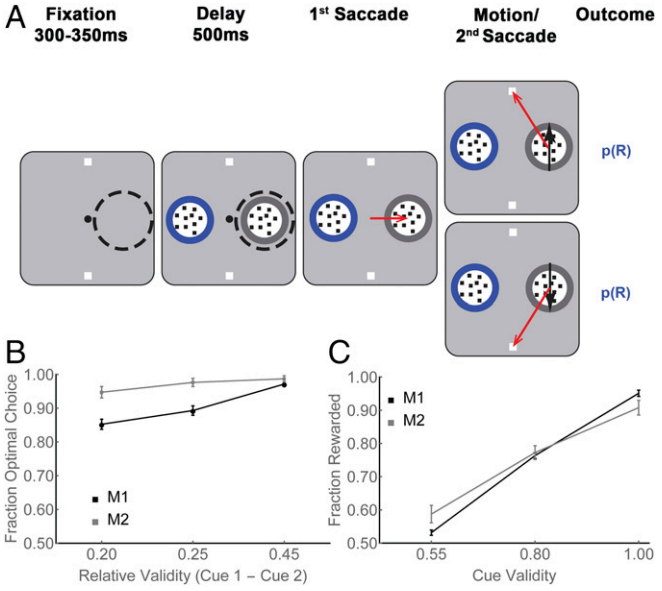

Fig. 1. Free-choice task and behavior. (*A*) Each trial began when the monkeys achieved fixation of a central spot (small black circle) placing the RF of an LIP cell (dashed circle) on an eccentric screen location. (A representative RF in the right hemifield is shown for illustrative purposes, but it was not visible to the monkey during the experiment.) Two targets were then presented outside the RF (white squares) followed by two cues, of which, one was inside the RF and the other was at the diametrically opposite location (round apertures containing small dots). The monkeys viewed the display for a 500-ms delay period, after which the fixation point disappeared, and the monkeys made a first saccade to one of the cues (freely chosen; third panel, red arrow). At the end of this saccade, the chosen cue delivered its information in the form of 100% coherent dot motion directed toward one of the targets (fourth panel, black arrows). After motion onset, the monkeys were free to indicate their final decision by making a second saccade to a target (fourth panel, red arrows), and the trial ended with a probabilistic reward [p(R)]. (*B*) The fraction of choices of the optimal cue at the first saccade step, as a function of the difference in validity of the available cues, for M1 and M2. The *x* axis shows the difference in validity for the three possible pairs of unequal validity cues, and each point shows the mean and SEM across all recording sessions. (*C*) The rates of reward as a function of the validity of the chosen cue, in the same format as *B*.

When presented with free-choice trials containing two unequal validity cues, both monkeys chose to inspect the higher-validity cue with >90% probability, suggesting that they used a validity-based sampling policy (each monkey, $P < 0.0001$ relative to 0.5; Fig. 1*B*). After selecting a cue, the monkeys tended to obey its instruction and choose the target that was congruent with the motion direction. Consequently, the reward rates approached the maximum expected rates in the task and, as expected in an instrumental context, scaled with the validity of the chosen cue (Fig. 1*C*, two-way ANOVA, $P < 10^{-9}$ for the main effect of validity, $P < 10^{-8}$ for monkey, and $P = 0.004$ for monkey × validity interaction).

LIP Neurons Encode EIG. To examine the encoding of cue validity at the level of individual cells, we identified individual target-selective LIP neurons (*Methods*) and customized the display so that, during the initial delay period, one of the cues fell inside the receptive field (RF) of the recorded cell while the other stimuli were outside the RF, allowing us to capture activity related to the first saccade decision. Note that, after the delay period, the RF of the cell moved away from the visual display by virtue of the first saccade, and we could no longer observe meaningful responses to the second saccade or motion discrimination.

During the delay period before the first saccade, LIP neurons ($n = 50$) had robust visual responses to the onset of the cues, followed by sustained saccade-related responses that were higher if the saccade was directed inside versus opposite the RF (Fig. 2*A*; two traces in each panel). However, the directional response—the difference in firing for the two directions—was not constant but scaled as a function of the relative validity of the available cues (Fig. 2*A*, compare across the three panels), suggesting that the neurons multiplex visual and saccade-related responses with a sensitivity to validity/EIG.

To quantitatively characterize the EIG response, we regressed each neuron's firing rates as a function of the validity of the RF cue, the validity of the opposite RF cue, saccade direction, saccade latency, saccade velocity, and saccade accuracy (Eq. **S3**). The validity of the RF cue had a prominent effect in the early part of the delay period (Fig. 2 *B* and *C*, 100 ms to 300 ms after cue onset), with an average coefficient of 1.51 ± 0.2 ($P < 0.0001$, corresponding to 12.08 spikes per second (sp/s) in terms of raw firing rates). During this interval, the validity of the RF cue had a much stronger effect than the validity of the opposite cue (orange; -0.36 ± 0.11, $P = 10^{-4}$) or saccade direction (green; 0.28 ± 0.17, $P = 10^{-6}$) and a much higher prevalence among individual cells (with 72% showing significant encoding of the validity of the RF cue, compared with only 16% for the opposite RF cue, 10% for saccade direction, and 0%, 0%, and 2%, respectively, for saccade velocity, latency, and accuracy; Fig. 2*C*, *Left*). During the later delay period (300 ms to 500 ms), the average effect of saccade direction increased and became comparable to that of cue validity (direction, 1.08 ± 0.24 vs. validity, 0.85 ± 0.13, $P = 0.68$), consistent with the fact that LIP neurons reflect oculomotor planning late in a

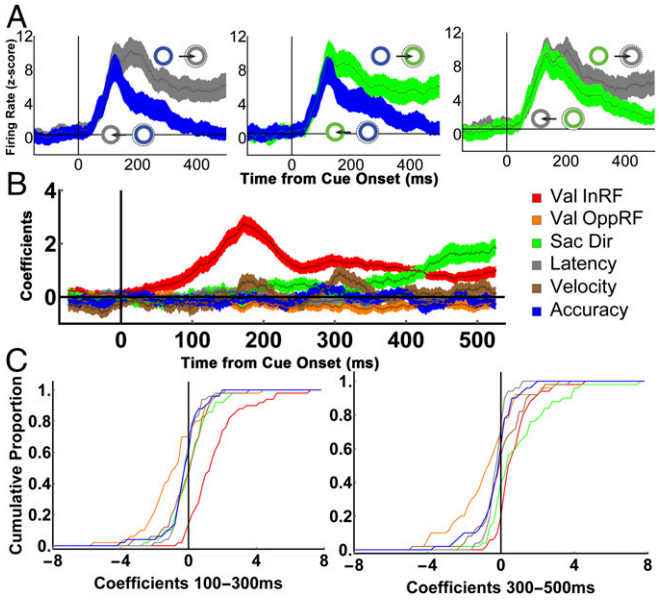

Fig. 2. Quantitative analysis of EIG effects on two-cue trials. (*A*) Population responses on two-cue trials in which the alternative cues had (*Left*) large, (*Center*) medium, and (*Right*) low differences in validity, sorted according to saccade direction (toward or opposite the RF). Throughout the paper, we use gray, green, and blue to represent, respectively, 100%, 80%, and 55% valid cues (although, in practice, the colors differed by monkey). Traces show mean and SEM of activity across all of the cells ($n = 50$). Firing rates were *z*-scored for each cell using the mean and SD across all correct trials (*Methods*), and can be compared across all trial groups. (*B*) Time-resolved regression coefficients (sliding window of 50-ms width, 1-ms step) estimating the effects of the validity of the RF cue (Val InRF), the validity of the opposite RF cue (Val OppRF), and saccade direction (Sac Dir), velocity, latency, and accuracy across the trials shown in *A*. (*C*) Cumulative distribution of coefficients for individual cells in 200-ms bins spanning the early and late parts of the delay period.

decision epoch (3). However, the validity response remained highly prevalent, with 84% of cells showing significant sensitivity to the validity of the RF cue, compared with 48% for saccade direction.

Further analysis verified that EIG modulations were robust in a subset of the two-cue trials selected so as to differ only in the validity of the RF cue but be matched for saccade direction and the validity of the opposite cue (Fig. S1). Finally, model comparisons using Akaike and Bayesian Information Criteria (23) showed that, among the set of all possible models based on the regressors of interest, models that contained a term for the validity of the RF cue provided a superior fit (Fig. S2). Together, these findings suggest that LIP neurons robustly encoded the EIG of an RF cue in combination with, but independently of, their previously described responses to visual onsets, saccade planning, and competitive normalization by non-RF cues.

EIG Is Encoded Independently of Expected Reward. Behavioral evidence suggests that informational factors influence gaze allocation during sensorimotor tasks, supporting the idea that the validity responses we find encode EIG (24). However, an alternative possibility is that these responses reflect reward expectation, because validity in an instrumental context is, by definition, correlated with the reward probability of subsequent actions (Fig. 1*C*). To distinguish between these possibilities, we compared the responses of a sample of LIP cells ($n = 69$, including the 50 cells described in Fig. 2) to two types of RF stimuli that were equated for reward expectation and differed in whether or not they brought decision information.

We tested the neurons in two modified versions of the original task, in which the stimulus appearing in the RF was, respectively, informative or uninformative for the final action. The informative condition was identical to the original two-cue task, with the exception that a single informative cue appeared inside the RF, and monkeys were forced to sample the information conveyed by this cue (Fig. 3*A*; forced choice, informative). The "uninformative" condition was equivalent in all respects, except that the monkeys received action and reward information from a precue, and the RF cue delivered no additional information (Fig. 3*B*). Each trial in this condition began with the presentation of a precue that was located opposite the RF, had a colored border indicating validity, and contained coherent motion toward one of the targets, thus signaling the trial's final target and reward probability (Fig. 3*B*, first panel). After the precue was extinguished, a second uninformative stimulus appeared in the RF, and, after a 500-ms delay, the monkeys made a first saccade to this stimulus, viewed random (0% coherence) motion inside it, and made a second saccade to a target.

In the uninformative condition, therefore, the first saccade to the RF cue was equated to that in the informative task in its direction, timing, and reward expectation, and differed only in that it was not expected to bring decision information. The equivalence in reward expectation was ensured by the fact that, in the uninformative condition, the saccade to the RF item was necessary to harvest the reward, and reward probability was governed by the validity of the initial precue. To ensure that reward associations were consistent over longer time scales, we consistently paired each type of uninformative item with a specific informative precue (e.g., a red border uninformative item always followed a 55% valid precue and hence had a ∼55% long-term reward expectancy, whereas a cyan border uninformative item always followed an 80% valid precue, and thus had an ∼80% long-term expectancy; Fig. 3*B*). (Note that, although monkey 1 (M1) performed the task with all three pairs of yoked informative–uninformative cues, monkey 2 (M2) was limited to pairs with a 55% and 80% precue (*Methods*). We therefore focus our analysis on these two trial types, which were tested in both monkeys (and separately analyze the three-cue data from monkey M1).

To confirm that the informative and uninformative items had equivalent reward value, we simulated a reinforcement learning (RL) model that used the same state transitions and reward contingencies as the behavioral task, and learnt the values of each state by trial and error using a temporal difference algorithm (see *SI Methods*). As shown in Fig. 3*C* (*Left*), the model predicts that the value of the first saccade will depend on cue validity and that this dependence will be identical for informative and uninformative items (Fig. 3*C*, two-way ANOVA, $P < 10^{-10}$ for cue validity, $P > 0.8$ for task type and interaction). This prediction follows from the fundamental structure of model-free RL algorithms, in which the final reward of an action sequence confers value to all previous steps, subject to temporal discounting but impervious to informational factors (25, 26), and confirms that our task design appropriately equated the reward values of informative and uninformative items.

Contrary to model predictions, LIP neurons discriminated only the validity of informative cues but not the reward associations of uninformative items (Fig. 3*C*, "Data"). A two-way ANOVA revealed significant effects of validity ($P < 0.001$) and validity × task interaction ($P < 0.001$) and, in post hoc comparisons, a significant effect of validity for informative cues but not for uninformative items ($P < 0.001$ vs. $P = 0.6$; average responses during the period of peak modulation, 125 ms to 250 ms after cue onset).

A time-resolved regression analysis (Fig. 4 *B* and *C* and Eq. S4) yielded average regression coefficients 3.4 ± 0.57 *z*-score (SD) units for informative cues vs. 0.81 ± 0.39 *z*-score units for uninformative stimuli ($P < 0.0003$), a result that was robust for each monkey (M1, 3.4 ± 0.65 vs. 0.75 ± 0.44, $P = 0.0013$; M2, 3.57 ± 0.9 vs. 1.2 ± 0.6, $P = 0.048$). This difference was consistent at the level of individual cells, where more than twice as many neurons showed significant effects of validity than showed encoding of

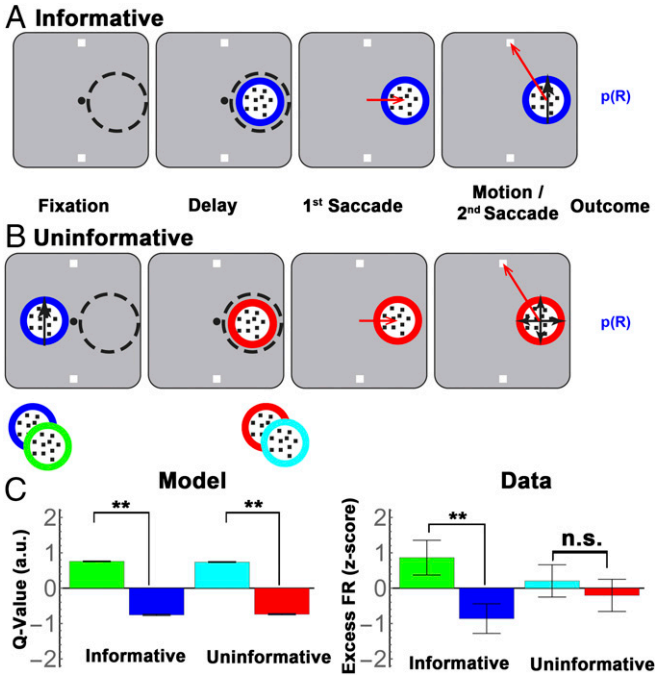

Fig. 3. The informative/uninformative stimulus test. (*A*) Trial stages in the informative and uninformative task. The Informative task was identical to the cue choice task except that a single cue appeared in the RF, forcing the monkeys to complete the trial based on this cue. (*B*) In the uninformative condition, a precue containing moving dots appeared opposite the RF simultaneous with target onset, and conveyed both the reward probability of the trial (by virtue of its colored border) and the instruction about the final action (through the dot motion; first panel). The precue then disappeared and was replaced by an uninformative stimulus inside the RF (second panel). After an additional 500-ms delay period, the monkeys were required to make a saccade to the RF stimulus (third panel) before making their final saccade to a target (fourth panel). Note that, although the uninformative stimulus delivered no information (but only random, 0% coherence motion), a saccade to this stimulus was still valuable because it was necessary to obtain the reward. As depicted by the cartoon at the bottom of *B*, we used two types of uninformative stimuli, which were distinguished by the color of their borders and were each paired with a constant informative precue to establish long-term reward associations. Throughout the paper, we depict the uninformative stimulus that followed a 55% valid precue in red and the stimulus that followed an 80% valid precue in cyan (although in practice the colors were randomized for each monkey). (*C*) (*Left*) A standard model-free RL simulation (*SI Methods*) predicts that cue value scales as a function of validity, and that this scaling is identical for the uninformative stimuli. Note that the model predictions refer to the relative modulations across the different cues, but the absolute *Q* values are arbitrary. (*Right*) LIP responses to the informative and uninformative items (mean and SEM of the traces shown in Fig. 4*A*, averaged across 125 ms to 250 ms after cue onset). (**$P < 0.001$; two-way ANOVA with pairwise post hoc comparisons; n.s., $P = 0.6$). Although the model predicts identical value scaling for informative and uninformative items, LIP neurons show much stronger scaling for the informative cues.

reward expectation (32/69 vs. 15/69 cells; $P < 0.002$, $z = 3.05$ test of proportions; Fig. 4*B*, colormaps, and Fig. 4*C*). We further verified that the sensitivity to cue type in the informative condition was independent of saccade metrics (Fig. S3), was replicated when reanalyzed in terms of the Shannon EIG (SEIG) (Fig. S3), and was replicated in M1 using the three yoked pairs of informative and uninformative cues (Fig. S4).

Note that, because the theoretical predictions pertain to the relative rather than absolute response magnitudes, our analysis so far has focused on the mean-subtracted neuronal activity, which factors out the visual onset response and reveals the relative responses to the different levels of validity/reward expectation (Fig. 4*A*). Examination of the full (not mean-subtracted) responses showed that the validity modulations rode on top of a visual onset response that was strong for both informative and uninformative cues and could be factored out by linear regression (Fig. S5*A* and Eq. **S4**). Finally, analysis of the raw (not *z*-scored) responses showed that the effects were very robust, with average regression coefficients of 24.15 ± 0.57 sp/s for informative cues vs. 4.68 ± 1.2 sp/s for uninformative stimuli ($P = 0.000008$; Fig. S6 *A* and *B*).

Although the above results suggest that LIP neurons encode EIG, it is important to rule out confounds based on spurious differences between the informative and uninformative tasks. An important concern is that, although our tasks nominally equated reward expectations, variations in the monkeys' performance may have influenced the rewards that were de facto experienced for the different cues. The latencies of the monkeys' saccades to uninformative items scaled inversely as a function of expected reward ($P < 0.01$ in each monkey), confirming that the monkeys were sensitive to expected reward even when making a saccade to an uninformative item. However, the reward rate differences between the 55% and 80% cues tended to be slightly higher for informative cues relative to uninformative items (across all sessions—M1: informative, 53.1 ± 0.7% vs. 76.3 ± 0.08%; uninformative: 54.7 ± 1.1% vs. 73.6 ± 1.0%, two-way ANOVA, $P < 10^{-59}$ for validity, $P = 0.53$ for condition and $P = 0.008$ for interaction; M2 informative, 58.1 ± 2.6 vs. 77.3 ± 2.2%; uninformative, 51.6 ± 2.0% vs. 61.1 ± 1.5%, two-way ANOVA, $P < 10^{-4}$ for validity, $P < 10^{-2}$ for condition and $P = 0.045$ for interaction). Two observations show that this performance difference could not explain the neuronal results. First, as documented above (Fig. 4*B*), the neuronal validity effects were similar in the two monkeys despite their different performance. Second, we found no correlation between the experienced reward differential and the EIG modulations in individual cells in either monkey for informative cues (Spearman rank coefficient—M1, $r = -0.1$, $P = 0.44$; M2, $r = 0.16$, $P = 0.68$) and, most importantly, for uninformative items (Fig. 5*A*; Spearman rank coefficient—entire sample $r = -0.04$, $P = 0.76$; M1, $r = -0.07$, $P = 0.61$; M2, $r = 0.34$, $P = 0.39$). Finally, additional regression analyses that incorporated terms for trial by trial reward history (Eqs. **S5** and **S6**) showed that the effects of prior trial rewards were minimal and limited to the baseline period before cue onset (Fig. 5*B*), ruling out artifacts related to the experienced reward rates or reward history.

An additional concern is that variations in the monkeys' postsaccadic motion viewing times may have affected trial length and hence temporal discounting. However, viewing times were considerably longer for informative relative to uninformative cues in both monkeys (Fig. 5*C*; each monkey, $P < 10^{-28}$), a difference that is consistent with the informativeness of the different cues but contrary to an explanation in terms of temporal discounting (according to which we should see weaker neural modulations for informative cues, due to the greater discounting associated with these cues). On a trial by trial basis, presaccadic firing rates were not sensitive to postsaccadic viewing times, further arguing against confounds related to the postsaccadic viewing or motion discrimination (Fig. 5*D* and Eq. **S7**).

Finally, we considered potential effects of task geometry, related to the fact that, in the uninformative condition, attention had to be reoriented toward the RF after having been engaged by a precue at the opposite location (Fig. 3*B*). As shown in Fig. S5*A*, reorienting was associated with a slight enhancement of the visual response in our sample of cells (compare gray traces for informative and uninformative items), consistent with previous results (27), but the magnitude of this enhancement was not correlated with the neurons' validity/reward modulations (Fig. S5*C*, black dots).

Together, these findings rule out spurious explanations related to the complexities of double-step tasks, and suggest that LIP neurons encode EIG independently of reward expectations.

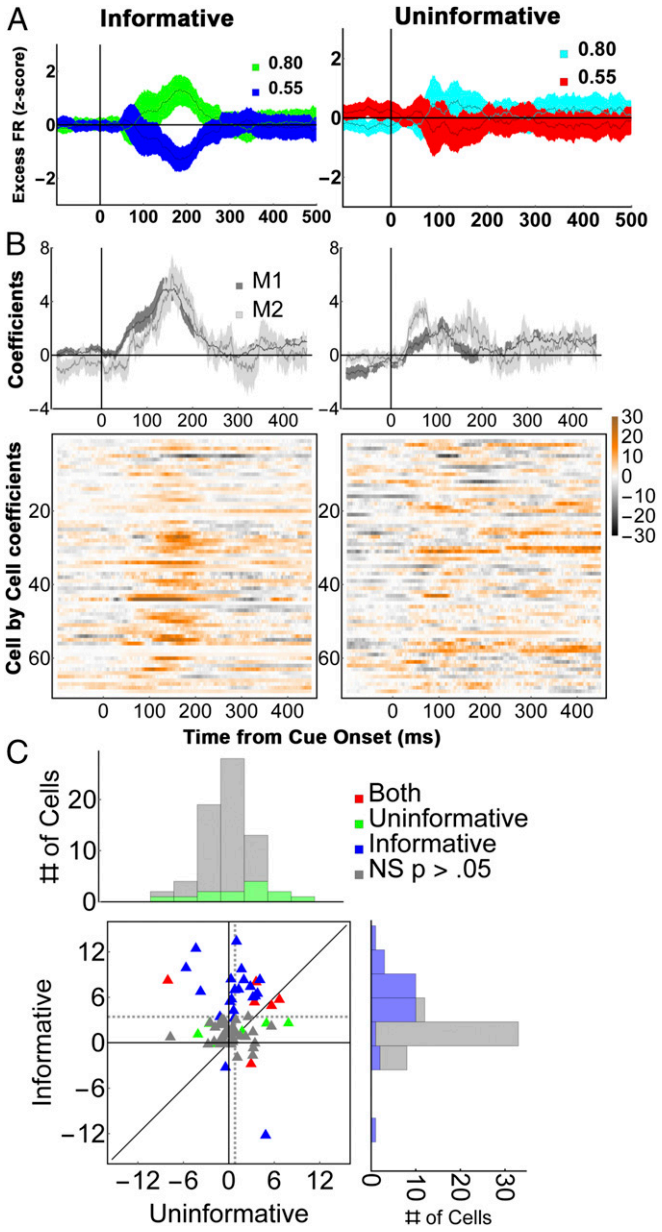

Fig. 4. LIP neurons encode validity but not the cumulative future rewards of uninformative cues. (*A*) (*Left*) Average firing rates ($n = 69$ cells) for 55% and 80% valid cues, and (*Right*) their yoked uninformative stimuli. To highlight the cue-related modulation, firing rates were *z*-scored after subtracting the average activity for each stimulus class (we use the term "Excess" to indicate mean subtraction). Error bars show SEM across cells. (*B*) Average regression coefficients for the validity/reward responses in (*Left*) informative and (*Right*) uninformative trials for (*Top*) each monkey and (*Bottom*) each cell (colormaps). Note that the regression coefficients estimate the size of the neural effects across the entire validity range (50 to 100%) and are thus nearly twice as large as the difference in responses between the 80% and 55% cues, which span only half of this range. (*C*) Cell-by-cell comparison of validity and reward effects. Each point shows the average validity/reward coefficients of one cell (125 ms to 250 ms), color coded according to its significance along the *x* and *y* axes. In the marginal histograms, significant points are indicated by darker colors. Dotted lines show sample means. Note that all of the cells were used to compute the marginal histograms and the means indicated by the dotted lines, but one outlier that had coefficients of 25.2 for informative cues and 11.4 for uninformative cues was left out of the plot for clarity of presentation (this cell came from M1 and can be seen on row 44 of *B* as showing very high modulation for informative cues in analysis epoch). Recomputing the statistics without this outlier did not change the results (both monkeys, average and SEM for informative vs. uninformative cues: 3.1 ± 0.48 vs. 0.65 ± 0.36, $P = 0.00035$; M1, 3.0 ± 0.54 vs. 0.57 ± 0.41, $P = 0.0015$; both monkeys, *z* test of proportions for the incidence of significant cells, $z = 3.09$, $P = 0.001$).

EIG Is Distinct from Reward Prediction Errors. Although, in *EIG Is Encoded Independently*, we considered an explanation in terms of cumulative future rewards—the payoffs that an agent can expect to receive after taking an action—an additional question is whether our findings can be explained in terms of RPE, defined as a change in reward expectation relative to a prior state (28). Even though previous investigations have not tested whether LIP neurons are sensitive to RPEs, such a sensitivity could potentially explain the neurons' lack of modulation for uninformative cues. In the informative condition, the monkeys began each trial with a prior reward expectation of ∼78% (based on the average reward rates across the different validity cues), and the appearance of an informative cue signaled an increase (for 100% validity), no change (80% validity), or decrease (55% validity) in reward probability relative to this prior expectation. However, the appearance of an uninformative item did not alter prior beliefs (Fig. 3*B*), potentially explaining the lack of neuronal modulation. To examine whether the cells respond to RPE, we conducted an additional cue-change test in which a first informative cue established the monkey's initial reward expectations and a second cue modified these expectations, producing RPEs that were independent of cue validity.

The majority of trials in the cue-change task were identical to the one-cue condition of the standard task (Fig. 3*A*), in that the monkeys received a single cue opposite the cell's RF and completed the trial based on the information they sampled in this cue. In the remaining, critical 25% of trials, the initial cue disappeared before the first saccade and was replaced with another informative cue inside the RF; the monkeys viewed the second cue for an additional 500 ms during central fixation, and completed the trial by harvesting information from this cue (Fig. 6*A*). In these cue-change trials, therefore, monkeys would form a reward expectation based on the validity of the initial cue, and change these expectations, producing an RPE, based on the second cue. Consider, for instance, the subset of trials in which the initial cue had 100%, 80%, or 55% validity, and was followed by a second cue of 80% validity (Fig. 6*A*). Even though this second cue had constant validity, it signaled different RPEs according to the validity of the initial cue (i.e., –20%, 0%, and 25% after initial cues of, respectively, 100%, 0%, and 25% validity). RL model simulations (*SI Methods*) verified that the RPEs in the cue-change task had equivalent magnitudes to those associated with the informative cues (Fig. 6*B*). Therefore, if LIP neurons encoded RPE, they should show significant modulation in response to the 80% cue, which should be equivalent to the response modulations across the informative cues.

Contrary to this prediction, the neurons modulated much more strongly as a function of validity than as a function of RPE. Focusing again on the mean-subtracted firing rates to remove the common visual response (Fig. 7*A*), we found a significant effect of validity but not RPE (Fig. 6*B*, "Data," two-way ANOVA, $P < 0.0004$ for validity and validity × task interaction; post hoc comparisons, $P < 0.001$ for validity, $P = 0.87$ for RPE). Regression analysis (Eq. **S8**) confirmed that the cells had much stronger modulations according to validity than according to RPE in each monkey [Fig. 7*B*, *Top*; average coefficients 125 ms to 250 ms after cue onset were 3.5 ± 0.57 for informative cues vs. 0.74 ± 0.3 for RPE in the entire sample ($n = 24$, $P < 0.0001$); M1, 3.67 ± 0.7 vs. 0.91 ± 0.4, $P < 0.002$, $n = 18$; M2, 2.97 ± 0.9 vs. 0.23 ± 0.37; $P < 0.024$, $n = 6$]. Significant effects of validity and RPE were found in, respectively, 15/24 vs. 5/24 cells (62.5% vs. 21%; *z*-test of proportions, $z = 2.3$, $P = 0.01$).

Control analyses ruled out explanations based on spurious factors. Reaction times for the first saccade showed a significant reduction as a function of RPE (one-way ANOVA, $P = 0.0001$ in

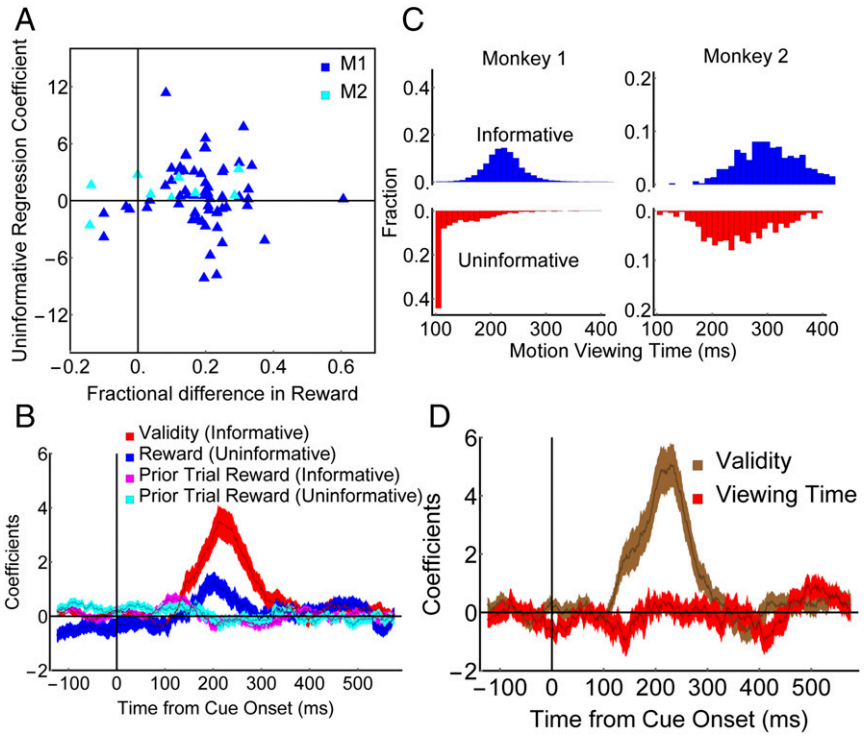

Fig. 5. Selective encoding of validity cannot be explained by task-related confounds. (*A*) Reward differentials. Each point represents one neuron. The *y* axis shows the reward coefficient on uninformative trials, and the *x* axis shows the difference in the fractional rewards obtained for the 80% versus the 55% uninformative stimuli during the recording of that neuron. The extent to which different neurons modulate for uninformative cues is not correlated with the difference in the average rewards experienced for those cues across or within monkeys. (*B*) Regression coefficients (mean and SEM) dissociating the effects of validity/reward and prior trial rewards (Eq. S5). During the 125- to 250-ms epoch of peak validity modulation, fewer than 10% of cells were significantly modulated by the prior trial reward (seven and four, respectively, in the informative and uninformative conditions); this result was confirmed if we restricted the history regressor to only include trials of the same cue type (Eq. S6) or added terms to capture the effects of up to three previous trials. (*C*) Distributions of motion viewing times, showing longer times in the informative condition for both monkeys. In the uninformative condition, M1 shows a prominent peak at 100 ms, the minimum viewing time imposed by the experiment. (*D*) Regression coefficients (mean and SEM) showing that the presaccadic responses encode the validity of informative cues but not the postsaccadic viewing times (Eq. S7).

each monkey), showing that the monkeys were sensitive to and motivated by the RPE associated with the second cue. Second, we controlled for the possibility that neurons only modulated for the first step in an action sequence, by using an additional subset of "change to self" trials where the first and second cues had equal validity of 100%, 80%, or 55% (*Methods*). On these trials, the neurons showed a significant modulation according to the validity of RF cue ($P < 0.0009$), verifying that a validity response could be elicited at the second step in a sequence. Finally, examination of the full (not mean-subtracted) response showed that the neurons had a robust visual response to the RF cues, which, similar to the uninformative task, showed a slight enhancement due to reorienting that was uncorrelated with the sensitivity to validity (Fig. S5 *B* and *C*).

Together, the findings rule out spurious explanations, and suggest that LIP neurons encode EIG independently of changes in reward expectations.

Discussion

We show that, in a task in which monkeys could select which stimulus to sample before choosing a final action, LIP neurons encoded the expected gains in decision information associated with alternative cues, and this encoding was distinct from the cells' well-known visual, saccade, and reward modulations, and from an encoding of RPE. Although replicating the essential features of instrumental sampling requires relatively complex sequential paradigms (10, 29), extensive analyses and control conditions ruled out confounds that may arise in such paradigms. We discuss the implication of the findings from the perspectives of the value-based and priority-based interpretations of LIP function.

Value-Based Decisions. A prominent interpretation of the LIP target selection responses is that they encode the relative reward values of competing options, which can be read out by downstream mechanisms to select reward-maximizing action policies (11, 22). Our results are broadly consistent with a decision-based interpretation, as the responses we found provided a presaccadic, validity-based ranking of the alternative cues that could guide the decision of which cue to sample.

Our key finding, however, is that the ranking based on EIG could not be explained by the reward mechanisms that have been considered in previous investigations. LIP neurons are sensitive to future rewards and are thought to encode the cumulative future value of an action or state, consistent with the predictions of model-free RL mechanisms (21, 30). In our task, however, the cells distinguished between informative and uninformative items with equivalent reward expectations, and did not modulate as a function of RPE [whose encoding is well established for DA cells (31, 32) but had not been investigated in LIP], indicating that they encode the EIG of visual cues in a manner that is not captured by model-free reinforcement mechanisms.

It is important to note that, although our findings imply that expected rewards are not sufficient to explain the LIP response, they leave open the possibility that rewards contribute to constructing this response. Indeed, a small fraction of the cells we examined showed significant modulations for uninformative stimuli and in the change-cue task, and may potentially encode

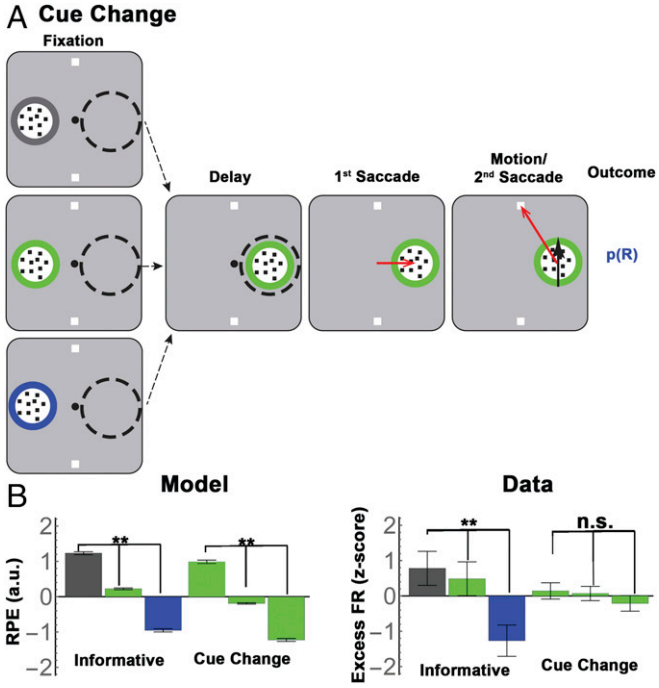

Fig. 6. Validity responses cannot be explained by RPE. (*A*) Trial stages for the cue-change trials. All conventions are as in Fig. 3*A*. The monkeys first viewed an informative cue, which had stationary dots but a known validity (border color), and appeared simultaneously with the targets, opposite the RF ("Fixation"). On the cue-change trials depicted here (which were 25% of all trials), the initial cue was replaced with a different RF cue. Thereafter, the trial proceeded identically to the one-cue informative condition, with a 500-ms delay period, followed by a saccade to the RF cue and a second saccade to a target. Note that the first cue never delivered its motion instruction; its purpose was to establish an initial reward expectation (by virtue of its validity), which could be modified by the second cue, producing an RPE. (*B*) (*Left*) RL simulations confirm that informative cues were associated with RPEs that were proportional to their validity ("Informative"). Cues of 80% validity that appeared on cue-change trials ("Cue Change") were associated with similar RPEs by virtue of following an initial cue of 55%, 80%, or 100% validity. The bars show mean and SEM of simulated RPEs, *z*-scored across all conditions. Green bars for the cue-change condition are arranged in order of RPE. (*Right*) LIP responses modulated for informative cues but not 80% valid cues with matched RPE. The bars show the mean and SEM of the mean-subtracted, *z*-scored firing rates shown in Fig. 7*A*, averaged across the interval of peak effect (125 ms to 250 ms after cue onset) and all 24 cells tested in this task. (**$P < 0.001$; two-way ANOVA with post hoc comparisons; n.s., $P = 0.87$).

reward expectation and/or RPE, consistent with the integrative nature of the LIP target-selective response (14, 33, 34).

Our conclusion that LIP neurons encode higher-order forms of utility beyond simple reward expectation is consistent with two previous studies demonstrating a sensitivity to values based on social status (35) and the motivational salience of a punishment-predicting cue (36). Interestingly, social status and motivational salience identify stimuli that govern future actions, raising the question of how social and motivational factors shape active sensing policies.

Our findings raise important questions about how the brain may compute EIG. One debated question is whether EIG estimates rely on explicit measures of uncertainty (37) or higher-order effects of rewards [such as convex utility function or nonstandard effects of RPE that have yet to be characterized in individual cells (38–40)]. A second key question is whether EIG is computed dynamically based on the uncertainty of each forthcoming action or relies on long-term estimates of average validity. A dynamic "look ahead" mechanism affords high flexibility, but it may be computationally expensive and may not be used consistently in all behavioral contexts (10, 29). An important question for future research is to what extent EIG computations are flexible and responsive to rapid changes in context, or rely on stored validity representations to produce routine-based sampling policies (13, 41, 42).

What benefits might the brain derive from computing a reward-independent response to EIG, given that information

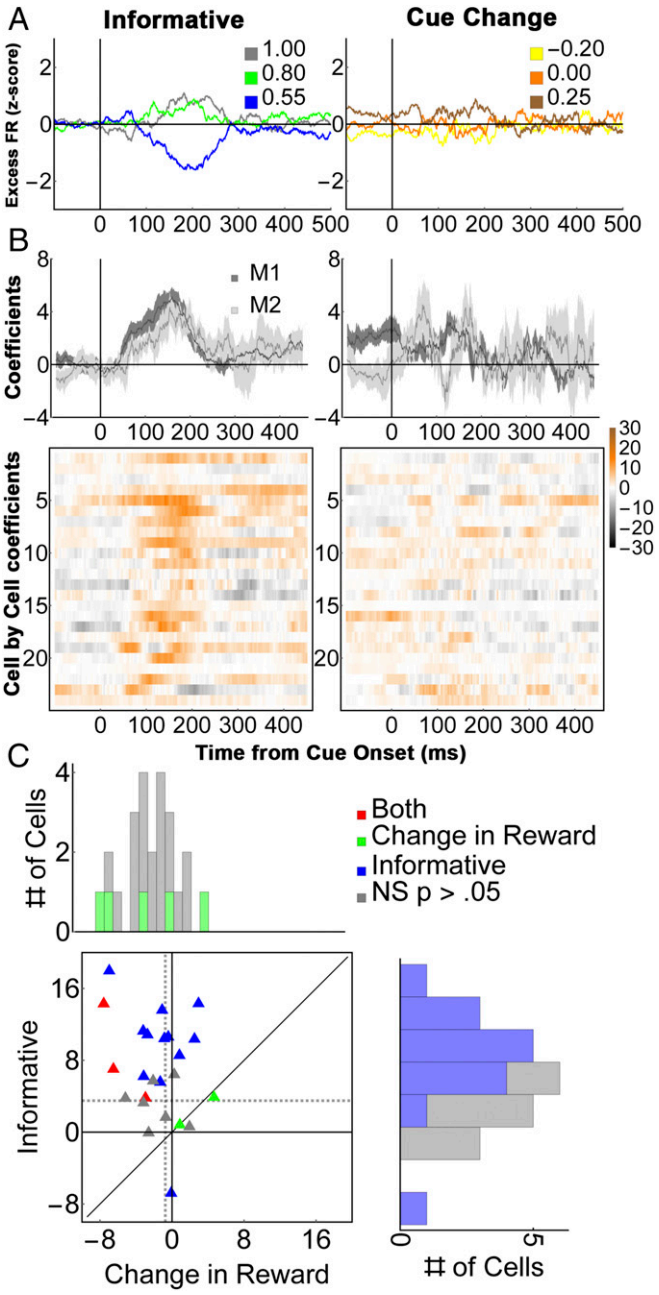

Fig. 7. Neural responses on the reward change test. (*A*) Average firing rates for (*Left*) valid cues and (*Right*) 80% cues with matched RPEs ($n = 24$ cells). To highlight the cue-related modulation, firing rates were *z*-scored after subtracting the average activity for each stimulus class (the term "Excess" indicates mean subtraction). The legends in each panel show the validity of the informative cues and the equivalent RPEs for the 80% cues on cue-change trials. (*B*) Average regression coefficients for the validity/reward change responses for (*Top*) each monkey and (*Bottom*) individual cells (colormaps). (*C*) Paired comparison of the validity and RPE effects in individual cells. All conventions are as in Fig. 4.

sampling must ultimately support reward maximization? One possible answer to this question comes from studies of cognitive control, which suggest that actions that require attentive control are those associated with new information, whereas actions that have low uncertainty are habitual and can be performed with little attention (43–45). This implies that the brain must triage cognitive effort according to not only future rewards but also the informational demands of a situation, requiring a reward-independent sensitivity to EIG.

A second answer to this question comes from the fact that the rewards associated with sampling are very indirect and, as was the case in our task, contingent on postsampling actions. In many conditions, the postsampling decisions may be quite complex (2, 41), and their rewards may be ambiguous or fully unknown [as in curiosity-based exploration (6)]. In such complex conditions, the brain may derive a significant advantage from computing decision variables over shorter time scales, related to reducing the uncertainty of a proximate action.

Attention and Gaze. A second common interpretation of the LIP target selection response is in terms of a "priority" map that ranks competing visual cues for saccades or attention (14, 33, 46–52). Similar to the value interpretation, the priority hypothesis describes LIP as encoding a common currency for visual selection. However, the priority hypothesis goes beyond the value framework in making the specific proposal that LIP cells provide topographically organized attentional feedback facilitating early sensory discrimination (14, 53).

The EIG modulations we found cannot be explained by a number of previously described attention/priority effects. First, these modulations cannot provide topographically organized attentional feedback, as they were not aligned in time or space with the perceptual discrimination: Although the EIG responses described a cue at a peripheral location before the saccade, the motion discrimination occurred at a foveal location after the saccade. In addition, EIG responses were uncorrelated with reorienting of spatial attention (27) or saccade motor factors.

Within the priority framework, our results are broadly consistent with the idea that LIP cells encode "relevance" or "top-down" target selection (14, 54). However, it is important to note that no studies, whether at the level of neural responses or in the behavioral/computational literatures, have attempted to give a computational definition of this form of selection, producing great difficulties in our ability to model top-down attention and gaze (41, 55). A key contribution of our results, therefore, is that they reveal a specific neural signal of task relevance based on expected gains in decision information, making these signals amenable to computational modeling in future investigations.

Methods

Data were collected from two adult male rhesus monkeys (*M. mulatta*) using standard techniques (56), approved by the Animal Care and Use Committees of Columbia University and New York State Psychiatric Institute as complying with the guidelines within the Public Health Service Guide for the Care and Use of Laboratory Animals. Eye position was recorded with Applied Science Laboratories eye tracking system digitized at 240 Hz. Visual stimuli were presented at 57 cm viewing distance on a Sony GDM-FW900 Trinitron monitor (viewing area of 30.8° by 48.2°), and their onset was measured by a diode that detected the onset of a refresh cycle.

Task. For all of the task versions, cues and uninformative stimuli were round patches that measured 3.5° in diameter and contained small dots (0.2°), and the targets were small squares of 0.4° on a side. Cues of different validity correctly indicated the rewarded target on, respectively, 100%, 80%, and 55% of trials and signaled the erroneous target on the remaining, randomly interleaved trials. The three validities were signaled with equiluminant gray, blue, and green borders, with validity-color mappings held constant for each monkey and randomized across monkeys. The display was adjusted for each LIP cell so that, when the monkeys held central fixation, one of the cues fell inside the RF (typically at eccentricities of 8° to 12°) while the other cue was at the diametrically opposite location, and the two targets were at equal eccentricities around an axis orthogonal to that linking the cues (also outside the RF). To ensure that the monkeys used the motion to guide their second saccade, we imposed minimum motion viewing times before making the second saccade (100 ms for M1, and 100 ms to 300 ms for M2). Rewards, if given, arrived at a fixed interval of 200 ms after the end of the second saccade.

Each trial began when the monkeys achieved and maintained central fixation for 400 ms to 600 ms. After this interval, the two targets appeared, on their own on standard informative trials (Figs. 1*A* and 3*A*, first panels), or simultaneous with the opposite-RF precue on uninformative and change-cue trials (Figs. 3*B* and 6*A*, first panels). The target/precue period lasted 300 ms for M1 and 350 ms for M2, and was followed by the appearance of the RF cues, the 500-ms delay period, and the two-saccade sequence as described in *Results*. Therefore, the three task versions differed only in whether or not a precue appeared together with the targets, while the timing was identical across the different tasks. In the two-cue choice task (Fig. 1*A*), trials with the three possible pairs of cues with unequal EIG were presented in random order, with the location of the higher-validity cue randomized to fall inside or opposite the RF.

In the uninformative cue condition (Fig. 3*B*), initial testing showed that the monkeys had difficulties performing single-cue informative and uninformative trials if these were interleaved. In addition, M2 had relatively low performance for the 100% valid cue (even though he reliably selected this cue on free-choice trials; Fig. 1*B*). To compensate for these difficulties and minimize reward confounds associated with performance differences across the two tasks, we ran the single-cue informative and uninformative trials in short interleaved trial blocks (collecting at least 10 completed trials for each condition and randomizing the order of the blocks), and based the analysis on the 55% and 80% valid cues that were tested in both monkeys.

In the cue-change task (Fig. 6), it was necessary for the monkeys to believe that the initial cue signaled reward probability on most trials, which meant that we had to maintain a low frequency of cue-change trials, which we set at 25%. This low frequency, in turn, prevented us from exhaustively examining all of the nine possible combinations of the first and second cues, and we decided to focus on the five sequences that were most relevant for our purposes. These were sequences in which the first cue had validity of 100%, 80%, or 55% and the second, RF, cue had validity of 80%, and two more sequences in which the first and second cues had equal validities of 100% and 55%. These trial types allowed us to test whether the responses to an 80% RF cue modulated according to the reward expectations set by a prior cue (Fig. 6*A*) and how neurons responded to cues that had different validities but zero RPE (100, 100%; 80, 80%; and 55, 55%). These five types of cue-change trials appeared with equal probability and were randomly interleaved with no-change trials. A block continued until at least five trials were completed for each trial type.

Neural Recordings. Single electrodes were advanced into the intraparietal sulcus (IPS) using a Kopf Microdrive (David Kopf Instruments), and the data were recorded using the Advanced Processing Module for neural signal recording Fred Haer (FHC, Inc.), and MatLab (MathWorks) and Mathematica (Wolfram) were used for off-line data analysis.

Neurons were identified as belonging to LIP based on accepted anatomical and physiological criteria. We used structural MRI to guide electrode placement to the appropriate level of the IPS and, during recordings, restricted our recordings depths to 3 mm to 8 mm below the cortical surface. At the end of the recordings, we acquired a final structural MRI with electrodes inserted at the anterior and medial margins of the region from which we had obtained cells, and found that the trajectories of these electrodes were fully in the lateral bank. Thus, we can be confident that all of the other recording locations—which were posterior and lateral to these landmarks—were also in the lateral bank. For physiological verification, we screened each isolated cell and only tested it further if it had significant, spatially tuned delay period activity during a memory-guided saccade task (one-way ANOVA, $P < 0.05$). These well-established criteria conclusively distinguish LIP from neighboring areas that are located in the medial bank of the IPS (medial intraparietal area), on the lateral cortical surface (7a), or at the bottom of the IPS (ventral intraparietal area), and which have visual and postsaccadic responses but much weaker delay period activity before memory-guided saccades (57, 58). Note also that, because we were biased toward recording cells with high delay period activity, we cannot speak to any correlations (or lack thereof) that may exist between this activity and validity modulations.

Data Analysis. Analysis was based on 69 well-isolated neurons (40 in M1) that were tested in the one-cue informative/uninformative task, of which 50 were also tested with the two-cue task, and 24 were tested with the cue-change task. Incomplete trials (in which the monkey did not make a second saccade)

were removed from the analysis. Complete trials were analyzed whether or not they received a reward. Saccades were analyzed using a velocity-based algorithm (59), and saccade onset was calculated from the earliest sample of continuous acceleration. All statistical analyses were preceded by tests of normality and symmetry ($P < 0.05$). If the data met the criteria of normality and symmetry, a paired-sample *t* test was used. If only the symmetry criterion was met, a Wilcoxon-signed-rank test was used. If neither criterion was met, a Mann–Whitney *u* test was computed.

Analysis of the neural responses was always conducted on unsmoothed rates. Time-resolved regression analyses were computed using standardized coefficients, on firing rates measured in a 50-ms window stepped by 1 ms throughout the delay period. For graphical displays, the value for each time bin was plotted in the middle of the 50-ms window (e.g., 0 ms to 50 ms is plotted at 25 ms). Note that the coefficients are signed (not converted to absolute values), and thus a positive coefficient indicates a true coding of that parameter (rather than spurious effects that may masquerade as positive coding if absolute values are taken). For display purposes only, firing rates were convolved with the right half of a Gaussian kernel of 20 ms SD that was centered on the true spike time, smearing the signal only forward in time.

Our main analyses are based on *z*-scored firing rates, which normalizes for overall differences in firing levels across cells. However, we obtained equivalent results using nonnormalized rates (Fig. S6). To *z*-score firing rates within a cell, we computed the average firing rate in each trial obtained from that cell, measured from the time of fixation point onset to the end of the trial (~200 ms after the end of the reward period). We then computed the mean and SD of these firing rates across all of the trials that were included in a given analysis, and transformed the trial by trial firing rate by subtracting the mean and dividing by the SD of this distribution. Individual trial *z* scores were averaged for a cell and then averaged across all cells to obtain the average and SEs of the population response.

ACKNOWLEDGMENTS. We thank Latoya Palmer and Cherise Washington for expert administrative assistance; Girma Asfaw for superior veterinary care; and Genevieve Price, Santiago Alonso Diaz, Kirsten Quiles, and Richard Meehan for help with animal training. The work was supported by the Kavli Institute for Brain Science at Columbia University, National Institutes of Health (NIH) Training Grants EY013933-10 and T32 MH015174-35, and NIH Grants 5R01MH098039 (to J.G.) and R24 EY015634 (to J.G.).

1. Nelson JD (2005) Finding useful questions: On Bayesian diagnosticity, probability, impact, and information gain. *Psychol Rev* 112:979–999.
2. Yang SC, Lengyel M, Wolpert DM (2016) Active sensing in the categorization of visual patterns. *eLife* 5:e12215.
3. Gold JI, Shadlen MN (2007) The neural basis of decision making. *Annu Rev Neurosci* 30:535–574.
4. Blanchard TC, Hayden BY, Bromberg-Martin ES (2015) Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. *Neuron* 85:602–614.
5. Bromberg-Martin ES, Hikosaka O (2009) Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron* 63:119–126.
6. Gottlieb J, Oudeyer PY, Lopes M, Baranes A (2013) Information-seeking, curiosity, and attention: Computational and neural mechanisms. *Trends Cogn Sci* 17:585–593.
7. Nelson JD, McKenzie CR, Cottrell GW, Sejnowski TJ (2010) Experience matters: Information acquisition optimizes probability gain. *Psychol Sci* 21:960–969.
8. Najemnik J, Geisler WS (2008) Eye movement statistics in humans are consistent with an optimal search strategy. *J Vis* 8:4.1–14.
9. Johnson L, Sullivan B, Hayhoe M, Ballard D (2014) Predicting human visuomotor behaviour in a driving task. *Philos Trans R Soc Lond B Biol Sci* 369:20130044.
10. Eckstein MP, Schoonveld W, Zhang S, Mack SC, Akbas E (2015) Optimal and human eye movements to clustered low value cues to increase decision rewards during search. *Vis Res* 113:137–154.
11. Sugrue LP, Corrado GS, Newsome WT (2005) Choosing the greater of two goods: Neural currencies for valuation and decision making. *Nat Rev Neurosci* 6:363–375.
12. Nakamura K (2006) Neural representation of information measure in the primate premotor cortex. *J Neurophysiol* 96:478–485.
13. Hayhoe M, Ballard D (2014) Modeling task control of eye movements. *Curr Biol* 24: R622–R628.
14. Bisley JW, Goldberg ME (2010) Attention, intention, and priority in the parietal lobe. *Annu Rev Neurosci* 33:1–21.
15. Platt ML, Glimcher PW (1999) Neural correlates of decision variables in parietal cortex. *Nature* 400:233–238.
16. Lee D, Seo H, Jung MW (2012) Neural basis of reinforcement learning and decision making. *Annu Rev Neurosci* 35:287–308.
17. Louie K, Grattan LE, Glimcher PW (2011) Reward value-based gain control: Divisive normalization in parietal cortex. *J Neurosci* 31:10627–10639.
18. Louie K, Glimcher PW (2010) Separating value from choice: Delay discounting activity in the lateral intraparietal area. *J Neurosci* 30:5498–5507.
19. Kable JW, Glimcher PW (2007) The neural correlates of subjective value during intertemporal choice. *Nat Neurosci* 10:1625–1633.
20. Dorris MC, Glimcher PW (2004) Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44:365–378.
21. Sugrue LP, Corrado GS, Newsome WT (2004) Matching behavior and the representation of value in the parietal cortex. *Science* 304:1782–1787.
22. Kable JW, Glimcher PW (2009) The neurobiology of decision: Consensus and controversy. *Neuron* 63:733–745.
23. Burnham KP, Anderson DR (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (Springer, New York), 2nd Ed.
24. Gottlieb J, Hayhoe M, Hikosaka O, Rangel A (2014) Attention, reward, and information seeking. *J Neurosci* 34:15497–15504.
25. Dayan P, Daw ND (2008) Decision theory, reinforcement learning, and the brain. *Cogn Affect Behav Neurosci* 8:429–453.
26. Sutton RS, Barto AG (1998) *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, MA).
27. Robinson DL, Bowman EM, Kertzman C (1995) Covert orienting of attention in macaques. II. Contributions of parietal cortex. *J Neurophysiol* 74:698–712.
28. Niv Y, Schoenbaum G (2008) Dialogues on prediction errors. *Trends Cogn Sci* 12:265–272.
29. Morvan C, Maloney LT (2012) Human visual search does not maximize the post-saccadic probability of identifying targets. *PLOS Comput Biol* 8:e1002342.
30. Seo H, Barraclough DJ, Lee D (2009) Lateral intraparietal cortex and reinforcement learning during a mixed-strategy game. *J Neurosci* 29:7278–7289.
31. Waelti P, Dickinson A, Schultz W (2001) Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412:43–48.
32. Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47:129–141.
33. Ipata AE, Gee AL, Bisley JW, Goldberg ME (2009) Neurons in the lateral intraparietal area create a priority map by the combination of disparate signals. *Exp Brain Res* 192:479–488.
34. Park IM, Meister ML, Huk AC, Pillow JW (2014) Encoding and decoding in parietal cortex during sensorimotor decision-making. *Nat Neurosci* 17:1395–1403.
35. Klein JT, Deaner RO, Platt ML (2008) Neural correlates of social target value in macaque parietal cortex. *Curr Biol* 18:419–424.
36. Leathers ML, Olson CR (2012) In monkeys making value-based decisions, LIP neurons encode cue salience and not action value. *Science* 338:132–135.
37. Daddaoua N, Lopes M, Gottlieb J (2016) Intrinsically motivated oculomotor exploration guided by uncertainty reduction and conditioned reinforcement in non-human primates. *Sci Rep* 6:20202.
38. Beierholm UR, Dayan P (2010) Pavlovian-instrumental interaction in 'observing behavior.' *PLOS Comput Biol* 6:e1000903.
39. Kreps DM, Porteus EL (1978) Temporal resolution of uncertainty and dynamic choice theory. *Econometrica* 46:185–200.
40. Iigaya K, Story GW, Kurth-Nelson Z, Dolan RJ, Dayan P (2016) The modulation of savouring by prediction error and its effects on choice. *eLife* 5:e13747.
41. Tatler BW, Hayhoe MM, Land MF, Ballard DH (2011) Eye guidance in natural vision: Reinterpreting salience. *J Vis* 11:5–25.
42. Desrochers TM, Jin DZ, Goodman ND, Graybiel AM (2010) Optimal habits can develop spontaneously through sensitivity to local cost. *Proc Natl Acad Sci USA* 107: 20512–20517.
43. Cavanagh JF, Frank MJ (2014) Frontal theta as a mechanism for cognitive control. *Trends Cogn Sci* 18:414–421.
44. Shenhav A, Botvinick MM, Cohen JD (2013) The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron* 79:217–240.
45. Fan J (2014) An information theory account of cognitive control. *Front Hum Neurosci* 8:680.
46. Mirpour K, Bisley JW (2012) Dissociating activity in the lateral intraparietal area from value using a visual foraging task. *Proc Natl Acad Sci USA* 109:10083–10088.
47. Bisley JW, Mirpour K, Arcizet F, Ong WS (2011) The role of the lateral intraparietal area in orienting attention and its implications for visual search. *Eur J Neurosci* 33:1982–1990.
48. Mirpour K, Ong WS, Bisley JW (2010) Microstimulation of posterior parietal cortex biases the selection of eye movement goals during search. *J Neurophysiol* 104:3021–3028.
49. Foley NC, Jangraw DC, Peck C, Gottlieb J (2014) Novelty enhances visual salience independently of reward in the parietal lobe. *J Neurosci* 34:7947–7957.
50. Gottlieb J, Snyder LH (2010) Spatial and non-spatial functions of the parietal cortex. *Curr Opin Neurobiol* 20:731–740.
51. Gottlieb J, Kusunoki M, Goldberg ME (2005) Simultaneous representation of saccade targets and visual onsets in monkey lateral intraparietal area. *Cereb Cortex* 15:1198–1206.
52. Gottlieb JP, Kusunoki M, Goldberg ME (1998) The representation of visual salience in monkey parietal cortex. *Nature* 391:481–484.
53. Reynolds JH, Heeger DJ (2009) The normalization model of attention. *Neuron* 61: 168–185.
54. Gottlieb J, Balan P (2010) Attention as a decision in information space. *Trends Cogn Sci* 14:240–248.
55. Baluch F, Itti L (2011) Mechanisms of top-down attention. *Trends Neurosci* 34: 210–224.
56. Oristaglio J, Schneider DM, Balan PF, Gottlieb J (2006) Integration of visuospatial and effector information during symbolically cued limb movements in monkey lateral intraparietal area. *J Neurosci* 26:8310–8319.
57. Barash S, Bracewell RM, Fogassi L, Gnadt JW, Andersen RA (1991) Saccade-related activity in the lateral intraparietal area. II. Spatial properties. *J Neurophysiol* 66: 1109–1124.
58. Snyder LH, Batista AP, Andersen RA (2000) Intention-related activity in the posterior parietal cortex: A review. *Vision Res* 40:1433–1441.
59. Nyström M, Holmqvist K (2010) An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behav Res Methods* 42:188–204.