Understanding active sampling strategies: empirical approaches and implications for attention and decision research

Jacqueline Gottlieb[1,2]

[1]Department of Neuroscience, Columbia University

[2]Zuckerman Mind Brain Behavior Institute, Columbia University

Abstract

In natural behavior we actively gather information using attention and active sensing behaviors (such as shifts of gaze) to sample relevant cues. However, while attention and decision making are naturally coordinated, in the laboratory they have been dissociated. Attention is studied independently of the actions it serves. Conversely, decision theories make the simplifying assumption that the relevant information is given to the decision maker, and do not attempt to describe how she may learn and implement active sampling policies. In this paper I review recent studies that address questions of attentional learning, cue validity and information seeking in humans and non-human primates. These studies suggest that learning a sampling policy involves large scale interactions between networks of attention and valuation, and that these policies are motivated by reward maximization, uncertainty reduction and the intrinsic utility of cognitive states. I discuss the importance of using such paradigms for formalizing the role of attention and devising more realistic theories of decision making that capture a broader range of empirical observations.

Introduction

The oculomotor system of humans and non-human primates holds a privileged status in neuroscience research. Motivated by the relative simplicity of the eye motor plant, the relative ease of measuring eye movements in the laboratory and the high degree of similarity between humans and non-human primates, scores of investigations have examined saccades – the rapid shifts of gaze that primates use to scan visual scenes – and characterized the neural pathways involved in their generation.

However, while these studies have elucidated many of the sensorimotor mechanisms involved in saccades, progress has stalled in explaining the cognitive aspects of saccades and

Corresponding author: Jacqueline Gottlieb, PhD, Department of Neuroscience, Columbia University, 1051 Riverside Drive, Kolb Research Annex, Rm. 569, New York, NY 10032, Phone: 212-543-6931, ext. 500, Fax: 212-543-5816, jg2141@columbia.edu.

attention – specifically, how the brain selects task-relevant cues. While behavioral evidence makes it clear that gaze is under strong task-related control – with humans deploying gaze very selectively to stimuli that are relevant to their immediate actions with minimal influence from salient distractors (Tatler, Hayhoe, Land, & Ballard, 2011; Yarbus, 1967) - computational models of gaze allocation are based primarily on bottom-up saliency (Berg, Boehnke, Marino, Munoz, & Itti, 2009; White BJ, Berg DJ, Marino RA, Itti L, & DP, 2017) with many fewer attempts to model task-related control (Navalpakkam & Itti, 2005; Tatler et al., 2011).

This gap in our understanding is particularly vexing for neurobiological investigations of oculomotor structures implicated in the selection of targets for attention or gaze, which include the superior colliculus, the frontal eye field (FEF) and the lateral intraparietal area (LIP) (Bisley & Goldberg, 2010; Krauzlis RJ, Lovejoy LP, & A., 2013; Thompson & Bichot, 2005). While abundant evidence shows *that* neurons in these areas encode top-down visual selection – selectively signaling the locations of task-relevant visual stimuli - we have little insight into *how* these responses arise. How do target selection neurons "know" which target to select? What is the computational definition of a "task-relevant" cue? While several lines of research have linked target selection responses in LIP with simple decisions based on perceptual evidence or rewards (Hanks TD & C., 2017; Kable & Glimcher, 2009; Sugrue, Corrado, & Newsome, 2005), these studies have yet to consider the unique information sampling nature of gaze (Gottlieb, Hayhoe, Hikosaka, & Rangel, 2014) and leave persistent unresolved questions about the selection process encoded by the cells (Gottlieb, 2012; Maunsell, 2004).

In this article I argue that, to understand task-related control, we must acknowledge the essential role of attention and gaze in sampling information – the fact that, in natural conditions, gaze and attention implement an *active sensing policy* that is coordinated with the decision maker's beliefs, goals or actions. While the informational – or epistemic - nature of saccades and attention is recognized by theoretical frameworks of predictive coding, which emphasize the imperative of minimizing surprise or free energy (Karl Friston, 2010; K. Friston & Ao, 2012; Karl Friston et al., 2015; Schwartenbeck, Fitzgerald, Dolan, & Friston, 2013) or by expanded reinforcement learning theories (Iigaya, Story, Kurth-Nelson, Dolan, & Dayan, 2016) – we have scant empirical data that can constrain or refine these theories.

I this paper I review the few studies that have addressed this question, with a focus on the features of behavioral paradigms that can probe the logic of active sensing policies and the key current findings regarding these policies in humans and monkeys. I argue that, although these approaches are relatively new to the field, developing them is essential for expanding our current understanding of both attention and decision making and bringing about a closer integration of research on these topics.

What does an observing decision entail?

Because questions of active sampling are relatively unfamiliar in the study of oculomotor control, it is useful to start by considering the computations that they may entail. Active

sampling is an ubiquitous aspect of natural behavior, and a core building block of the perception-action cycle: before deciding what to do at an intersection we *look* at the traffic (or a traffic sign, or a traffic light) and, before deciding whether to reach for the peanut butter jar we *look* at the jar. As these example illustrate, understanding active sampling requires us to consider two related decisions: the selection of a task-relevant cue, followed by the decision of which action to take based on that cue.

Sequential decisions of this kind are typically analyzed (e.g., in reinforcement learning frameworks) using a Markov decision chain such as that illustrated in Fig. 1A, which specifies a sequence of states that the decision maker expects to traverse in a task, and the probabilistic actions and transitions that are possible from each state. In the case of a pedestrian reaching an intersection (Fig. 1A), the chain may start with the decision of whether to look at the traffic light or a cloud, followed by the decision of whether to stop or proceed followed by the observation of an outcome (e.g., staying safe, operationalized as a reward probability).

Our concern is with the first decision in this chain – the determination of which stimulus to sample – and the diagram in Fig. 1A illustrates three key points about this step: it depends on prior knowledge of the task structure, it may be guided by both expected rewards and the prospect of resolving uncertainty, and it requires the agent to estimate the desirability of the available cues in advance of the full sensory discrimination. I discuss each feature in turn.

Model-based selection

One of the most important features of active sampling policies is that, like other types of decisions, they depend on prior knowledge of the task structure. This knowledge is embodied in a task model such as that shown in Fig. 1A, which specifies the states and actions involved in a task, as well as the relation between stimuli and subsequent states. It is only based on this knowledge that the agent can estimate the probability (or uncertainty) of competing actions and the *meaning* of sensory cues as well as the information that the cues may bring about future states (e.g., that the colors of the traffic light are associated with crossing or waiting). Mechanistically, this feature implies a hierarchical process whereby prior knowledge organizes local sampling strategies. As we will see in the following sections, the role of hierarchical learning in taskrelated saccade and attention control is an important topic for further investigation.

Dependence on reward and uncertainty

A second critical feature is that, in the context of a task model, there are two possible mechanisms for distinguishing between informative and uninformative cues: the reward expectations associated with a cue, and the prospect that a cue will alter the decision maker's beliefs about future states, and these mechanisms may play differential roles according to the context (Johnson, Sullivan, Hayhoe, & Ballard, 2014; Sullivan, Johnson, Rothkopf, Ballard, & Hayhoe, 2012).

In conditions where the decision maker can act based on the sampled information – so called *instrumental* sampling paradigms - an informative cue is *by definition* one that signals the more desirable action, and thus the reliability of a cue is closely correlated with the chance

of success in the task. In the example shown in Fig. 1A, if the pedestrian looks at the traffic light and takes the action signaled by that light – be it to stop or proceed – she has correctly estimated the state of the world and has a high chance of success in the task (high reward probability). However, should the pedestrian decide to look at the cloud, she will choose her actions at random and is likely to have a much lower reward probability. Although the saccade or shift of attention is not the reward-harvesting action and not the decision maker's primary goal, these actions acquire indirect reward value, because the eventual probability of success of the action sequence is larger if one starts by sampling an informative rather than an uninformative cue.

However, it is important to note that, while informativeness is closely aligned with reward associations in some instrumental contexts, this relationship is not obligatory in all task conditions. These conditions include *non-instrumental* paradigms in which agents may simply want to *know* but cannot act on the outcome (discussed in the following sections), as well as natural behaviors involving curiosity that are beyond the scope of this review (Gottlieb, Oudeyer, Lopes, & Baranes, 2013).

While reward associations are inconsistent markers for cue reliability, a more useful quantity is the potential of a cue to alter the decision maker's beliefs about future task states – or in other words, the expected information gains (EIG) associated with sampling a cue. The pedestrian in Fig. 1A starts the task with uncertainty about which action to take (e.g., may believe that stopping or crossing are equally desirable alternatives) and expects that her beliefs about the most appropriate action will be modified if she observes the traffic - but not if she observes the cloud. Note that the EIG associated with different sampling strategies can be precisely computed *based on knowledge of a task model*, by comparing the distributions of possible beliefs before and after observing a cue – for instance, using common information metrics such as KL divergence or differences in Shannon entropy (Fig. 1A legend). In the following sections I will describe experiments illustrating how such computations may help allocate gaze.

Prospective nature

A final important aspect of information sampling is that it requires the agent to make a sampling decision *before* discriminating the sampled information (Navalpakkam & Itti, 2005). The pedestrian in Fig. 1A must decide to look at the traffic light before knowing whether the light is red or green or whether she will stop or proceed – indeed, she must decide to attend *in order to enable* the perceptual discrimination.

This is an important feature that has been ignored in traditional paradigms. By instructing participants to attend to well defined cues (e.g., "look at the red horizontal line") these paradigms confound the sensory discrimination of the detailed stimulus features (i.e., is it red and horizontal) with the determination of task-relevance (i.e., that color and orientation are the relevant dimensions to monitor in the task, rather than, for instance, size or location). To understand active sampling therefore, we must think of shifts of gaze or attention as *proactive requests for information* – or, to use other terms, *questions* that we pose to the world or the *opening of information channels* – and attempt to understand how the agent

selects *which question to ask* based on the estimated costs and benefits associated with the possible answers.

Armed with these considerations, we can now review some of the empirical studies devoted to active sampling mechanisms.

Learning and encoding of cue validity

A few seminal experiments in humans and monkeys have probed some of the features I described in the previous section, including the use of information gains to guide saccade sampling decisions, the neural mechanisms involved in estimating and updating cue reliability, and the explicit encoding of reliability in cortical area LIP. I will review each in turn.

To understand whether human observers guide saccadic scanning based on estimates of information gains, Yang and colleagues trained participants to use visual scanning to infer which type of pattern – zebra-like stripes or cheetah-like spots – was lurking underneath a masked visual display (Yang, Lengyel, & Wolpert, 2016). Participants made a series of saccades before classifying the pattern and received information in a gaze-contingent fashion – through transient removals of the mask in a local region around their current fixation (Fig. 2A). To model optimal behavior, the authors used a Bayesian active sampling (BAS) model that (1) updated its estimates of the posterior probability of each candidate pattern based on the information acquired in each successive fixation, and (2) estimated the expected information gains (EIG) of each potential (next) fixation location based on its current beliefs and the known statistics of the candidate patterns. The key finding was that gaze allocation, while falling somewhat short of optimal predictions, was strongly biased toward locations with high expected gains in information (Fig. 2B, C). The findings are consistent with previous investigations (Najemnik & Geisler, 2005; Renninger, Verghese, & Coughlan, 2007) and clearly illustrate how saccades may be guided by mathematically defined measures of EIG.

The idea that cue reliability – equivalent to EIG - impacts saccades and attention is supported by a separate series of studies using an extension of the Posner cueing paradigm in which cue was systematically manipulated (S. Vossel et al., 2014; S. Vossel, C., K.E., & J., 2015; S. Vossel, Thiel, & Fink, 2006). The key behavioral finding of these studies is that participants showed a robust sensitivity to reliability, such that the cueing effects on reaction times (RT) – the RT difference for target detection on valid versus invalid trials - increased in proportion of reliability (Fig. 1D). Even though the participants in this paradigm did not have a choice of which cue to sample, they nevertheless adjusted the weight they afforded to a predictive cue based on its estimated reliability. A second important behavioral finding is that participants flexibly updated their estimates of cue reliability when tested in a dynamic regime in which this quantity changed throughout a session in an unannounced fashion. This dynamic updating was successfully modeled using a hierarchical Bayesian framework which represented, at successive levels, beliefs about the immediate target location, beliefs about the current validity of the cue and beliefs about the volatility of the environment (the extent to which cue reliability was expected to change) (S. Vossel et al., 2015).

Functional magnetic imaging (fMRI) suggested that the updating of cue reliability involves several areas, most notably the temporal parietal junction (TPJ), putamen, the frontal eye fields (FEF) and the intraparietal sulcus (IPS) (S. Vossel et al., 2014; S. Vossel et al., 2015; S. Vossel et al., 2006). These areas had stronger responses on *invalid* relative to valid trials - and these responses, as well as the connectivity between the TPJ and FEF, putamen and intraparietal sulcus (IPS) increased as a function of cue reliability – i.e., were highest when the reliability was high and invalid trials were rare. Together with supporting evidence from a study using trans cranial magnetic stimulation (Mengotti, Dombert, Fink, & Vossel, 2017) the findings support the idea that the TPJ signals a reliability-weighted visual prediction error - the "surprisingness" of a target at an invalidly cued location - that is used to update estimates of validity based on trial by trial observations. Additional evidence suggests that the learning rates in this paradigm are sensitive to cholinergic tone (S Vossel et al., 2014) and that different updating mechanisms may be involved depending on the type of information conveyed by the cues (e.g., spatial versus feature cueing (Dombert, Fink, & Vossel, 2016; Dombert, Kuhns, Mengotti, Fink, & Vossel, 2016); or sensory versus motor cueing (Kuhns, Dombert, Mengotti, Fink, & Vossel, 2017)). These studies therefore, reveal some of the mechanisms involved in learning and updating the predictive properties of visual cues.

A more recent study by Leong and colleagues suggest that, in addition to sensory prediction errors, learning to attend may also be sensitive to the rewards of the task (Leong, Radulescu, Daniel, DeWoskin, & Niv, 2017). Participants performed a dynamic decision making paradigm in which they learnt to choose one of several options that differed in their reward probabilities. To track attention allocation, the experimenters constructed each option as a triplet of images - a face, a tool and a landmark – in which one of the images, initially unknown to participants, was predictive of a high reward probability. Using a combined gaze and fMRI measure of attention, the authors could measure learning not only at the level of choice (the extent to which participants learnt to choose the most valuable triplet) but also at the level of attention (the extent to which they learnt to attend to the predictive image within a triplet). The findings suggested that attention used a simple win-stay/lose shift strategy, tending to persist on a feature if that feature was associated with recent rewards but shift to another feature/dimension after reward omission. Switches of attention were associated with enhanced connectivity between the dorsal fronto-parietal network and the ventromedial prefrontal cortex (vmPFC), suggesting that, at least in some cases, learning to attend is sensitive to reward mechanisms.

Although these fMRI investigations highlight the neural systems that may be recruited during reliability updating, they have yet to identify explicit signals encoding cue reliability or link them explicitly with the neural systems generating shifts of gaze or attention. In a recent study in our laboratory, we asked whether cue reliability may be encoded in monkey area LIP, which is implicated in task-related saccadic control and is sensitive to both rewards and informational factors (Gottlieb et al., 2014).

To examine this questions we trained monkeys on a novel paradigm in which they made two contingently related saccades on each trial - a first saccade to gather information from a visual cue, and a second saccade to report a decision based on that information (Fig. 3A). As

in the study of Yang et al., the monkeys received the cue information in a gaze-contingent fashion - only after making a saccade to a cue - and made their saccadic decision based on advance information regarding cue reliability. The monkeys were trained on cues of 100%, 80% and 55% validity that remained stable throughout the recording sessions, so that we examined the steady-state encoding rather than dynamic updating of cue reliability.

When afforded the opportunity to choose which cue to sample, the monkeys consistently chose to inspect the more accurate cue, verifying that they adopted a reliability-based sampling policy. Consistent with this finding, LIP neurons showed robust modulations by cue reliability on both forced choice trials (in which a single cue appeared inside their receptive field; RF) and, importantly, in free-choice trials (Fig. 3B) in which the monkeys had the opportunity to freely select a cue. To see whether these responses were explained merely by reward probability, we compared the neurons' responses to informative cues (which had different reward expectations by virtue of their reliability (cf Fig. 1A) and a set of uninformative stimuli (which were not expected to convey decision information but had equal reward expectation as the informative cues; Fig. 4A). Reinforcement model simulations showed that, if neurons encoded only reward probability, they should discriminate equally well between the different uninformative stimuli and informative cues, as these items had equivalent reward expectation. However, this prediction was disconfirmed by the data. LIP neurons robustly discriminated between cues that had high or low reliability, but did not discriminate between uninformative stimuli with high or low reward expectations (Fig. 4B). Together with an additional control experiment ruling out the possibility that the cells respond to reward prediction errors, these findings establish that LIP cells encode a *bona fide* representation of cue reliability that is independent of simple reward mechanisms.

In sum, these findings suggest that in humans, the learning and updating of cue reliability involves large scale interactions between several systems including cholinergic systems and networks involved in attention and reward valuation. In monkeys, explicit signals of cue reliability are encoded in parietal oculomotor cells and can contribute to the top-down orienting of attention and gaze.

Incentive salience and non-instrumental information seeking strategies

The reliability-based effects discussed in the previous section are consistent with a reward maximizing strategy, because they arise in contexts in which the participants made decisions based on the sampled information. Converging evidence however, shows that animals – including humans, pigeons and monkeys – also seek out cues that are reward predictive even if they cannot take actions based on the sampled information (e.g., (Eliaz & Schotter, 2007; Falk & Zimmermann; Zentall & Stagner, 2012)).

Fig. 1B shows the structure of an observing paradigm that was recently used to reveal such non-instrumental information seeking in monkeys (Bromberg-Martin & Hikosaka, 2009). In this paradigm the monkeys given the opportunity to sample one of two cues, A or B, which had different reliabilities in predicting the size of a reward that would be given at the end of the trial. If the monkeys chose the informative cue (A in Fig. 1B), this cue changed to one of

two patterns (A1 or A2) that provided advance information about reward size. However, if the monkeys chose the uninformative item (B, Fig. 1B), the ensuring two patterns, B1 and B2, had only a random relation to reward size. Similar to instrumental settings therefore, the monkeys were free to observe cues of different reliabilities (cf Fig. 1A and 1B) but in contrast with such settings, they could take no action to alter the reward probability – so that the cues were associated with equivalent reward expectations. A non-instrumental setting is thus a useful laboratory tool for experimentally manipulating reliability independently of reward associations.

The monkeys tested in this paradigm developed a consistent preference for the informative cues, which was associated with a small increase of activity in midbrain dopamine (DA) cells (Bromberg-Martin & Hikosaka, 2009). A subsequent study extended this finding by showing that monkeys are even willing to sacrifice juice reward to view predictive cues, and neurons in the orbitofrontal cortex (OFC) encoded the value that the monkeys placed on this information (Blanchard, Hayden, & Bromberg-Martin, 2015). The responses in DA and OFC cells arose at the time when the sampling decisions are made – before the monkeys discriminated the specific reward information – and thus could motivate an information sampling policy.

Since the payoffs of the informative and uninformative cues were, on average, equal, *prima facie*, these results seem to imply a mechanism that is sensitive strictly to the early resolution of uncertainty (Blanchard et al., 2015; Bromberg-Martin & Hikosaka, 2009). However, because the cues provided information *about* a reward, an alternative possibility is that the monkeys were motivated by the mere desire to reveal a positive, reward associated cue independently of gains in information. This possibility is consistent with a large literature showing that stimuli that have a prior history as reward predictors gain salience and the ability to automatically capture attention even if they do not provide currently relevant information (Anderson, 2016; Foley, Jangraw, Peck, & Gottlieb, 2014; Peck, Suzuki, Efem, & Gottlieb, 2009).

A recent experiment by Daddaoua et al. shows that both motives – the reduction of uncertainty and the seeking of positive cues - shape saccadic information seeking strategies (Daddaoua, Lopes, & Gottlieb, 2016). In this paradigm, monkeys were given the opportunity to search for an informative reward cue in trials that had a 0%, 50% or 100% prior reward probability (Fig. 5A). Importantly, the gaze during the search phase was not under instrumental control and was free to express the monkeys' genuine interest in the additional cue. Search rates, and the probability of revealing the cue, were higher if the trial had a 50% relative to a 100% reward probability, confirming a sensitivity to the prospect of reducing uncertainty. Importantly however, the monkeys also searched if they had *no* uncertainty, provided that the initial cue signaled a 100% rather than 0% reward probability (Fig. 5C).

The monkeys' motivation to reveal the second cue in the 100% condition was striking, because this cue brought no increase in operant gains or a further reduction in uncertainty. The results are consistent with the findings on reward salience, and support the broader idea that humans have intrinsic preferences over cognitive states – preferring states in which they can "savor" anticipated positive events and avoiding those in which they dread anticipated

negative events (Golman & Loewenstein, 2016). A computational model based on savoring has recently proposed that the value associated with reward anticipation can be boosted by the positive reward prediction error (RPE) produced by a predictive cue (Iigaya et al., 2016). However, it remains to be determined whether this idea can explain the robust sampling found by Daddaoua et al. in conditions of 100% prior probability, when the sought-after cue was redundant and thus not associated with RPEs.

Implications and future prospects

The experiments I reviewed above are among the first attempts to probe the logic and neural substrates of information sampling policies in the context of visual attention and gaze – arguably the most intensively investigated active sensing mechanism. While our understanding of these questions is in its infancy, continuing these efforts, I propose, has potentially far reaching implications for both attention and decision research. I close by discussing some of these implications.

Computational description of attention

As I emphasized through the paper, the study of information sampling can be critical for describing attention in concrete computational terms. The lack of well accepted computational models of top-down attention creates confusion between "attention" and "decision"- related neural responses (Gottlieb, 2012; Maunsell, 2004) and makes "attention" seem to be a superfluous cognitive construct. The studies I reviewed are consistent with the longstanding views of attention as responding to informational constraints (Dayan, Kakade, & Montague, 2000; Karl Friston, 2010; K. Friston & Ao, 2012; Karl Friston et al., 2015; Yu & Dayan, 2005), and highlight empirical approaches that can provide much needed verification and refinement of these theoretical perspectives.

I also emphasized the fact that, beyond its well investigated effects on sensory perception and actions, the control of attention entails the *proactive opening of an information channel* based on the costs and benefits that the channel is estimated to bring. The evidence available so far suggests that three types of motives influence this type of decision. One motive relates simply to the extent to which information is expected to increase the operant rewards of a task; a second motivation is related to reducing the uncertainty of belief states; and yet a third motive may be related to the intrinsic utility or dis-utility of anticipating a positive or a negative outcome (savoring or dread). Characterizing these motives, their neural mechanisms and their relative contributions to sampling in different behavioral contexts will be a central goal of future investigations.

Finally, the evidence I reviewed highlights the fact that information sampling has significant costs, which may be related both to processing the sampled information and to learning an efficient sampling policy – learning to recognize the most reliable cues. Behavioral studies suggest that humans have highly efficient routines for sampling information in overlearned tasks (e.g., driving or preparing a sandwich; (Tatler et al., 2011)) but can perform poorly in conditions requiring rapid changes in sampling strategies (Morvan & Maloney, 2012). Understanding how we learn a sampling policy, what are the most helpful feedback regimes,

and what costs are entailed in sampling information, are central questions for further investigation.

Decision making

In addition to its relevance to attention research, an expanded understanding of active information sampling may be critical for devising realistic, cognitively grounded decision theories.

To date, the vast majority of decision research starts from the simplifying assumption that decision makers have little freedom on how they define a decision situation. In laboratory paradigms, participants may be given some amount of control about whether and for how long to sample decision-relevant information (Hanks TD & C., 2017), but the *identity* of the relevant sources is rarely in question. In studies of perceptual decisions, subjects make decisions based on a well-defined perceptual cue, while in studies of value-based decisions, they decide based on well-defined outcomes. Signal detection theory, which is foundational to much decision research, considers how an agent discriminates signal from noise – but makes no mention of the fact that an agent can also determine *which* source of information is a signal and which one is noise.

Realistic decision makers, however, can choose to which aspect of a decision situation to devote more attention, and thus have a tremendously important degree of freedom that is not accounted for by current theories. The evidence that information sampling is based not only on normative strategies for uncertainty reduction or reward maximization but also on idiosyncratic preferences over belief states and cognitive effort, suggests that attentional and information sampling strategies can significantly extend decision theories, and explain some of the marked individual variability and apparent irrationalities that are not currently explained by these theories (Bordalo, Gennaioli, & Shleifer, 2013; Polonio, DiGuida, & Coricelli, 2015; Reis, 2006; Sims, 2003; Woodford, 2009).

While a number of economic theories (e.g., (Caplin & Dean, 2015; Sims, 2003), similar to predictive coding theories, have recognized the importance of informational constraints, these theories are in dire need of empirical data to constrain and refine their predictions. Providing such data is, I propose, essential both for understanding elusive cognitive constructs such as selective attention, and for devising cognitively grounded decision theories that can account for a larger range of observations.

Acknowledgments

References

Anderson B. The attention habit: how reward learning shapes attentional selection. Ann N Y Acad Sci. 2016; 1369(1):24–39. [PubMed: 26595376]

Berg DJ, Boehnke SE, Marino RA, Munoz DP, Itti L. Free viewing of dynamic stimuli by humans and monkeys. J Vis. 2009; 9(5):19, 11–15. [pii]. DOI: 10.1167/9.5.19/9/5/19/

Bisley JW, Goldberg ME. Attention, intention, and priority in the parietal lobe. Annual Review of Neuroscience. 2010; 33:1–21.

Blanchard TC, Hayden BY, Bromberg-Martin ES. Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. Neuron. 2015; 85(3):602–614. [PubMed: 25619657]

Bordalo P, Gennaioli N, Shleifer A. Slience and Consumer Choice. Journal of Political Economy. 2013; 121:803–843.

Bromberg-Martin ES, Hikosaka O. Midbrain dopamine neurons signal preference for advance information about upcoming rewards. Neuron. 2009; 63(1):119–126. doi: S0896-6273(09)00462-0 [pii]. DOI: 10.1016/j.neuron.2009.06.009 [PubMed: 19607797]

Caplin A, Dean M. Revealed Preference, Rational Inattention and Costly Information Acquisition. American Economic Review. 2015; 105(7):2183–2203.

Daddaoua N, Lopes M, Gottlieb J. Intrinsically motivated oculomotor exploration guided by uncertainty reduction and conditioned reinforcement in non-human primates. Sci Rep. 2016; 6(20202)doi: 10.1038/srep20202

Dayan P, Kakade S, Montague PR. Learning and selective attention. Nat Neurosci. 2000; (3 Suppl): 1218–1223. DOI: 10.1038/81504 [PubMed: 11127841]

Dombert PL, Fink GR, Vossel S. The impact of probabilistic feature cueing depends on the level of cue abstraction. Experimental Brain Research. 2016; 234(3):685–694. [PubMed: 26586268]

Dombert PL, Kuhns A, Mengotti P, Fink GR, Vossel S. Functional mechanisms of probabilistic inference in feature- and space-based attentional systems. Neuroimage. 2016; 142:553–564. [PubMed: 27523448]

Eliaz K, Schotter A. Experimental testing of intrinsic preferences for noninstrumental information. American Economics Review. 2007; 97:166–169.

Falk A, Zimmermann F. Beliefs and Utility: Experimental Evidence on Preferences for Information. CESifo Working Paper No. 6061. 2016

Foley NC, Jangraw DC, Peck C, Gottlieb J. Novelty enhances visual salience independently of reward in the parietal lobe. J neurosci. 2014; 34(23):7947–7957. DOI: 10.1523/JNEUROSCI. 4171-13.2014 [PubMed: 24899716]

Friston K. The free-energy principle: a unified brain theory? Nature Reviews Neuroscience. 2010; 11(2):127–138. [PubMed: 20068583]

Friston K, Ao P. Free energy, value, and attractors. Comput Math Methods Med. 2012; 2012:937860.doi: 10.1155/2012/937860 [PubMed: 22229042]

Friston K, Rigoli F, Ognibene D, Mathys C, Fitzgerald T, Pezzulo G. Active inference and epistemic value. Cognitive neuroscience. 2015:1–28. [PubMed: 25185800]

Golman, R., Loewenstein, G. Information Gaps: A Theory of Preferences Regarding the Presence and Absence of Information. Decision, Advance online Publication. 2016. doi: htto://dx.doi.org/10.1037/dec0000068

Gottlieb J. Attention, learning, and the value of information. Neuron. 2012; 76(2):281–295. S0896-6273(12)00888-4 [pii]. DOI: 10.1016/j.neuron.2012.09.034 [PubMed: 23083732]

Gottlieb J, Hayhoe M, Hikosaka O, Rangel A. Attention, reward and information seeking. Journal of Neuroscience. 2014; 34(46):15497–154504. [PubMed: 25392517]

Gottlieb J, Oudeyer PY, Lopes M, Baranes A. Information seeking, curiosity and attention: computational and empirical mechanisms. Trends in Cognitive Science. 2013; 17(11):585–593. DOI: 10.1016/j.tics.2013.09.001

Hanks TD, C S. Perceptual Decision Making in Rodents, Monkeys, and Humans. Neuron. 2017; 93(1): 15–31. [PubMed: 28056343]

Iigaya K, Story GW, Kurth-Nelson Z, Dolan RJ, Dayan P. The modulation of savouring by prediction error and its effects on choice. eLife. 2016 Apr.21(5):e13747. pii. doi: 10.7554/eLife.13747

Johnson L, Sullivan B, Hayhoe M, Ballard DH. Predicting human visuomotor behavior in a driving task. Phil. Trans. R. Soc. B. 2014; 369:20130044. doi: http://dx.doi.org/10.1098/rstb.2013.0044. [PubMed: 24395971]

Kable JW, Glimcher PW. The neurobiology of decision: consensus and controversy. Neuron. 2009; 63(6):733–745. doi: S0896-6273(09)00681-3 [pii]. DOI: 10.1016/j.neuron.2009.09.003 [PubMed: 19778504]

Krauzlis RJ, Lovejoy LP, A Z. Superior colliculus and visual spatial attention. Annual Reviews of Neuroscience. 2013; 36:165–182.

Kuhns AB, Dombert PL, Mengotti P, Fink GR, Vossel S. Spatial Attention, Motor Intention, and Bayesian Cue Predictability in the Human Brain. Journal of Neuroscience. 2017; 37(21):5334–5344. [PubMed: 28450541]

Leong Y, Radulescu A, Daniel R, DeWoskin V, Niv Y. Dynamic Interaction between Reinforcement Learning and Attention in Multidimensional Environments. Neuron. 2017; 93(2):451–463. [PubMed: 28103483]

Maunsell JH. Neuronal representations of cognitive state: reward or attention? Trends Cogn Sci. 2004; 8(6):261–265. [PubMed: 15165551]

Mengotti P, Dombert PL, Fink GR, Vossel S. Disruption of the Right Temporoparietal Junction Impairs Probabilistic Belief Updating. Journal of Neuroscience. 2017; 37(22):5419–5428. [PubMed: 28473647]

Morvan C, Maloney L. Human visual search does not maximize the post-saccadic probability of identifying targets. PLoS Comput Biol. 2012; 8(2):e1002342.doi: 10.1371/journal.pcbi.1002342 [PubMed: 22319428]

Najemnik J, Geisler WS. Optimal eye movement strategies in visual search. Nature. 2005; 434(7031): 387–391. [PubMed: 15772663]

Navalpakkam V, Itti L. Modeling the influence of task on attention. Vision Res. 2005; 45(2):205–231. doi: S0042-6989(04)00392-X [pii]. DOI: 10.1016/j.visres.2004.07.042 [PubMed: 15581921]

Peck CJ, Jangraw DC, Suzuki M, Efem R, Gottlieb J. Reward modulates attention independently of action value in posterior parietal cortex. J Neurosci. 2009; 29(36):11182–11191. doi: 29/36/11182 [pii]. DOI: 10.1523/JNEUROSCI.1929-09.2009 [PubMed: 19741125]

Polonio L, DiGuida S, Coricelli G. Strategic sophistication and attention in games: an eye tracking study. Games and Economic Behavior. 2015; 94:80–96.

Reis R. Inattentive Producers. Review of Economic Studies. 2006; 73:793–821.

Renninger LW, Verghese P, Coughlan J. Where to look next? eye movements reduce locatl uncertainty. Journal of vision. 2007; 7(3):6.

Schwartenbeck P, Fitzgerald T, Dolan R, Friston K. Exploration, novelty, surprise and free energy minimization. Front Psychol. 2013; 4:710.doi: 10.3389/fpsyg.2013.00710 [PubMed: 24109469]

Sims CA. Implications of rational inattention. Journal of Monetary Economics. 2003; 50:665–690.

Sugrue LP, Corrado GS, Newsome WT. Choosing the greater of two goods: neural currencies for valuation and decision making. Nat Rev Neurosci. 2005; 6(5):363–375. [PubMed: 15832198]

Sullivan BT, Johnson L, Rothkopf CA, Ballard D, Hayhoe M. The role of uncertainty and reward on eye movements in a virtual driving task. J. Vis. 2012; 12(13)doi: 10.1167/12.13.19

Tatler BW, Hayhoe MN, Land MF, Ballard DH. Eye guidance in natural vision: reinterpreting salience. J Vis. 2011; 11(5):5–25.

Thompson KG, Bichot NP. A visual salience map in the primate frontal eye field. Prog Brain Res. 2005; 147:251–262. [PubMed: 15581711]

Vossel S, Bauer M, Mathys C, Adams RA, Dolan RJ, Stephan KE, Friston KJ. Cholinergic Stimulation Enhances Bayesian Belief Updating in the Deployment of Spatial Attention. Journal of Neuroscience. 2014; 34(47):15735–15742. [PubMed: 25411501]

Vossel S, C M, J D, M B, J D, J FK, E SK. Spatial Attention, Precision, and Bayesian Inference: A Study of Saccadic Response Speed. Cerebral Cortex. 2014; 24(6):1436–1450. [PubMed: 23322402]

Vossel S, C M, KE S, J FK. Cortical Coupling Reflects Bayesian Belief Updating in the Deployment of Spatial Attention. Journal of Neuroscience. 2015; 35(33):11532–11542. [PubMed: 26290231]

Vossel S, Thiel CM, Fink GR. Cue validity modulates the neural correlates of covert endogenous orienting of attention in parietal and frontal cortex. NeuroImage. 2006; 32(3):1257–1264. [PubMed: 16846742]

White BJ, Berg DJ, K J, Marino RA, Itti L, DP M. Superior colliculus neurons encode a visual saliency map during free viewing of natural dynamic video. Nature Communications. 2017; 8:14263.doi: 10.1038/ncomms14263

Woodford M. Information-Constrained State-Dependent Pricing. Journal of Monetary Economics. 2009; 56(S):100–124.

Yang SC, Lengyel M, Wolpert DM. Active sensing in the categorization of visual patterns. eLife. 2016; :e12215. doi: pii. doi: 10.7554/eLife.12215 [PubMed: 26880546]

Yarbus, AL. Eye Movements and Vision. New York: Plenum; 1967.

Yu AJ, Dayan P. Uncertainty, neuromodulation, and attention. Neuron. 2005; 46(4):681–692. doi: S0896-6273(05)00362-4 [pii]. DOI: 10.1016/j.neuron.2005.04.026 [PubMed: 15944135]

Zentall TR, Stagner JP. Do pigeons prefer information in the absence of differential reinforcement? Learning & Behavior. 2012; doi: 10.3758/s13420-012-0067-5
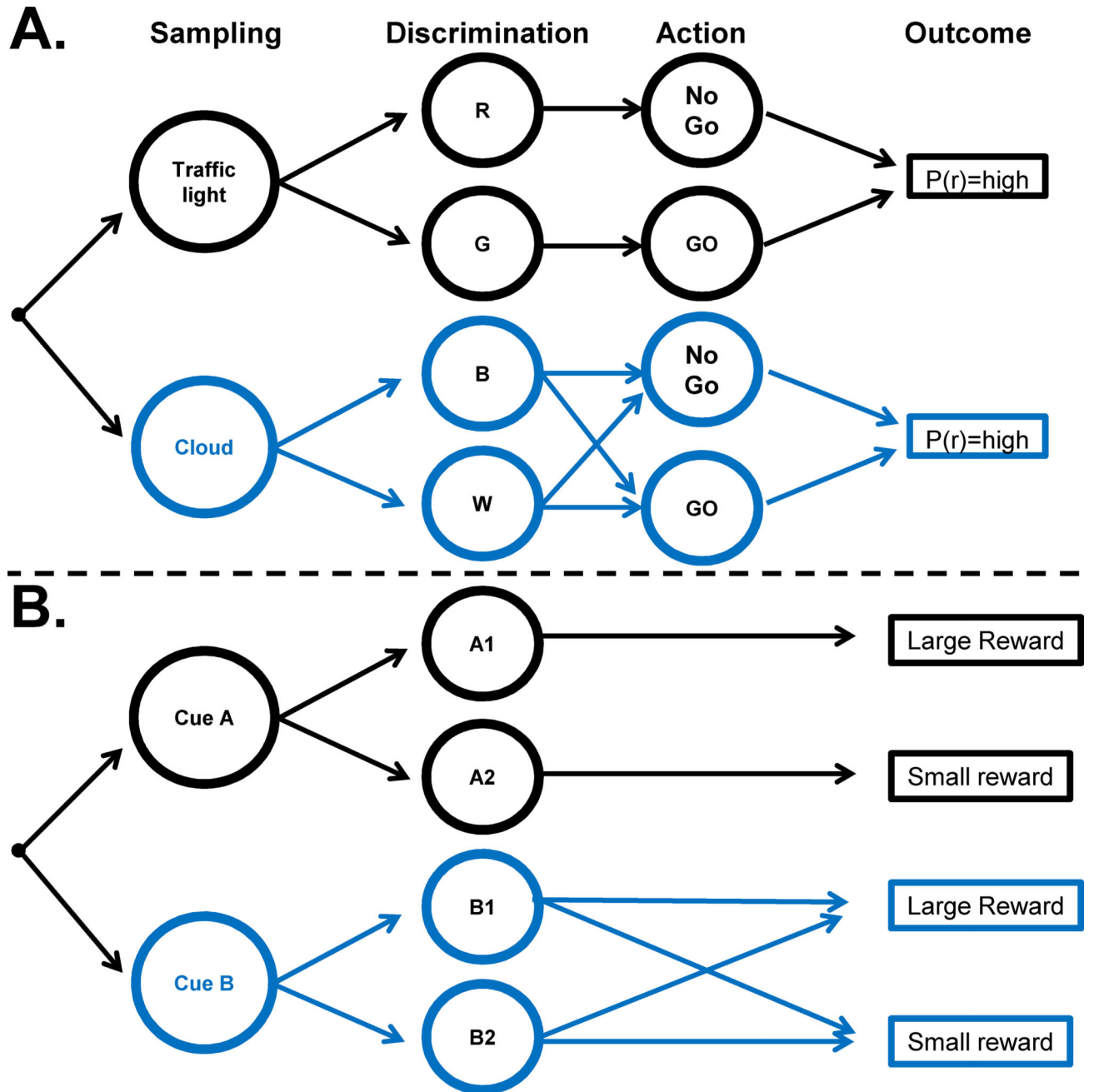
Figure 1. Decision chains for sampling and actions

(A) Instrumental sampling: the agent makes a decision of which cue to sample
("Sampling"), discriminates the properties of the selected cue ("Discrimination"), decides
which action to take based on the discrimination ("Action") and realizes an outcome
("Outcome", or reward, r) with probability (P(r)). In the specific example, a pedestrian
decides whether to sample a traffic light or a cloud, discriminates the colors of the sampled
stimuli (red/green for the light and blue/white for the cloud), and takes the decision to stop
or proceed (NoGo/Go) in order to be safe (reward, r). The Shannon entropy of the possible
actions is high before sampling either cue as well as after sampling the cloud (1 bit if the

Go/NoGo actions are equally likely) but becomes much lower depending on the reliability of the cue (e..g, 0 if the cue produces perfect certainty about the optimal action).

(B) Non-Instrumental Sampling: The cues indicate a pre-ordained outcome but the agent cannot alter the outcome.The agent makes the decision whether to sample cue A or B, and discriminates the signal given by the sampled cue. Signals A1 and A2, produced upon sampling cue A, predict with certainty whether the reward will be large or small. Signals B1 and B2, produced by cue B are random and do not reduce the uncertainty about reward size.
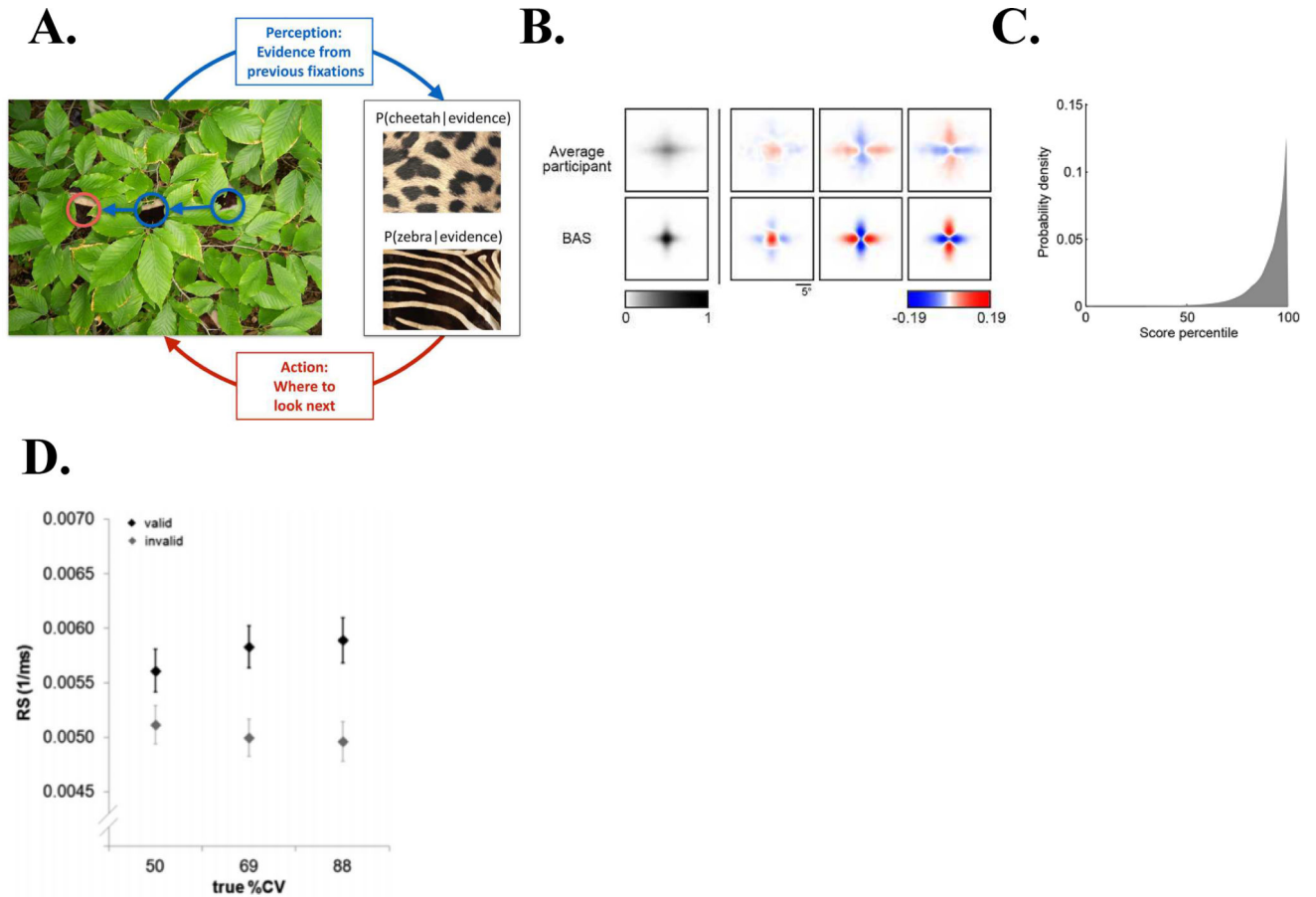
Figure 2. Behavioral measures of information-based policies

(A–C) Saccadic sampling for a categorization decision (**A**) When trying to categorize whether a fur hidden behind foliage (left) belongs to a zebra or a cheetah, evidence from multiple fixations (blue, the visible patches of the fur, and their location in the image) needs to be integrated to generate beliefs about fur category (right, here represented probabilistically, as the posterior probability of the particular animal given the evidence). Given current beliefs, different potential locations in the scene will be expected to have different amounts of informativeness with regard to further distinguishing between the categories, and optimal sensing involves choosing the maximally informative location (red). In the example shown, after the first two fixations (blue) it is ambiguous whether the fur belongs to a zebra or a cheetah, but active sensing chooses a collinearly located revealing position (red) which should be informative and indeed reveals a zebra with high certainty. (**B**) Revealing density maps for participants and the BAS model. The first column shows mean reveal density and the last three columns show mean subtracted densities for each of the three underlying image types (patchy, horizontal stripes, vertical stripes). Bottom: color scales used for all mean densities (left), and for all mean-corrected densities (right). All density maps use the same scale, such that a density of 1 corresponds to the peak mean density across all maps. (**C**) Histogram showing the distribution of percentile values of informativeness (as derived by the BAS algorithm) across all participants, trials and fixations.

(**D) Validity effects in a Posner cueing paradigm.** The difference in mean response speed (RS, the reciprocal of RT) to detect targets that were validly vs invalidly cued increases as a function of the reliability of the cue (%cue validity, %CV). Reproduced with permission from Yang et al., 2016 (A–C) and Vossel et al., 2015 (D).
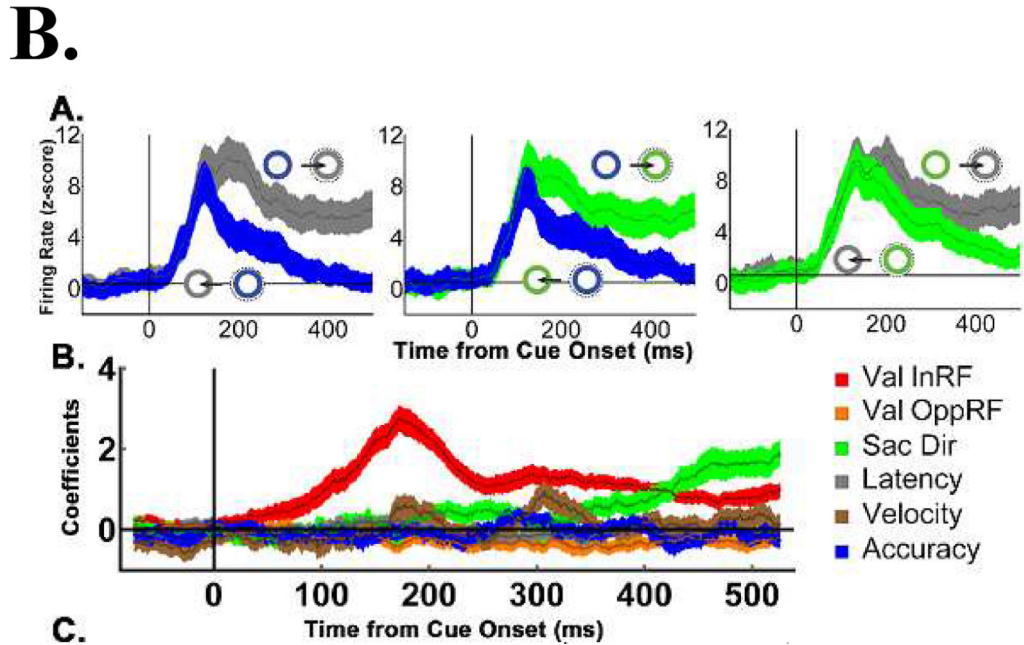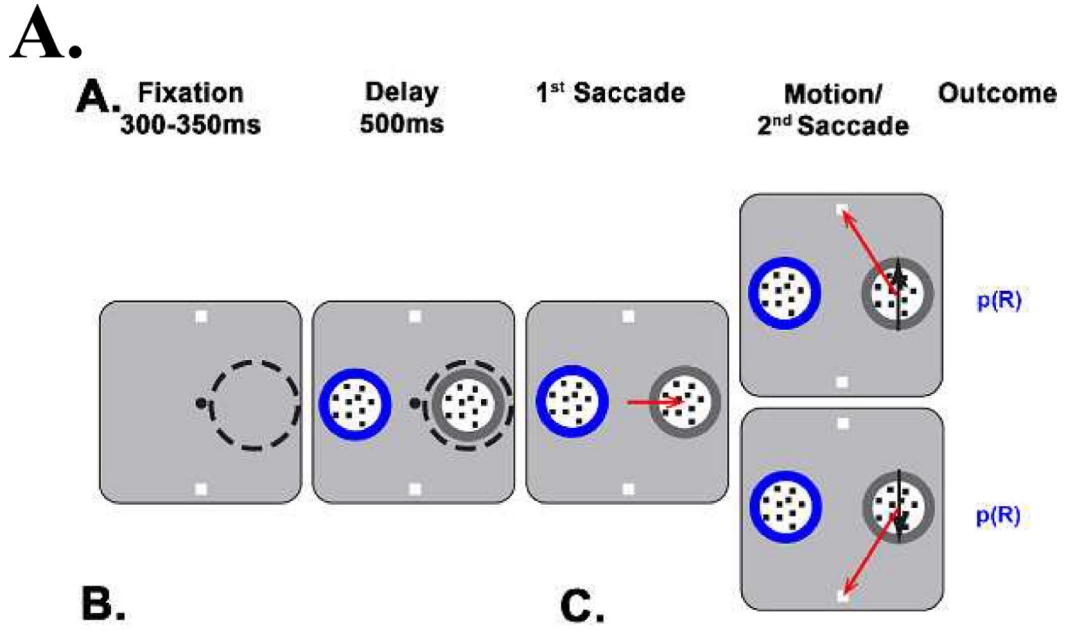
A.

B.

Figure 3. Two step sampling task

(**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 was not visible to the monkey during the experiment.) After the monkey achieved fixation, the display was presented, containing two targets outside the RF (white squares) and two cues of which one was inside the RF and the other 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 a cue of

their choice (third panel, red arrow). At the end of the first saccade, the chosen cue delivered its information in the form of 100% coherent dot motion directed toward one of the targets (last panel, black arrows). After motion onset, the monkeys were free to indicate their final decision by making a second saccade to a target (last panel, red arrows), and the trial ended with a probabilistic reward (p(R)).

(**B**) Population responses on 2-cue trials, sorted according to the difference in validity and the saccade direction. Gray, blue and green traces indicate, respectively, 100%, 80% and 55% valid cues. The cartoons show the cue that was chosen by the monkeys' saccade (the higher validity of the pair) and whether that cue was inside the RF (dashed circle, higher firing rates) or at the opposite location (lower firing rates).. The saccade response (difference in firing rates between the two saccade directions) scaled with relative validity, being highest for the largest validity difference (100% vs 55% cues, left) and lowest when the cues were similar in validity (100% vs 80%, right panel).

(**C**) Time-resolved regression coefficients (sliding window of 50 ms width, 1 ms step) estimating the effects of the validity of the RF cue, the validity of the opposite RF cue, and saccade direction, velocity, latency and accuracy across the trials shown in **A**. Reproduced with permission from Foley et al., 2017.
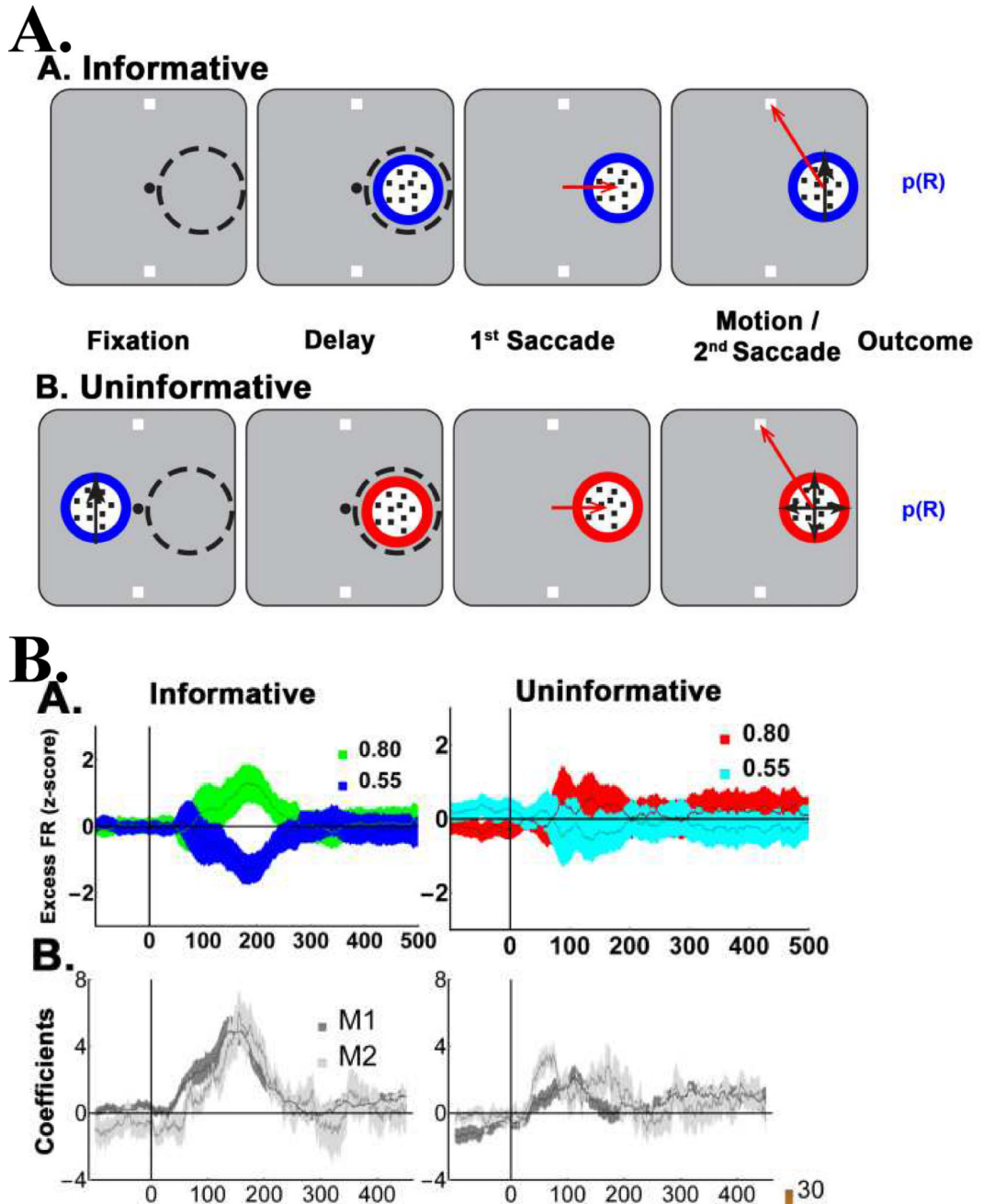
Figure 4. The informative/uninformative stimulus test

(A)Top row: 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. **Bottom row:** In the uninformative condition, a pre-cue 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; leftmost panel). The pre-cue 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 (4th panel). Note that, while 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.

(**B**) LIP neurons encode validity but not the cumulative future rewards of uninformative cues. **Top row:** Average firing rates (n = 69 cells) for 55% and 80% valid cues (left), and their yoked uninformative stimuli (right). 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. Average regression coefficients for the validity/reward responses in informative and uninformative trials for each monkey (M1 and M2). 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. Reproduced with permission from Foley et al., 2017.
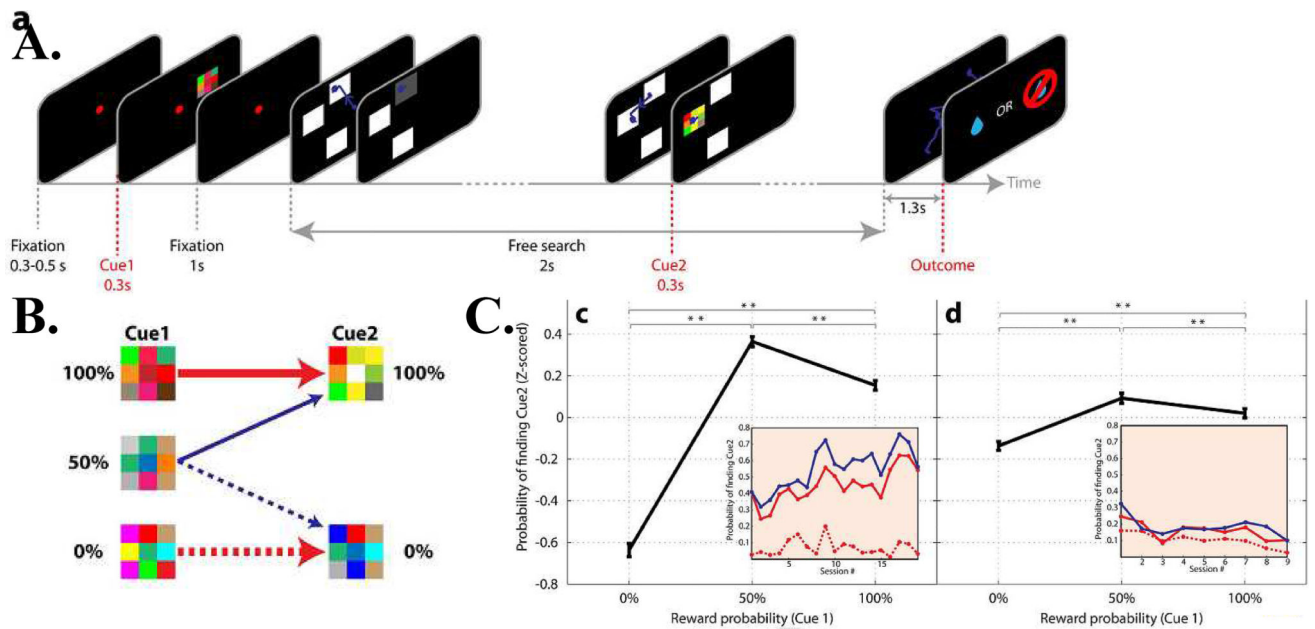
Figure 5. Non instrumental search behavior

(**A**) Task stages. The monkey initiated each trial by fixating a central point and maintaining gaze on it while cue1 was shown in the periphery for 0.3 s. After an additional second, the fixation point was replaced with a search display containing 3 white masks that the monkey could freely scan. If the monkey maintained gaze on a mask for 300 ms, this triggered a reveal of the underlying pattern (a gray square or an additional cue). In the example illustrated, the monkey first uncovered an uninformative gray square and later found cue 2 at the middle location. The search display then disappeared, and after an additional 1.3 s delay (blank screen) the trial ended with a tone that was accompanied by the outcome (a reward or a lack of reward according to the probability signaled by cue 1). The search behavior was entirely unconstrained and had no bearing on the final outcome.

(**B**) Transition statistics between cue 1 (which could signal 0%, 50% or 100% reward likelihood) and cue 2 (which signaled a 0% or 100% probability). If cue 1 signaled 100% or 0% reward, cue 2 merely confirmed this prediction, bringing no new information. If cue 1 signaled a 50% reward likelihood, cue 2 brought new information, and was equally likely to signal a positive or a negative outcome.

(**C**) The probability of finding cue 2 as a function of the prior probability signaled by cue 1, for two subjects (right and left panels). Points show the mean and standard errors of these probabilities, z-scored across all sessions. Stars indicate p < 0.025 (Wilcoxon test). The **insets** in each panel show the average of the raw data per session. The dotted red trace indicates 0% cue 1, the solid red trace shows 100% cue 1 and the solid blue trace shows 50% cue 1. Reproduced with permission from Daddaoua et al., 2016.