

Attention, Learning, and the Value of Information

Jacqueline Gottlieb^{1,2,*}

¹Department of Neuroscience

²The Kavli Institute for Brain Research

Columbia University, New York, NY 10032, USA

*Correspondence: jg2141@columbia.edu

<http://dx.doi.org/10.1016/j.neuron.2012.09.034>

Despite many studies on selective attention, fundamental questions remain about its nature and neural mechanisms. Here I draw from the animal and machine learning fields that describe attention as a mechanism for active learning and uncertainty reduction and explore the implications of this view for understanding visual attention and eye movement control. I propose that a closer integration of these different views has the potential greatly to expand our understanding of oculomotor control and our ability to use this system as a window into high level but poorly understood cognitive functions, including the capacity for curiosity and exploration and for inferring internal models of the external world.

Introduction

Long ago defined by William James as “the focusing of the mind,” selective attention is simultaneously one of our most pervasive and most baffling cognitive functions. On one hand attention is recruited for nearly every behavior and has been investigated in humans, monkeys, mice, and rats. On the other hand despite this wealth of research, significant questions remain about the nature of attention, its purpose and neural mechanisms.

In humans and nonhuman primates, much of our knowledge of the mechanisms of attention comes from the system of vision and eye movement control. Intensive research into this system has shown that attention affects sensory representations at all levels of the visual hierarchy, starting from low-level areas such as the lateral geniculate nucleus, through high-level cortical areas in the inferior temporal lobe (Reynolds and Heeger, 2009; Saalmann and Kastner, 2011). These studies also suggest that the source of attentional modulations lies, at least in part, in sensorimotor areas associated with rapid eye movements (saccades). Two areas that have been particularly well investigated are the lateral intraparietal area and the frontal eye field (shown in Figure 1A for the macaque monkey brain). Neurons in these areas have spatial receptive fields and saccade-related responses and respond selectively to stimuli that are likely to attract attention in a variety of tasks. Not specifically sensory or motor, these cells seem to encode the specific act of target selection, and can provide feedback regarding this selection both to earlier visual areas and to downstream movement structures that generate shifts of gaze.

Important questions remain however, about the significance and computations underlying this target selection response. Historically, two frameworks have been used to explain this response. One line of research describes target selection in motor decision terms, as the integration of evidence toward, and eventual commitment to a shift of gaze (Gold and Shadlen, 2007; Kable and Glimcher, 2009). An alternative interpretation describes it as *stimulus* selection—the act of focusing on a sensory cue that may drive attentional modulations of the sensory response (Bisley and Goldberg, 2010; Gottlieb and Balan, 2010). While earlier studies have attempted to dissect

the visual versus the motor components of target selection, more recent studies have emphasized the decision—free choice—aspect of the saccadic response. However, the decision framework has remained largely separate from an attentional interpretation and it is unclear to what extent the two frameworks are compatible or distinct (Maunsell and Treue, 2006).

In this perspective, I propose a broader approach that integrates elements of both explanations and considers the cognitive aspects of eye movement control. Consistent with the decision framework, I propose that the neural response to target selection can be viewed as an internal decision that seeks to maximize a utility function (i.e., increase a benefit and minimize a cost). However, consistent with an attention interpretation I emphasize that, as a system controlling a *sensory* organ—the eye—this decision must be optimized for sampling *information*. Therefore, the distinction between visual and motor selection, which may seem trivial in sensorimotor terms, becomes highly significant in a decision perspective. To understand oculomotor decisions we must tackle the complex and little understood question of how the brain ascribes value to sources of information, and how this may differ from value determined by primary reward.

The question of active information selection is rarely studied as a distinct topic (and even more rarely in individual cells), but it arises repeatedly in learning and memory research. Recent evidence from computational and behavioral studies makes it clear that processes of information selection tap into some of our highest cognitive functions, involving, among others, intrinsic curiosity and the ability for advance planning and forming internal models of complex tasks (e.g., Gershman and Niv, 2010; Johnson et al., 2012). My goal in this perspective is to consider these processes and their relevance to vision and eye movement control. I begin with a brief overview of target selection responses in monkey frontal and parietal cortex and their relation with attention and eye movement control. I then consider the possible relation between target selection and information selection, drawing particularly on three areas that have been traditionally separate from oculomotor research—namely, studies on associative learning in humans and rats (Holland

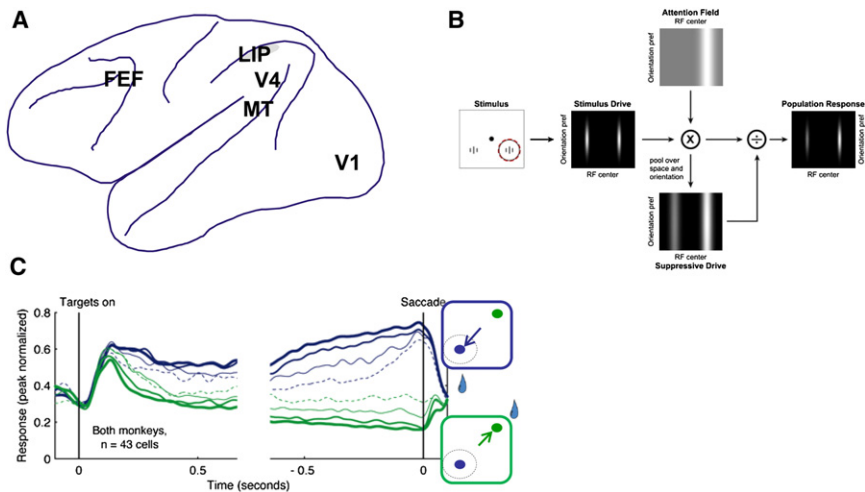


Figure 1. Current Approach to Attention Research

(A) Cortical areas investigated in relation to attention. Lateral view of the macaque monkey cortex showing some of the areas that have been investigated in relation with attention, including primary visual cortex (V1), area V4, the middle temporal area (MT), and two sensorimotor areas, the lateral intraparietal area (LIP) and the frontal eye field (FEF).

(B) Normalization model of attention. The model includes two populations of cells: feature selective neurons that are sensitive to stimulus location and features (e.g., orientation) and respond to a stimulus with both excitatory and suppressive drives (black panels, lower row) and attention neurons that are selective only for location and provide a selective multiplicative gain (gray, top row). Reproduced with permission from Reynolds and Heeger (2009).

(C) Reward sensitive target selection activity in a decision task. Monkeys were trained to direct gaze to one of two possible targets for receipt of

a juice reward, and the targets were placed so as to fall inside or opposite the receptive field of an intraparietal cell (dashed oval). Traces show the average responses of a cell population, aligned on the time of target presentation and the monkeys' subsequent choice. The neurons encoded the direction of the chosen saccade, responding more for saccades directed toward versus away from the receptive field (blue versus green). Directional selectivity however became stronger as a function of the difference in expected reward (dotted, thin solid, and thick solid traces show progressively larger differences in expected reward). Reproduced with permission from Sugrue et al. (2004).

and Maddux, 2010; Le Pelley, 2010; Pearce and Mackintosh, 2010); studies of eye movement control in natural behaviors (Hayhoe and Ballard, 2005; Tatler et al., 2011), and computational studies in the machine learning field (Dayan and Daw, 2008; Dayan et al., 2000; Oudeyer et al., 2007).

Because of the complexity and vastness of the topic, my discussion will be necessarily incomplete. I will eschew circuit-level mechanisms (most of which are currently unknown), and detailed mathematical considerations (for which excellent descriptions can be found elsewhere [Dayan and Daw, 2008; Dayan et al., 2000; Oudeyer et al., 2007]). Despite these limitations however, I hope that it will become clear in the forthcoming discussion that appreciating the cognitive dimensions of eye movement control is both a necessity and a source of strength. Gaining this appreciation is necessary for explaining a range of observations regarding the neural responses to target selection, which have no good explanation in sensory or motor terms. More importantly perhaps, broadening our perspective will strengthen the field of oculomotor research and allow us to use the full power of this system as a window into high-level but poorly understood cognitive functions.

From Vision to Eye Movements, and an Intermediate Stage

Research on selective attention in humans and nonhuman primates spans numerous studies, using a vast array of psychophysical and neurophysiological techniques. While these studies differ widely in their specific details, many share the common feature that they direct subjects to attend to a specific item—be it an object, feature, or location—and measure the effects of attentional selection on perception or action. These studies have shown that attention produces widespread effects throughout early and late visual areas, which collectively increase the signal from the attended item and suppress noise from unattended distractors (Reynolds and Heeger, 2009). A shift of attention can remain covert—generating only an im-

provement in perceptual discrimination—or can be accompanied by saccades—rapid eye movements that place the fovea on the attended item. The oculomotor component of an attentional response is generated by a network of cortical and subcortical structures that includes portions of the basal ganglia, the superior colliculus, and the frontal eye field (Schall et al., 2011; Stanford et al., 2010).

Neurophysiological studies have also shown that, interposed between visual processing and saccade production is an intermediate layer of target selection, which has been most intensively investigated in the frontal eye field and the lateral intraparietal area (Figure 1A). A large fraction of neurons in these areas have spatial receptive fields and respond both to visual stimuli and/or to a planned saccade. Rather than being selective for a visual features, these cells encode a more abstract quantity of target selection—i.e., discriminate between targets and distractors in a variety of tasks (Gottlieb and Balan, 2010; Thompson and Bichot, 2005). Experiments that manipulate the salience or relevance of visual cues show that target selection cells respond very selectively to stimuli that are likely to be attended, either because of their physical salience or behavioral relevance (Gottlieb et al., 1998; Thompson and Bichot, 2005). Experiments that dissociate visual selection from motor output show that neural responses to target selection can be flexibly linked with action—for example, being coupled with a shift of gaze, with a skeletal response or with no immediate motor action (Balan et al., 2008; Bisley and Goldberg, 2003; Schall et al., 2011). Experiments involving direct manipulations (i.e., through microstimulation or reversible inactivation) show that these two areas produce both feedforward effects—specifying potential plans for a saccadic response—and feedback influences—driving the perceptual effects of attention that are expressed either in visual neural responses (Moore and Armstrong, 2003; Noudoost and Moore, 2011) or in psychophysical reports (Balan and Gottlieb, 2009; Wardak et al., 2006; Wardak et al., 2004).

The Missing Link

Having thoroughly characterized the target selection response, these studies set the stage for tackling the next critical question: how does the brain generate this selective response, and how do parietal and frontal cells “know” where to attend (Baluch and Itti, 2011)? Surprisingly, despite the wealth of attention research, few studies have addressed this question. To appreciate this gap, let us consider three classes of computational models that synthesize empirical findings on various aspects of selective attention.

One substantial body of investigation has examined the sensorimotor transformation for eye movement control—the chain of events through which visual selection generates an eye movement response. Recent models synthesizing these findings have proposed a process of gated accumulation, whereby the accumulation of information in saccade movement cells is insulated from visual selection unless (or until) an eye movement becomes appropriate (Lo and Wang, 2006; Purcell et al., 2012; Schall et al., 2011). The model captures a host of findings related to visual and motor selection and the brain’s ability flexibly to link attention with action. However, the model does not attempt to explain target selection itself; it simply asks how visual selection, once it has been generated, gives rise to an overt saccade.

A similar stance is adopted by models focusing on sensory responses, which ask how parietal or frontal signals of target selection may produce sensory attentional effects. A recent “normalization” model of attention has been particularly successful in explaining a large number of sensory effects using a simple biologically-plausible circuit (Reynolds and Heeger, 2009). As illustrated in Figure 1B, the model proposes that a spatially selective “attention field” is fed back to the visual system and multiplicatively scales visual inputs in spatially specific fashion. Followed by divisive normalization based on local competition with other visual inputs (“suppressive drive”), this attentional influence results in a biased visual representation where the attended stimulus is more strongly represented (“population response”). The “attention field” conforms to the properties of the target selection response—i.e., it is sensitive to spatial location but not visual features. However, this drive is portrayed as a box with an output but no inputs; in other words, the model focuses on its sensory effects, but not on how the drive is itself generated.

And finally, a similar stance is adopted by models describing the links between attention and decision formation. A common theme in these models is that attention influences the accumulation of evidence toward the attended option, making the subject more likely to select that option (Krajovich et al., 2010). These models begin by assuming that attention exists, but do not explain how it may come to be—e.g., why subjects may attend to a specific object in the first place.

These computational efforts therefore, reflecting the state of the art in empirical research, uniformly treat attention as an external bias term. They portray attention as a “cognitive force” that has widespread influences on perception and action but which is itself external to, rather than emergent from, these latter functions.

Target Selection as a Value Representation

A notable exception to this theoretical stance comes from an unexpected source—a line of studies that have not addressed

attention per se but have used the eye movement system as an experimental platform for studying decision formation. These studies start from the premise that the ultimate goal of any act of selection is to maximize an organism’s biological fitness. Therefore it seems likely that, as specific types of selection, eye movements and attention would also satisfy a utility function—i.e., seek to maximize a benefit and minimize a cost. Guided by this idea, decision studies have trained monkeys to choose between eye movement targets that deliver various amounts of juice reward. By placing the targets inside and opposite the receptive field of a target selective cell, these studies evoke the target selection response and study its properties to gain insight into decision formation.

A consistent outcome revealed by these investigations (which have been typically carried out in the lateral intraparietal area) is that the signal of target selection is not stereotyped but increases as a function of the relative desirability of the alternative options (Kable and Glimcher, 2009; Sugrue et al., 2005). An example of this result is shown in Figure 1C in a task where monkeys had to choose between two alternative targets whose payoffs varied dynamically from trial to trial (Sugrue et al., 2004). Monkeys apportioned their choices in proportion to the recent history of reward, and neurons in the lateral intraparietal area increased their selective responses in proportion with the target’s expected reward: firing for a saccade directed toward the receptive field increased monotonically (blue traces, dashed to solid), while firing for a saccade to a different location decreased monotonically as a function of reward expectation (green traces). Similar results are obtained in tasks that manipulate the desirability of a target using different methods, for example by controlling the relative magnitude, probability or delay of its expected reward (Bernacchia et al., 2011; Louie et al., 2011; Sugrue et al., 2004; Yang and Shadlen, 2007). Taken together these studies suggest the powerful hypothesis that target selection neurons encode the relative value of alternative actions, and that they integrate multiple sources of evidence pertinent to this estimation.

This utility-based view of target selection is particularly attractive not only because of its parsimony and elegance, but also because it has straightforward theoretical interpretations in economic and reinforcement learning terms. The computational framework of reinforcement learning, originally developed in the machine learning field (Sutton and Barto, 1998), has been particularly successful in explaining behavioral and neuronal results. The core idea in this framework is that agents (be they animals or machines) constantly estimate the values of alternative options based on their repeated experience with these options. This intuition is captured in the Rescorla-Wagner equation, which states that the estimated value at time t (V_t) is based on the estimate at the previous step (V_{t-1}) plus a small learning term ($\beta * \delta$):

$$V_t = V_{t-1} + \beta * \delta \quad (\text{Equation 1})$$

As described above, parietal neurons encoding target selection are thought to report an action value representation—the term V in the Rescorla-Wagner equation—and to update this representation in dynamic fashion (Sugrue et al., 2004). This

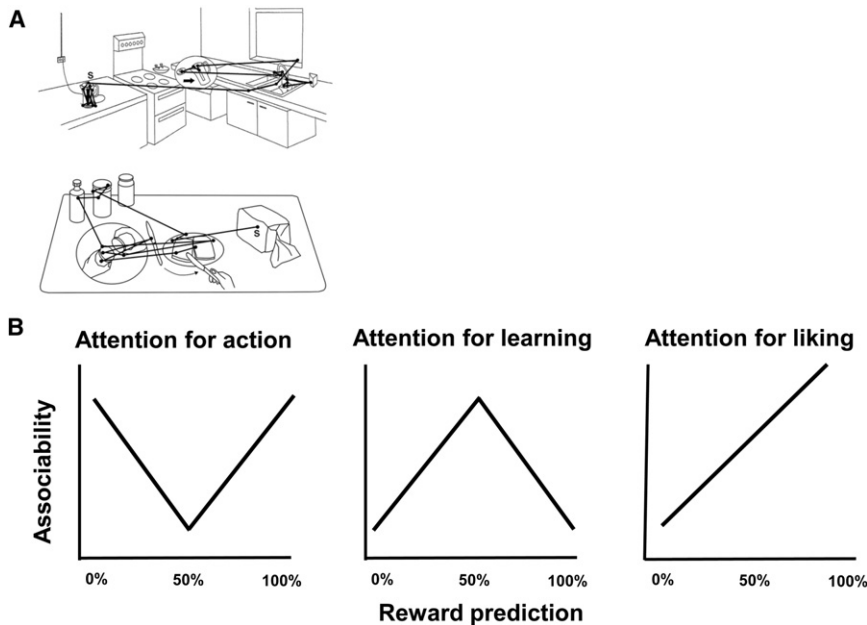


Figure 2. Attention as Information Selection

(A) Gaze behavior in naturalistic tasks where a subject fills a kettle for preparing tea (top) or prepares a peanut butter sandwich (bottom). Gaze is directed to task relevant locations that reduce the subject's uncertainty, and precede the skeletal actions. Reproduced with permission from Land (2009).

(B) Three putative attentional mechanisms that assign associability according to the reliability (left), uncertainty (middle), or reward probability (right) predicted by a cue.

value response could then be used by downstream motor mechanisms such as those in the basal ganglia or the superior colliculus, to select optimal (reward maximizing) actions.

The right-hand—learning—term in the equation in turn has been more closely linked with modulatory systems, in particular noradrenaline and dopamine, and is composed of two quantities. One quantity, β , is a learning rate that takes values between 0 and 1 and determines how quickly the agent updates its predictions. This rate may depend on global task properties such as the volatility or uncertainty of a given task and could be conveyed through neuromodulation (Cohen et al., 2007; Nassar et al., 2012).

The second quantity is the *prediction error* term (δ), which describes how “surprised” the agent is by a particular outcome—i.e., how well or poorly it had predicted that outcome. This quantity, defined as the difference between the agent's estimate and the actual outcome at the previous step ($\delta = r - V_{t-1}$), provides a trigger for learning—updating expectations so as to reduce future errors in prediction. A by-now classic series of results suggests that the reward prediction error is encoded by midbrain dopamine cells (Glimcher, 2011; Schultz, 2006; Waelti et al., 2001). An example of this prediction error response is shown in Figure 3B, in an experiment in which monkeys were initially uncertain about the size of a reward and at the time marked “Cue” received a visual signal that conveyed information about the expected reward (Bromberg-Martin and Hikosaka, 2009). Dopamine cells had a transient excitatory response to a stimulus that signaled a larger-than-expected reward (“Info-big”) and a transient inhibition to a stimulus that signaled a lower-than-expected reward (“Info-small”) but had nearly no response to a stimulus that provided no new information (“Rand,” blue traces). When the actual reward was delivered (“Reward”) the cells again had excitatory and inhibitory responses to, respectively, high or low reward, but only if these

reward were unexpected (“Rand,” but not “Info” conditions) precisely as expected from a prediction error term. As shown by the Rescorla-Wagner equation, such a signal of unexpected outcomes can drive an agent to increase or decrease its value estimates if the outcome it has experienced was, respectively, higher or lower than expected.

Taken together, these findings reveal a remarkable confluence between com-

putational and empirical results. They suggest an integrated account of learning and decision formation, whereby value representations are maintained in cortical and sensorimotor structures and are dynamically updated based on feedback from dopaminergic cells (Kable and Glimcher, 2009; Sugrue et al., 2005).

Eye Movements Select Information

Casting target selection as an internal value estimation would seem to bridge the conceptual gap in attention research. A straightforward implication of this idea is that, to decide where to shift gaze or where to attend, the brain may simply keep track of the values of the alternative options and make choices according to this value representation. A key challenge in making this link however, concerns the specific value that has been considered in the decision field. As I described in the preceding section, in all current studies of decision formation “value” is defined in terms of primary reward: the value of a saccade target in a laboratory task is defined by the juice that the monkey obtains by making the saccade (Figure 1C). In natural behavior however, eye movements rarely harvest primary reward. Instead, they sample *information*.

Consider for example the eye movements made by a subject in two everyday tasks—preparing a peanut butter sandwich or filling up a kettle to prepare some tea (Figure 2A). Like the monkey in a decision experiment, these subjects seek a reward—i.e., a sandwich or a cup of tea. Unlike the monkey, however, their rewards will not be realized by merely looking at a spot, no matter how intense their attention may be. Rather, the subjects use attention and gaze as intermediate steps that allow them to acquire *information*, which will only indirectly guide their future actions.

Computational studies of naturalistic behaviors show that the act of acquiring information—whether it is overt or remains

internal to the brain—may indeed have material value, as it increases the chance of success of a future action (Tatler et al., 2011). However, these studies also show that the processes required to compute information value differ markedly from those that have been so far considered in decision tasks. A salient property of this process is that information value depends critically on the subjects' uncertainty and, in the Rescorla-Wagner equation is more closely related with the right side of the equation—the act of *learning* or *modifying expectations*. As a simple illustration of this distinction, consider again the tea-making task in Figure 2B. To prepare and consume her tea, the subject must make both arm and leg actions, and in the reinforcement equation both actions would be assigned a high value term (V). The subject's gaze, however, is very selectively allocated to the targets of the arm and not the leg actions. This selectivity cannot be explained in terms of action value alone but reflects the fact that the arm movements have higher uncertainty and thus more to gain from new information. Thus, the drive that motivates a shift of gaze is not value per se but the need to learn—i.e., to update one's predictions through new information.

Independent support for a view of attention as a learning mechanism comes from an area of research that has been mostly separate from the oculomotor field (but see Le Pelley, 2010) but has directly addressed the cognitive aspects of information selection—namely, the question of how subjects learn from and about sensory cues (Pearce and Mackintosh, 2010). A central finding emerging from these studies is that subjects estimate the reliability of a sensory stimulus based on their prior experience with that stimulus and use this knowledge to modulate their future learning based on that cue. In the Rescorla-Wagner equation this process is implemented using an *associability* parameter, α , which is a stimulus-specific *learning rate* (Pearce and Mackintosh, 2010):

$$V_t = V_{t-1} + \alpha * \beta * \delta \quad (\text{Equation 2})$$

While, as we have seen above, the standard learning rate β is applied globally to a context or task, associability is a property of an individual cue and can differentially weight the available cues. As I discuss in detail in the following sections, this apparently simple modification entails a complex, hierarchical learning mechanism. It entails an executive process which, having previously learned the predictive validity of a sensory cue, guides the moment by moment information selection—i.e., has in effect learnt how to learn.

A final line of evidence for the information-bound nature of eye movement control comes from single-neuron studies of target selection that dissociate shifts of attention from overt shifts of gaze (Gottlieb and Balan, 2010). An example of such a study is the experiment shown in Figure 4A, in which we trained monkeys to report the orientation of a peripheral target (a right- or left-facing letter “E”) by releasing a bar (Oristaglio et al., 2006). Monkeys had to perform the task while maintaining their gaze straight ahead (on the central fixation point), so that overt saccades had *no value* and would have been punished with a loss of reward—and indeed, monkeys actively suppressed the saccades. Nevertheless the informative cue had value, and

neurons in the lateral intraparietal area continued selecting the cue, showing much higher activity if the “E” rather than a distractor was in their receptive field (Balan and Gottlieb, 2009; Balan et al., 2008; Oristaglio et al., 2006; Figure 4B). These neural responses are in some respect not surprising because the capacity for covert attention has been well-established in psychophysical research, and its correlates are found also in the frontal eye field (Schall et al., 2011; Thompson et al., 2005). However the findings are highly significant from a decision perspective: they highlight the fact that the decision variable for target selection hinges not on the value of a motor action, but on the properties of a sensory cue.

In sum, three lines of investigation conducted in very different fields—studies of eye movement control in natural behaviors, associative learning in humans and rats and target selection in the frontal and parietal lobes—converge on a common point. All these studies indicate that to understand oculomotor decisions we must describe how the brain assigns value to sources of information. What might this process entail?

Three Types of Attention

A useful way of organizing the discussion starts from the proposal advanced in the associative learning field that the brain has several types of attention mechanism. These systems are thought to have different neuronal substrates and to serve different behavioral roles and are dubbed, respectively “attention for action,” “attention for learning,” and “attention for liking.”

To gain an intuitive understanding of these types of attention, consider a hypothetical experiment in which you have a 50% prior probability of receiving a reward, and on each trial are shown a sensory cue that provides information about the trial's reward (Figure 2B). Some cues bring perfect information, indicating that you will definitely receive or not receive a reward (100% or 0% likelihood). Other cues make uncertain predictions, e.g., that you have a 50% chance of reward. This set of sensory cues can be characterized along two dimensions. One is the *expected reward* of the cue, which is defined as the product of reward magnitude and probability, and increases monotonically along the x axis. The second dimension is the *variance* or reliability of the cue's predictions. Variance is an inverted V-shaped function with a peak for the 50% cue (Figure 2B, center). The inverse of variance (reliability) has an upright-V profile, with a minimum at the 50% cue and maxima for 0% or 100% predictors (Figure 2B, left).

The associability hypothesis postulates that the systems of “attention for action” and “attention for learning” assign weight based, respectively, on the reliability and variance of a cue's predictions (Pearce and Mackintosh, 2010). As shown in the left panel of Figure 2B, the system of “attention for action” is thought to assign low weight (associability) to cues that predict an uncertain reward, but a *high* weight for cues that make consistent predictions. This system would enable an animal to attend to a familiar cue that makes consistent predictions, such as a traffic light at an intersection. The system of “attention for learning” on the other hand (Figure 2B, center) has the opposite weighting and assigns priority to an uncertain or variable cue (Pearce and Mackintosh, 2010). This system would enable an animal to attend to novel and uncertain stimuli such as a new

sign in a storefront. Importantly however, both systems are value-neutral in the sense that they do not depend on expected reward: they give equal weight to stimuli predicting low or high reward, provided these make equally reliable predictions.

The third system of “attention for liking” differs qualitatively from the first two because it assigns priority simply in proportion to the associated reward, directing more resources to a “good news” (100%) relative to a “bad news” (0%) cue (Figure 2B, right). Although not originally proposed in associative learning research, converging behavioral and neural observations bring strong evidence supporting this system (Hogarth et al., 2010; Vuilleumier, 2005).

In the following sections I discuss each system in turn, considering questions related to their implementation and contrasting the associability-based explanation with related proposals from the reinforcement learning field.

Attention for Action: Reliability, Relevance, and Reward

Although not typically discussed in relation with eye movement control, the system of “attention for action” that is proposed in studies of associative learning maps naturally on the purposive, task-related eye movements made by subjects in everyday tasks (e.g., Figure 2A). Quantitative studies show that practically all the eye movements made in naturalistic goal-directed behaviors can be interpreted as acquiring information to guide a forthcoming action (Tatler et al., 2011). According to the associability idea, to achieve this type of control, the brain will explicitly learn (and potentially represent) the reliability of the predictions generated by a cue (Pearce and Mackintosh, 2010). An alternative explanation, however, emerges from studies of eye movements in natural behaviors, which suggest that the value of an eye movement lies in reducing uncertainty and increasing the expected reward (probability of success) of a future action (Ballard and Hayhoe, 2009; Hayhoe et al., 2012; Rothkopf et al., 2007; Tatler et al., 2011). I consider the relationship between these ideas and their possible neural implementation.

Reliability

While support for the reliability hypothesis comes from behavioral and neuropsychological studies in humans and rats (Holland and Maddux, 2010; Pearce and Mackintosh, 2010), a key open question at the present time is whether (and how) reliability is encoded in individual cells. Perhaps the strongest neural evidence supporting this idea comes from studies of sensory perception, which show that the strength (signal to noise) of a sensory input can mediate a reliability-based form of sensory integration. For example, in tasks where monkeys are trained to estimate their heading direction based on a combination of vestibular and visual motion cues, the relative influence of the visual cue increases in proportion with the signal to noise of its motion signal. A number of studies have proposed ways in which stimulus strength, reflected in the width and strength of its sensory responses, can mediate optimal reliability-based cue integration (Fetsch et al., 2012; Ma et al., 2008; Vilares and Kording, 2011).

It is unclear, however, whether the brain encodes the more cognitive type of reliability that is postulated by the associative learning field, which is not embedded in the stimulus itself but requires learning of complex relationships between the stimulus

and the predicted events. This is the type of reliability that we may ascribe, for example, to a weather forecast, to the advice we receive from our physician or to an economic indicator. While a recent study using an “information choice task” proposed that this type of reliability is encoded in midbrain dopaminergic cells (Bromberg-Martin and Hikosaka, 2009), the findings remain open to alternative interpretations.

In the “information choice task” used by Bromberg-Martin and Hikosaka, monkeys began each trial with a 50% probability of obtaining a large or a small reward and were given the opportunity to obtain advance information about the size of the reward. As shown in Figure 3A, if the monkeys shifted gaze to one of the available targets (dubbed the “informative” target), this target gave way to one of the cues that reliably predicted whether the trial will yield a large reward or in a small reward (“Info”). However, if monkeys shifted gaze to the unreliable target (“Rand” in Figure 3A), this target produced a distinct set of subsequent cues that conveyed only uncertain (50%) information about the future reward. Notably, the reward outcomes themselves were on average equal and fixed in all conditions, so that monkeys could not increase their physical reward with a specific choice. Nevertheless, monkeys reliably selected the informative target suggesting that they had an intrinsic preference for information.

Dopamine neurons (Figure 3B) had two types of responses on the task. At the time of the actual information (marked “Cue”) in Figure 3B, the neurons emitted the customary prediction error response which, as described above, was excitatory for a “good news” (big reward) cue and inhibitory for a “bad news” (small reward) cue. Of particular interest however was a response that preceded the actual cue and seemed to signal the expected information. This response arose at the time of the monkeys’ selection (marked “Target” in Figure 3B) and was slightly stronger if the trial included an informative rather than an uninformative cue (red versus blue traces). This early response seems to signal a superordinate property of “informativeness” (or reliability) that is independent of a specific message, and to correspond to the monkeys’ behavioral preference for the informative cue.

Unfortunately however, because the information in this task was *about* a primary reward, the results do not conclusively rule out alternative explanations based on this reward. It is well known that monkeys modulate their anticipatory licking based on stimulus-reward associations and will stop licking when observing a low-reward cue (Fiorillo et al., 2003). In addition as I mentioned above, subjects direct attention based on stimulus-reward associations, and may have gazed for longer periods at the high-reward versus the low-reward cue (e.g., the green cross versus green wave in Figure 3A; Hogarth et al., 2010). It remains therefore possible that by selecting the informative cue the monkeys did not specifically seek information but simply sought to minimize their effort (by avoiding having to lick for or look at a low-reward pattern) or perhaps to bring about the motivationally salient, high-reward pattern (Beierholm and Dayan, 2010). At this time therefore it remains an open question whether the brain has a bona fide reliability representation.

Internal Models, Uncertainty, and Information

Rather than searching for an “intrinsic” preference for information, studies of eye movements in natural behaviors have

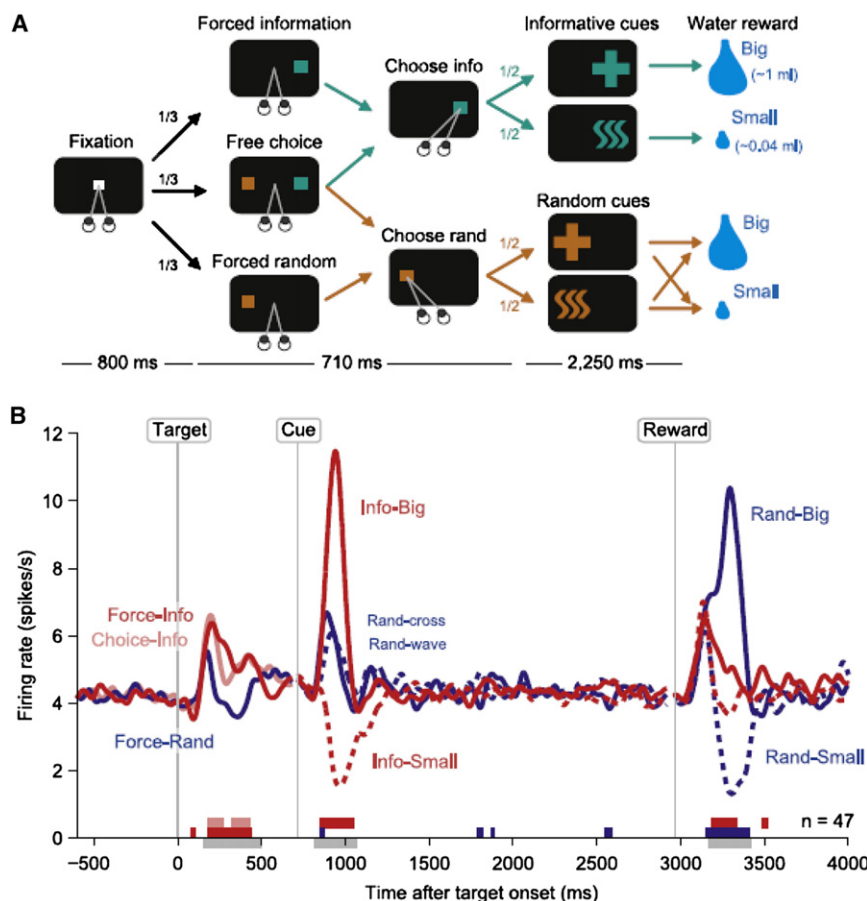


Figure 3. Dopamine Neuron Responses in an Information Choice Task

(A) On each trial after achieving central fixation monkeys viewed a target prefacing an informative (green) or uninformative (orange) cue. Single target trials (top and bottom) were interleaved with two-target trials where monkeys were free to select the target they wished to view. If monkeys shifted gaze to the informative target (green) they were shown two subsequent cues that were consistently associated with, respectively, a large or small water reward. If monkeys shifted gaze to the uninformative target (orange) they were shown two other cues that were inconsistently associated with the large or small reward (50% predictive validity). The large and small reward were equally likely to occur, so that the informative and uninformative targets had equal expected reward.

(B) Neural responses of DA cells on the information choice task. The traces show average activity in a population of DA cells, aligned on the time of target presentation, appearance of the reward cues and delivery of the final reward. At the time of target presentation the neurons had stronger responses when the display contained an informative target (dark and light red traces) than when it only contained the uninformative target (blue). After the information was revealed (cue) DA neurons had the expected reward prediction response. At the time of cue presentation they had excitatory and inhibitory responses to, respectively, the high and low reward predictive pattern, and small excitatory responses to the uncertain pattern announcing a 50% probability of reward. At the time of the reward, the neurons had excitatory and inhibitory responses upon receipt of, respectively, the large and small reward, but only if this reward was unpredicted (i.e., upon selection of the uninformative cue). Reproduced with permission from (Bromberg-Martin and Hikosaka, 2009).

adopted a more pragmatic approach and attempt to estimate the material value that an eye movement may bring (Hayhoe and Ballard, 2005; Tatler et al., 2011). The studies make use of so-called Markov decision chains—mathematical methods that allow one to formulate a task description as a sequence of steps and estimate the cumulative future reward that can be expected by traversing these steps. By including an estimate of the uncertainty that arises at each step, one can further calculate the costs of this uncertainty and the benefits of reducing it by obtaining information (Dayan and Daw, 2008; Rothkopf and Ballard, 2010; Sprague and Ballard, 2005; Tatler et al., 2011). For instance, in the tea-making task, one can calculate how uncertain one is about one's position and distance from the faucet, and what the benefit would be of reducing that uncertainty through a shift of gaze. These studies have shown how, when applied to complex tasks (such as an agent walking through an environment while avoiding obstacles and picking up litter) these methods can be applied to identify the uncertainty and informational requirements of intermediate steps (Rothkopf and Ballard, 2010; Sprague and Ballard, 2005).

However, even as they demonstrate the feasibility of this computation, the studies show that information selection can be remarkably complex. Most of the complexity derives from the fact that, because the benefits of information are only indirect, computing its value requires planning across a sequence

of steps. Moreover, this planning requires not only a simple knowledge of the order of various steps, but a sophisticated model of the task structure that specifies the hidden (causal) relationships between consecutive steps. Consider for example the simple act of directing gaze to the water faucet while preparing a tea (Figure 2A). To generate this apparently trivial act, the brain must know not only that the faucet is associated with the task (after all, so are the kitchen floor and the walls) but that lifting the handle will cause the water to flow, which in turn will have a determining influence on preparing the tea. In other words, to determine which sources of uncertainty should be optimally resolved, the brain must know which steps are causal or predictive of the future outcome (Gershman and Niv, 2010). In a simple scenario such as making a tea this computation may be greatly aided by extensive practice. In other behaviors, however, it requires much more difficult inferences on longer time scales. It can be prohibitively complex for example, to determine which one of the available stimuli is informative if one lands on Mars, or which economic indicator is truly consequential for a future outlook.

Converging evidence shows that humans indeed infer hidden models of complex tasks (Acuña and Schrater, 2010; Braun et al., 2010; Daw et al., 2011; Gershman and Niv, 2010; Yakushiji and Jacobs, 2011), and indirect evidence from tasks involving schemas or contextual associations suggests that lower

animals may also possess this capacity (Balan and Gottlieb, 2006; Braun et al., 2010; Johnson et al., 2012). Building internal models that identify the relevant steps is critical for specifying what subset of a very high-dimensional information stream should be considered at a given time. Such models, in other words, are necessary for deciding to what to attend. As mentioned above in relation with the associability equation (Equation 2), this process entails an executive mechanism that *learns how to learn*—that is, decides how to organize the moment by moment sampling of sensory information. The need for hierarchical learning has been discussed in relation to motor control and cognitive tasks (Braun et al., 2010; Johnson et al., 2012) and, as it is clear from this discussion, is also at the heart of attention control.

Given an appropriate model of a task structure, informative options (stimuli or actions) may be identified through a prediction error mechanism as those options which, by reducing uncertainty, increase the expected future reward. Importantly, however, the reward prediction errors that have been traditionally considered in dopamine cells are *model-free* quantities that only register changes in value between consecutive time steps (Niv and Schoenbaum, 2008). Such a mechanism can be prohibitively slow in complex tasks, and may erroneously assign credit to irrelevant steps (Rothkopf et al., 2007). A model free system for example may conclude that the decision to wear a white shirt was critical for obtaining a high grade on a test, simply because this decision was closer in time to the actual exam relative to the earlier act of studying for the exam. Recent evidence from functional imaging experiments in humans suggests that dopamine cells and their recipient structures also encode *model-based* prediction errors that take into account future actions (Daw et al., 2011; Morris et al., 2006; Takahashi et al., 2011) suggesting a potential involvement in model-based mechanisms. As I discuss in the final section, the distinction between model-free and model-based computations is fundamental and may explain key differences between an “attention for action” and “attention for liking” mechanism.

Executive Control and Target Selection

Although the neural mechanisms computing relevance are very poorly understood, lesion studies in monkeys and rats suggest that they depend on the frontal lobes. The studies implicate the dorsolateral prefrontal cortex, the anterior cingulate cortex, and the orbital frontal cortex in this computation (sometimes referred to as a “credit assignment” computation) (Kolling et al., 2012; Rossi et al., 2009; Rushworth et al., 2011; Walton et al., 2011) and suggest that these areas may convey the results to dopaminergic cells (Takahashi et al., 2011). Interestingly, converging evidence suggests that the parietal target selection response, which reflects the moment by moment deployment of attention, has a number of complex properties that may reflect an interface with executive mechanisms (Gottlieb and Snyder, 2010).

A good illustration of these complex properties comes from an experiment that I mentioned above, where we trained monkeys to report the orientation of a visual target by releasing a bar (Oristaglio et al., 2006). The task required monkeys to find a relevant target using covert attention as described above (Figure 4A) and in addition to apply a learnt stimulus-action association namely, to release a bar held in their right paw if the instructive

cue was oriented to the right (an “E”) or a bar held in the left paw if it was oriented to the left (a “3”). The task therefore did not require monkeys to orient to the attended location but rather report the information at that location using an arbitrary (symbolic) action, much as one would step on the brake when seeing a red traffic light or step on the gas when seeing a green light.

As I mentioned in the previous section, parietal neurons encoded the location of the relevant cue, and some of the cells had *only* a target selection response, responding more if the cue rather than a distractor was in the RF regardless of the manual release (Figure 4B). These simple spatial responses are consistent with the traditional view of attention control, whereby the top-down drive contains only spatial and not non-spatial information (e.g., Figure 1B). A sizeable fraction of cells however showed a combinatorial coding of both the attended location and the bar release. Some of the cells, like that shown in Figure 4C, responded selectively if the “E” was in their receptive field *and* instructed release of the left bar; other cells had the complementary preference, responding best if the “E” was in their receptive field *and* instructed release of the left bar (not shown). These manual modulations were not free-standing limb motor responses but modulatory effects on visual selection (i.e., the effects were not seen if a distractor appeared in the receptive field; Figure 4C, right), a conclusion consistent with the later finding that reversible inactivation produced visual but not skeletal motor defects (Balan and Gottlieb, 2009).

These findings are difficult to explain in a purely visual framework that casts target selection as a disembodied bias term (Figure 1B). They are also puzzling in an action based framework that asks whether parietal areas are involved in skeletal or ocular actions (Snyder et al., 2000). However, neural responses with combinatorial (mixed) properties are hallmarks of goal-directed cognitive control (Rigotti et al., 2010), and in the context of information selection may embody the bank of knowledge that is necessary for selecting cues. These results therefore raise the important question of how target selection interfaces with frontal processes of executive control and with visual learning mechanisms that assign meaning to visual cues (Albright, 2012; Freedman and Assad, 2011; Mirabella et al., 2007). One important question is what these complex responses imply for the nature of top-down control. Is the attentional feedback from the parietal lobe only carried by neurons with simple spatial responses, consistent with current assumptions that it only carries spatial information (e.g., Figure 1B)? Or, alternatively, does the top-down feedback carry higher bandwidth information regarding both stimuli and actions, conveyed by neurons with combined responses (Baluch and Itti, 2011)? A second question concerns the sophistication of the information conveyed by this combinatorial code: does this code reflect only coincidental associations between stimuli and contexts or actions, or do they reflect internal models of multielement tasks?

In sum, the preceding discussion has highlighted some of the complexities that can be entailed by a shift of gaze. Far from requiring a mere direct or habitual sensorimotor link, computing an effective scan path for sampling information requires an executive mechanism that infers the relevant steps in an extend task, and uses this inference to determine points of significant

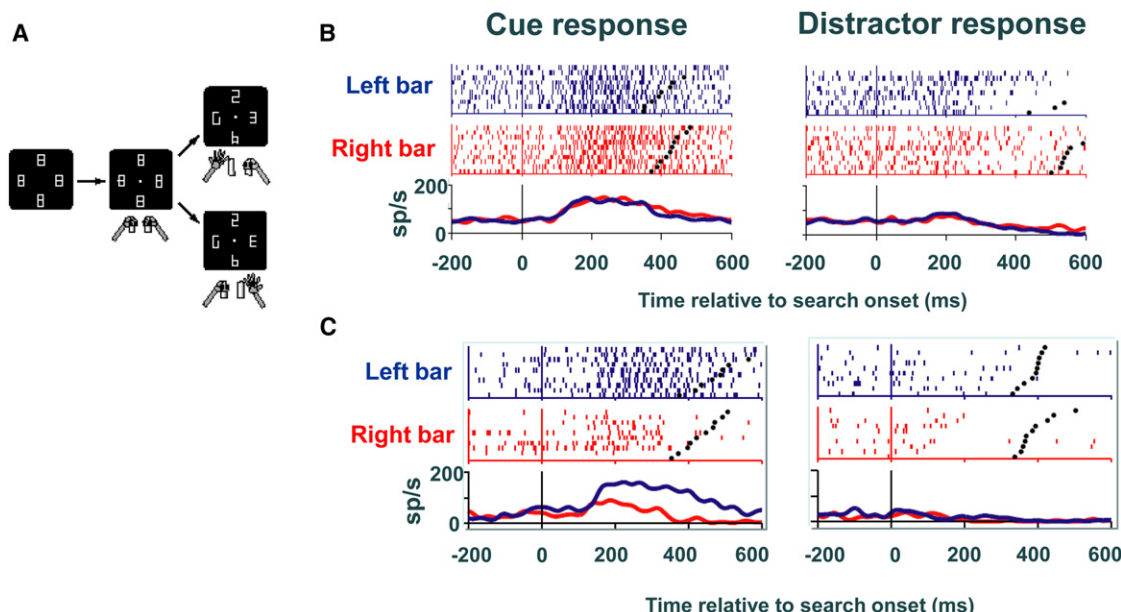


Figure 4. Lateral Intraparietal Neurons Combine Responses to Visual Selection and Visuomanual Associations

(A) Search task. An array of several figure-8 placeholders remained stable on the screen at all times. To begin a trial monkeys directed their eye to the central fixation point (dot) and grabbed two response bars. The search display was then revealed, and contained a cue (a right or left-facing letter “E”) that appeared at an unpredictable location in among letter-like distractors. Monkeys were trained to continue holding central fixation and release a bar held in the right or left hand to indicate whether the “E” was facing, respectively, to the right or to the left.

(B) A parietal neuron that was sensitive only to cue location. The panels show the activity of a lateral intraparietal neuron aligned on the time of target onset. In each row of action potential, the time of the manual release marked by a black dot. Left and right panels are sorted according to the location of the “E.” Blue and red traces refer to trials in which the “E” required release of, respectively, the left or right bar.

(C) A neuron sensitive to both cue location and manual release. The neuron encoded “E” location but was modulated by the manual release, responding more strongly if the monkey released the left rather than the right bar. Reproduced with permission from (Oristaglio et al., 2006).

uncertainty and sources of information that may reduce that uncertainty.

Attending to the Unknown

While the preceding discussion has focused on target selection in familiar tasks, an equally important and possibly more difficult role of attention is to discover new information – learn about new predictors that were previously unknown. We almost instinctively orient to a new sign in a store front or to a strange bird perched on a tree, and in laboratory tasks, gaze is drawn to novel or uncertain stimuli in familiar scenes (Brockmole and Henderson, 2005a, 2005b; Yang et al., 2009). As described in Figure 2B, studies of associative learning propose that exploratory attention is mediated by a separate system of “attention for learning” which, in contrast with “attention for action,” allocates resources to *uncertain* rather than reliable cues (Figure 2B, center panel). Model-based accounts however, suggest that this distinction may not be quite as clear cut, and that, even when the brain orients toward uncertain cues, it is with the goal of learning or reducing the uncertainty regarding that cue.

It has been previously noted that to generate adaptive exploration the brain must distinguish between at least two types of uncertainty (Oudeyer et al., 2007; Payzan-LeNestour and Bos-saerts, 2011; Yu and Dayan, 2005). *Reducible uncertainty* is due to the observer’s imperfect knowledge and can be eliminated by acquiring information—for example when we hear an

ambulance siren and turn to find out where it is. *Irreducible uncertainty* by contrast is built into a task and cannot be reduced through the observers’ effort—as in the case of white noise on a television screen. If “attention for learning” is specifically guided by reducible uncertainty (as it would optimally be) its goal need not be fundamentally different from that of an action-based mechanism. Neither form of attention values uncertainty per se. Instead, both may be information-seeking mechanisms that detect the presence of uncertainty and devise strategies for reducing that uncertainty (Dayan and Daw, 2008).

A difficult question however is how the brain distinguishes between reducible and irreducible uncertainty, as this is not a priori specified. When conducting scientific research, for example, humans are faced with vast sources of uncertainty which, despite significant effort, we are yet to resolve. What determines whether we continue our search or conclude that this is a fruitless task?

Several intriguing solutions have been proposed to this question in the machine learning field. One solution, emerging from the field of developmental robotics, is that the brain generates intrinsic reward when it senses learning progress (i.e., a decline in prediction errors over time) (Oudeyer et al., 2007). This mechanism may motivate learning even in the absence of an external reward, and has been very effective in producing curiosity-like behaviors—whereby robots remain spontaneously interested in activities of intermediate complexity where they improve their predictions but disengage from random (unlearnable) or from

overlearned and “boring” tasks. An alternative account is that goal directed exploration is not motivated by learning progress but by reward expectations that are generalized based on prior experience (P. Dayan, personal communication). For example, when deciding which experiment to pursue we may infer based on past knowledge that a particular approach will be more effective. Interestingly, this form of generalization may call upon the same executive mechanisms of “learning to learn” that we discussed the previous section: to generalize effectively the brain must recognize and compare the relevant (significant) aspects of the different tasks (Bavelier et al., 2012).

In addition to processes that generate targeted information search, exploratory mechanisms almost invariably include simpler strategies, based on random action selection or hard-wired heuristics. For instance, novelty has been proposed to act as an exploration bonus in reward seeking tasks (Wittmann et al., 2008) and to be encoded in dopamine cells as an intrinsic bonus for exploration (Redgrave and Gurney, 2006). This raises the possibility that other forms of automatic attention that are produced by salience or surprise (Boehnke et al., 2011; Karacan and Hayhoe, 2008; Wittmann et al., 2008), rather than being mere weaknesses of a control mechanism, are vital heuristics for allocating resources in very uncertain conditions, when the brain has not yet learnt how to learn.

Neuropsychological studies in rats suggest that task-related and exploratory attention rely on separate neural circuits that involve, respectively, the medial frontal cortex (Maddux and Holland, 2011) versus the substantia nigra, amygdala and the parietal lobe (Maddux et al., 2007). It would be of great interest to know whether this distinction also holds in the monkey and how it is expressed in individual cells—i.e., whether the frontal eye field mediates a system of “attention for action” while the parietal lobe is more closely related with an exploratory mechanism. Neural responses to uncertainty or surprise have been reported in multiple structures (den Ouden et al., 2010; Fiorillo et al., 2003; Kepecs et al., 2008; McCoy and Platt, 2005; O'Neill and Schultz, 2010; Preusschoff et al., 2006, 2008; Schultz et al., 2008; So and Stuphorn, 2012; Tobler et al., 2009) and have been linked with variables such as arousal, anxiety, risk preference, or global learning rates (Nassar et al., 2012; Preusschoff and Bossaerts, 2007). An important question is how these responses are related with selective attention and with the processes computing the uncertainty or information value of *specific* cues.

Attention for Liking

The final system shown in Figure 2B is the system of “attention for liking,” whereby subjects preferentially direct attention to pleasurable or high reward cues. Although not guided by reliability or expected information, this form of attention is a powerful mechanism, which automatically draws resources to stimuli that have intrinsic emotional or conditioned associations (Damaraju et al., 2009; Flagel et al., 2011; Hickey et al., 2010a, 2010b; Hogarth et al., 2010; Della Libera and Chelazzi, 2009; Vuilleumier, 2005). These attentional influences are difficult to overcome and may underlie maladaptive reactions in psychiatric disorders, such as the enhanced susceptibility of addicted patients to drug-related cues (Flagel et al., 2011).

The neural substrates of emotional attention are not very well understood, but a recent experiment in our laboratory suggests that they include the parietal lobe. The experiment, illustrated in Figure 5, tested how attention and parietal activity are influenced by stimuli that convey positive or negative reward information but do not instruct the monkey as to an appropriate action (Peck et al., 2009).

Monkeys began each trial with a 50% prior probability of reward and, at the onset of a trial were shown a reward cue—a conditioned stimulus that signaled whether the current trial will end in a reward (CS+) or a lack of reward (CS−) (Figure 5A). However, while the CS reliably signaled a 50% increase or a decrease in expected reward relative to prior expectations, they did not indicate the required action. To successfully complete the trial and progress to the next, monkeys had to make a saccade to an independent target that appeared after the disappearance of the CS and was located randomly either at the same or at the opposite location. An incorrect trial (where monkeys did not look at the target) was immediately repeated until correctly completed. This allowed us to distinguish between attentional orienting to the relevant target or to the initial, reward-predicting CS. An attention system that directs resources in goal-directed fashion would assign priority to the target regardless of the CS; by contrast, a system of “attention for liking” may automatically orient based on the value of the CS.

The behavioral and neural results revealed the influence of both mechanisms. In most trials monkeys accurately directed gaze to the target, showing that they had learnt its significance. This learning however was not perfect, and saccades were also biased by the preceding CS. The strongest effect was for saccades following a low-reward cue (CS−) (Figure 5D). If the target happened to appear at the location that had been occupied by a CS−, the monkeys' saccades had longer reaction times and lower accuracy relative to saccades to other locations. Notably, this interference was not due to lower motivation but was spatially specific, showing that attention was inhibited specifically at the CS− location.

This behavioral bias in the monkeys' saccades was correlated with CS evoked responses in the parietal lobe (Figures 5B and 5C). After presentation of a CS+ or CS− in their receptive field, lateral intraparietal neurons had a transient visual response that was higher for a positive relative to a negative cue, consistently with previously reported reward modulations (Figure 5B, blue versus red trace). Surprisingly, however—given that the CS had no action relevance—the neurons maintained a sustained response to the CS during the ensuing delay. Moreover, as seen for the behavioral effect, this persistent response did not reflect global changes in arousal or motivation, but a spatial bias toward or away from the CS location. Sustained activity following a CS+ was higher at the cue location relative to the opposite location, suggesting that attention lingered at the CS+ location (Figure 5C, top, black versus gray trace). By contrast, sustained activity following a CS− was *lower* at the cue's location relative to the opposite location (Figure 5C, bottom), consistent with the behavioral suppression at the CS− location. The CS− evoked inhibition interfered with the monkeys' performance and lowered their rate of reward. Nevertheless, the

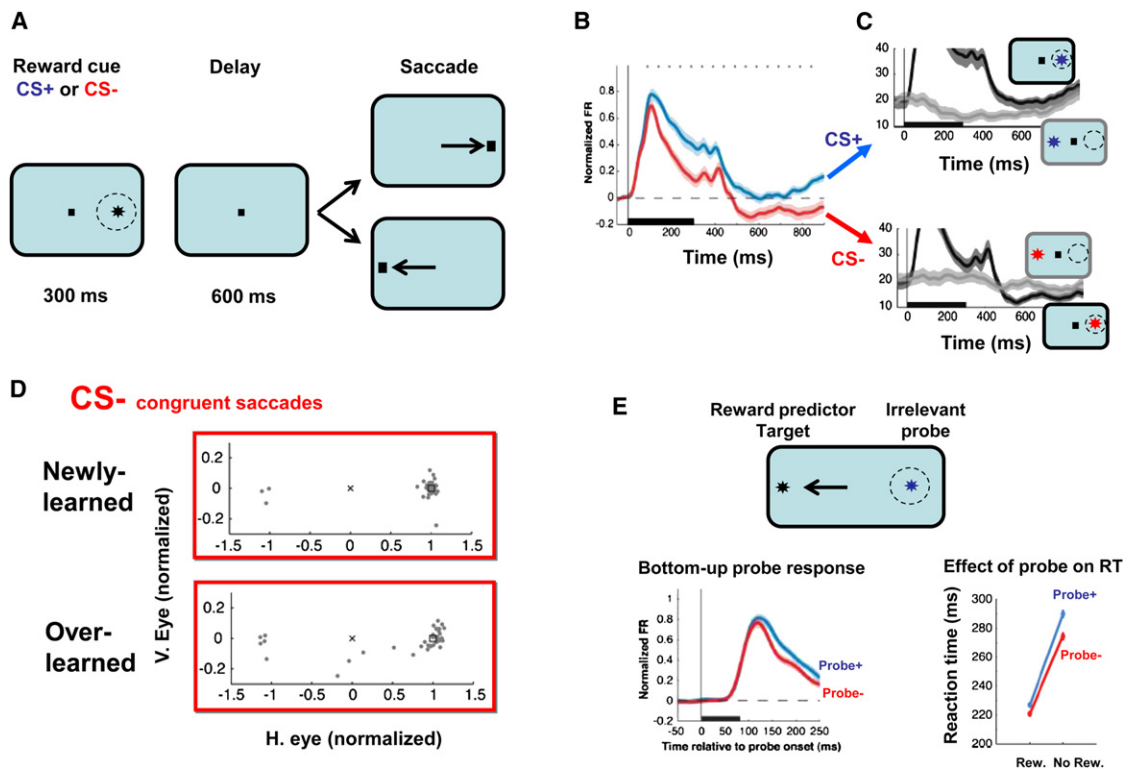


Figure 5. Pavlovian Attention in the Lateral Intraparietal Area

(A) Behavioral task. Each trial has a 50% prior probability of ending in a reward. After monkeys achieved central fixation a peripheral cue was flashed for 300 ms either inside the neuron's receptive field (dashed oval) or at the opposite location. Cues were abstract colored patterns that signaled with certainty whether the trial will receive a reward (CS+) or no reward (CS-). After a 600 ms delay period a second target appeared unpredictably at the same or opposite location relative to the CS and monkeys have to make an immediate saccade to this target to receive the outcome announced by the CS. An error trial is immediately repeated until correctly completed, so that monkeys have to perform each trial to progress in the task.

(B) Parietal responses to the reward cues. When a CS appeared in the receptive field the population of cells showed transient and sustained responses that were selective for cue value, being stronger for a positive cue predicting a reward (CS+, blue) relative to a negative cue predicting no reward (CS-, red). The stars show time bins with a significant difference between the two conditions. The bottom dashed line shows the pre-cue level of activity. Shading shows the standard error of the mean.

(C) CS-evoked responses were spatially specific. The dark traces in each panel show responses when the CS appeared in the receptive field and the gray traces, responses when the CS appeared at the opposite location. (The dark traces are the same as, respectively, the blue and red traces in A, but are shown on an expanded vertical axis.) Responses evoked by a receptive field cue are higher than (CS+) or lower than (CS-) those at the opposite, non-stimulated location, showing that they reflected a spatial bias, and not a global change in motivation.

(D) Saccadic effects of CS- cues. Eye movements in a representative session on unrewarded trials when the saccade target was spatially congruent with a CS-. The location of the CS and target is normalized as if falling horizontally on the right (coordinates of (1,0)) and each gray dot shows the endpoint of a single saccade. The bottom panel shows saccades that followed highly familiar, overlearned CS- (corresponding to the neural responses shown in B). The top panel shows responses on trial with newly learned CS- that were introduced and trained within a single session. Measurement of anticipatory licking showed that monkeys learned the value of the novel CS within the first 5–10 trials, and data collection began after this learning was complete. Presentation of a CS- impaired saccade accuracy if the target happened to be congruent with the CS- location, and the impairment was stronger for overlearned relative to newly learned CS-.

(E) Overtraining produces plastic changes in the visual response. Bottom-up responses to the trained CS were tested in a separate control condition where the previously trained CS were flashed as task-irrelevant probes. In this condition a first predictive CS and the saccade target appeared opposite the receptive field (top panel). Simultaneous with presentation of the saccade target a previously trained CS (the probe) was flashed briefly in the receptive field. The probes had prior reward associations but did not predict reward on these trials. For an overtrained pattern, the bottom-up response remained selective to previous reward associations (bottom left). This value dependent visual response produced differential interference with the saccade, as shown in the bottom right panel. Saccade reaction times (RT) were longer in the presence of a positive relative to a negative probe (blue versus red), reflecting the stronger interference by the positive pattern. Note that RT were longer on unrewarded relative to rewarded trials, showing that monkeys correctly inferred reward probability based on the first predictive CS (that had appeared opposite the receptive field) and not based on the irrelevant probe. Modified with permission from Peck et al. (2009).

effects grew rather than abating with training and, in both neural responses and behavior, were larger after familiar relative to novel CS (Figure 5D, bottom versus top). Moreover, after prolonged training the effects seemed to involve plasticity of the early visual response, since they became insensitive to context and automatically transferred to a different task in which the pre-trained CS no longer predicted reward (Figure 5E).

These findings describe a correlate of “attention for liking” phenomena described in behavioral research, whereby attention is automatically biased by the reward (conditioned) stimulus associations. The findings are consistent with several—not mutually exclusive—mechanisms. One possibility is that they are related to the phenomenon of inhibition of return, whereby attention is inhibited from revisiting recently examined locations

(e.g., Mirpour et al., 2009). A related possibility is that they reflect specific reinforcement mechanisms. The value-dependent orienting described in Figure 5 may arise through a modulation of visual activity by a dopamine reward prediction error response (e.g., Figure 3B) which, like the responses in the parietal lobe, is excitatory for a positive and inhibitory for a negative reward predictive cue. This modulation may also differ from that underlying goal-directed control in that it acts in model-free rather than model-based fashion. As I discussed in the previous section, a model-based allocation would assign priority to the target in the Peck et al. (2009) task, since this was the stimulus that was informative for the future action. A model-free mechanism by contrast would assign priority to the initial CS, since this was the stimulus that signaled a change in reward expectations.

Regardless of the specific answers to these questions (which remain to be determined by future research), the findings highlight the critical point that reward may influence attention through several *distinct* mechanisms. A goal-directed mechanism assigns value to stimuli based on their relevance to future actions, while an agnostic system simply prioritizes stimuli that signal changes in reward expectation.

Why would the brain possess an automatic “attention for liking” mechanism, if this can produce maladaptive effects? This question, which arises here in the context of emotional attention, can be equally applied to other forms of automatic orienting such as those based on salience, novelty or surprise, which can also interfere with ongoing tasks. The answer to this question is not fully known, but an important consideration may be the difficulty of an optimal (model-based) computation. As we have seen in the preceding sections, computing information value optimally is a costly and time-consuming operation that requires inference and advance planning for multiple future steps, and can itself be suboptimal in complex tasks (Wilson and Niv, 2011). Automatic forms of attention by contrast are based on much simpler heuristics. Therefore, the brain may have retained these systems as vital and useful tools for rapidly allocating resources to potentially significant information.

Conclusion: Who Needs Attention?

While all living organisms take actions that bring biological reward, a unique hallmark of higher intelligence is a vast capacity for learning and prediction (Friston, 2010). Here, I proposed that selective attention is intimately linked with these prediction mechanisms. I have argued that attention is the core cognitive system that mediates our *active* search for information—whether information is sought for a foreseeable, well-practiced action or in a more open-ended, exploratory fashion.

While this view is consistent with reinforcement learning research, it is not well integrated with studies of oculomotor control. A closer integration would be beneficial on several counts. First, as I described in the earlier sections, this integration has become necessary for understanding core open questions in attention control—i.e., how the brain decides when and to what to attend. To understand this question—as well as complex properties of the target selection response—we will need to understand the visual learning mechanisms by which the brain assigns meaning to visual cues, and the cognitive systems that assign value to these cues.

Second, by appreciating the cognitive dimension of eye movement control we can begin use the full power of this system as a window into cognitive function. As mentioned in the opening sections, existing research has used the oculomotor system to study cognitive variables involved in decision formation but have interpreted the results in a highly simplified framework of sensorimotor transformation. For example in a well-known motion discrimination paradigm, the direction of motion of a sensory cue is thought to be discriminated by cells in the middle temporal area, while lateral intraparietal cells select the appropriate action (e.g., a specific saccade) (Gold and Shadlen, 2007). This framework therefore explains oculomotor decisions as a sensory-to-motor transfer without invoking the concept of selective attention.

The need for selective attention, however, becomes clear when we consider that, in addition to analyzing visual information the brain must solve another highly complex task—namely, determine the significance and value of that information. As I have discussed above, this requires the brain to estimate its uncertainty and the ability of sensory cues to reduce that uncertainty. The processes involved in this selection include building internal models of external events, guiding behavior based on curiosity and exploration, and generating (and controlling) emotional biases in information processing. Some of these processes have been studied in behavioral paradigms and, by recognizing their tight links with selective attention we can use the oculomotor system to gain insight into their cellular substrates.

ACKNOWLEDGMENTS

I am deeply indebted to Peter Dayan and Mary Hayhoe for their detailed comments on several rounds of this manuscript, and to Eric Kandel and Tom Albright for their guidance in the final stages of its preparation. I also thank members of my laboratory, in particular Nicholas Foley, Himanshu Mhatre, and Adrien Baranes for their comments on several versions of this paper. Work from my own laboratory that is described in this research was supported by The National Eye Institute, The National Institute of Mental Health, The National Institute for Drug Abuse, The Keck Foundation, the McKnight Fund for Neuroscience, The Klingenstein Fund for Neuroscience, the Sloan Foundation, the National Alliance for Research on Schizophrenia and Depression, and the Gatsby Charitable Foundation.

REFERENCES

- Acuña, D.E., and Schrater, P. (2010). Structure learning in human sequential decision-making. *PLoS Comput. Biol.* 6, e1001003.
- Albright, T.D. (2012). On the perception of probable things: neural substrates of associative memory, imagery, and perception. *Neuron* 74, 227–245.
- Balan, P.F., and Gottlieb, J. (2006). Integration of exogenous input into a dynamic salience map revealed by perturbing attention. *J. Neurosci.* 26, 9239–9249.
- Balan, P.F., and Gottlieb, J. (2009). Functional significance of nonspatial information in monkey lateral intraparietal area. *J. Neurosci.* 29, 8166–8176.
- Balan, P.F., Oristaglio, J., Schneider, D.M., and Gottlieb, J. (2008). Neuronal correlates of the set-size effect in monkey lateral intraparietal area. *PLoS Biol.* 6, e158.
- Ballard, D.H., and Hayhoe, M.M. (2009). Modelling the role of task in the control of gaze. *Vis. Cogn.* 17, 1185–1204.
- Baluch, F., and Itti, L. (2011). Mechanisms of top-down attention. *Trends Neurosci.* 34, 210–224.

- Bavelier, D., Green, C.S., Pouget, A., and Schrater, P. (2012). Brain plasticity through the life span: learning to learn and action video games. *Annu. Rev. Neurosci.* 27, 391–416.
- Beierholm, U.R., and Dayan, P. (2010). Pavlovian-instrumental interaction in 'observing behavior'. *PLoS Comput. Biol.* 6, e1000903. <http://dx.doi.org/10.1371/journal.pcbi.1000903>.
- Bernacchia, A., Seo, H., Lee, D., and Wang, X.J. (2011). A reservoir of time constants for memory traces in cortical neurons. *Nat. Neurosci.* 14, 366–372.
- Bisley, J.W., and Goldberg, M.E. (2003). Neuronal activity in the lateral intraparietal area and spatial attention. *Science* 299, 81–86.
- Bisley, J.W., and Goldberg, M.E. (2010). Attention, intention, and priority in the parietal lobe. *Annu. Rev. Neurosci.* 33, 1–21.
- Boehne, S.E., Berg, D.J., Marino, R.A., Baldi, P.F., Itti, L., and Munoz, D.P. (2011). Visual adaptation and novelty responses in the superior colliculus. *Eur. J. Neurosci.* 34, 766–779.
- Braun, D.A., Mehring, C., and Wolpert, D.M. (2010). Structure learning in action. *Behav. Brain Res.* 206, 157–165.
- Brockmole, J.R., and Henderson, J.M. (2005a). Object appearance, disappearance, and attention prioritization in real-world scenes. *Psychon. Bull. Rev.* 12, 1061–1067.
- Brockmole, J.R., and Henderson, J.M. (2005b). Prioritization of new objects in real-world scenes: evidence from eye movements. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 857–868.
- Bromberg-Martin, E.S., and Hikosaka, O. (2009). Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron* 63, 119–126.
- Cohen, J.D., McClure, S.M., and Yu, A.J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362, 933–942.
- Damaraju, E., Huang, Y.M., Barrett, L.F., and Pessoa, L. (2009). Affective learning enhances activity and functional connectivity in early visual cortex. *Neuropsychologia* 47, 2480–2487.
- Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., and Dolan, R.J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69, 1204–1215.
- Dayan, P., and Daw, N.D. (2008). Decision theory, reinforcement learning, and the brain. *Cogn. Affect. Behav. Neurosci.* 8, 429–453.
- Dayan, P., Kakade, S., and Montague, P.R. (2000). Learning and selective attention. *Nat. Neurosci. Suppl.* 3, 1218–1223.
- Della Libera, C., and Chelazzi, L. (2009). Learning to attend and to ignore is a matter of gains and losses. *Psychol. Sci.* 20, 778–784.
- den Ouden, H.E., Daunizeau, J., Roiser, J., Friston, K.J., and Stephan, K.E. (2010). Striatal prediction error modulates cortical coupling. *J. Neurosci.* 30, 3210–3219.
- Fetsch, C.R., Pouget, A., DeAngelis, G.C., and Angelaki, D.E. (2012). Neural correlates of reliability-based cue weighting during multisensory integration. *Nat. Neurosci.* 15, 146–154.
- Fiorillo, C.D., Tobler, P.N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898–1902.
- Flagel, S.B., Clark, J.J., Robinson, T.E., Mayo, L., Czuj, A., Willuhn, I., Akers, C.A., Clinton, S.M., Phillips, P.E., and Akil, H. (2011). A selective role for dopamine in stimulus-reward learning. *Nature* 469, 53–57.
- Freedman, D.J., and Assad, J.A. (2011). A proposed common neural mechanism for categorization and perceptual decisions. *Nat. Neurosci.* 14, 143–146.
- Friston, K. (2010). Is the free-energy principle neurocentric? *Nat. Rev. Neurosci.* 11, 605–606.
- Gershman, S.J., and Niv, Y. (2010). Learning latent structure: carving nature at its joints. *Curr. Opin. Neurobiol.* 20, 251–256.
- Glimcher, P.W. (2011). Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proc. Natl. Acad. Sci. USA* 108(Suppl 3), 15647–15654.
- Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535–574.
- Gottlieb, J., and Balan, P.F. (2010). Attention as a decision in information space. *Trends Cogn. Sci.* 14, 240–248.
- Gottlieb, J.P., Kusunoki, M., and Goldberg, M.E. (1998). The representation of visual salience in monkey parietal cortex. *Nature* 391, 481–484.
- Gottlieb, J., and Snyder, L.H. (2010). Spatial and non-spatial functions of the parietal cortex. *Curr. Opin. Neurobiol.* 20, 731–740.
- Hayhoe, M., and Ballard, D. (2005). Eye movements in natural behavior. *Trends Cogn. Sci.* 9, 188–194.
- Hayhoe, M.M., McKinney, T., Chajka, K., and Pelz, J.B. (2012). Predictive eye movements in natural vision. *Exp. Brain Res.* 217, 125–136.
- Hickey, C., Chelazzi, L., and Theeuwes, J. (2010a). Reward changes salience in human vision via the anterior cingulate. *J. Neurosci.* 30, 11096–11103.
- Hickey, C., Chelazzi, L., and Theeuwes, J. (2010b). Reward guides vision when it's your thing: trait reward-seeking in reward-mediated visual priming. *PLoS ONE* 5, e14087.
- Hogarth L., Dickinson A., and Duka T., eds. (2010). *Selective Attention to Conditioned Stimuli in Human Discrimination Learning: Untangling the Effects of Outcome Prediction, Valence, Arousal, and Uncertainty* (Oxford: Oxford University Press).
- Holland, P.C., and Maddux, J.-M. (2010). Brain systems of attention in associative learning. In *Attention and Associative Learning*, C.J. Mitchell and M.E. Le Pelley, eds. (Oxford: Oxford University Press).
- Johnson, A., Varberg, Z., Benhardus, J., Maahs, A., and Schrater, P. (2012). The hippocampus and exploration: dynamically evolving behavior and neural representations. *Front. Hum. Neurosci.* 6, 216.
- Kable, J.W., and Glimcher, P.W. (2009). The neurobiology of decision: consensus and controversy. *Neuron* 63, 733–745.
- Karacan, H., and Hayhoe, M.M. (2008). Is attention drawn to changes in familiar scenes? *Vis. Cogn.* 16, 356–374.
- Kepecs, A., Uchida, N., Zariwala, H.A., and Mainen, Z.F. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature* 455, 227–231.
- Kolling, N., Behrens, T.E., Mars, R.B., and Rushworth, M.F. (2012). Neural mechanisms of foraging. *Science* 336, 95–98.
- Krajich, I., Armel, C., and Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nat. Neurosci.* 13, 1292–1298.
- Land, M.F. (2009). Vision, eye movements, and natural behavior. *Vis. Neurosci.* 26, 51–62.
- Le Pelley, M.E. (2010). *Attention and Associative Learning* (Oxford: Oxford University Press).
- Lo, C.C., and Wang, X.J. (2006). Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat. Neurosci.* 9, 956–963.
- Louie, K., Gratton, L.E., and Glimcher, P.W. (2011). Reward value-based gain control: divisive normalization in parietal cortex. *J. Neurosci.* 31, 10627–10639.
- Ma, W.J., Beck, J.M., and Pouget, A. (2008). Spiking networks for Bayesian inference and choice. *Curr. Opin. Neurobiol.* 18, 217–222.
- Maddux, J.M., and Holland, P.C. (2011). Dissociations between medial prefrontal cortical subregions in the modulation of learning and action. *Behav. Neurosci.* 125, 383–395.
- Maddux, J.M., Kerfoot, E.C., Chatterjee, S., and Holland, P.C. (2007). Dissociation of attention in learning and action: effects of lesions of the amygdala central nucleus, medial prefrontal cortex, and posterior parietal cortex. *Behav. Neurosci.* 121, 63–79.

- Maunsell, J.H., and Treue, S. (2006). Feature-based attention in visual cortex. *Trends Neurosci.* 29, 317–322.
- McCoy, A.N., and Platt, M.L. (2005). Risk-sensitive neurons in macaque posterior cingulate cortex. *Nat. Neurosci.* 8, 1220–1227.
- Mirabella, G., Bertini, G., Samengo, I., Kilavik, B.E., Frilli, D., Della Libera, C., and Chelazzi, L. (2007). Neurons in area V4 of the macaque translate attended visual features into behaviorally relevant categories. *Neuron* 54, 303–318.
- Mirpour, K., Arcizet, F., Ong, W.S., and Bisley, J.W. (2009). Been there, seen that: a neural mechanism for performing efficient visual search. *J. Neurophysiol.* 102, 3481–3491.
- Moore, T., and Armstrong, K.M. (2003). Selective gating of visual signals by microstimulation of frontal cortex. *Nature* 421, 370–373.
- Morris, G., Nevet, A., Arkadir, D., Vaadia, E., and Bergman, H. (2006). Midbrain dopamine neurons encode decisions for future action. *Nat. Neurosci.* 9, 1057–1063.
- Nassar, M.R., Rumsey, K.M., Wilson, R.C., Parikh, K., Heasly, B., and Gold, J.I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nat. Neurosci.* 15, 1040–1046.
- Niv, Y., and Schoenbaum, G. (2008). Dialogues on prediction errors. *Trends Cogn. Sci.* 12, 265–272.
- Noudoost, B., and Moore, T. (2011). Control of visual cortical signals by prefrontal dopamine. *Nature* 474, 372–375.
- O'Neill, M., and Schultz, W. (2010). Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. *Neuron* 68, 789–800.
- Oristaglio, J., Schneider, D.M., Balan, P.F., and Gottlieb, J. (2006). Integration of visuospatial and effector information during symbolically cued limb movements in monkey lateral intraparietal area. *J. Neurosci.* 26, 8310–8319.
- Oudeyer, P.Y., Kaplan, F., and Hafner, V.V. (2007). Intrinsic motivation systems for autonomous mental development. *IEEE Trans. Evol. Comput.* 11, 265–286.
- Payzan-LeNestour, E., and Bossaerts, P. (2011). Risk, unexpected uncertainty, and estimation uncertainty: Bayesian learning in unstable settings. *PLoS Comput. Biol.* 7, e1001048.
- Pearce, J.M., and Mackintosh, N.J. (2010). *Two Theories of Attention: A Review and a Possible Integration* (New York: Oxford University Press).
- Peck, C.J., Jangraw, D.C., Suzuki, M., Efem, R., and Gottlieb, J. (2009). Reward modulates attention independently of action value in posterior parietal cortex. *J. Neurosci.* 29, 11182–11191.
- Preusschoff, K., and Bossaerts, P. (2007). Adding prediction risk to the theory of reward learning. *Ann. N Y Acad. Sci.* 1104, 135–146.
- Preusschoff, K., Bossaerts, P., and Quartz, S.R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* 51, 381–390.
- Preusschoff, K., Quartz, S.R., and Bossaerts, P. (2008). Human insula activation reflects risk prediction errors as well as risk. *J. Neurosci.* 28, 2745–2752.
- Purcell, B.A., Schall, J.D., Logan, G.D., and Palmeri, T.J. (2012). From salience to saccades: multiple-alternative gated stochastic accumulator model of visual search. *J. Neurosci.* 32, 3433–3446.
- Redgrave, P., and Gurney, K. (2006). The short-latency dopamine signal: a role in discovering novel actions? *Nat. Rev. Neurosci.* 7, 967–975.
- Reynolds, J.H., and Heeger, D.J. (2009). The normalization model of attention. *Neuron* 61, 168–185.
- Rigotti, M., Ben Dayan Rubin, D., Wang, X.J., and Fusi, S. (2010). Internal representation of task rules by recurrent dynamics: the importance of the diversity of neural responses. *Front. Comput. Neurosci.* 4, 24.
- Rossi, A.F., Pessoa, L., Desimone, R., and Ungerleider, L.G. (2009). The prefrontal cortex and the executive control of attention. *Exp. Brain Res.* 192, 489–497.
- Rothkopf, C.A., and Ballard, D. (2010). Credit assignment in multiple goal embodied visuomotor behavior. *Front. Psychol.* 1, 173.
- Rothkopf, C.A., Ballard, D.H., and Hayhoe, M.M. (2007). Task and context determine where you look. *J. Vis.* 7, 11–20.
- Rushworth, M.F., Noonan, M.P., Boorman, E.D., Walton, M.E., and Behrens, T.E. (2011). Frontal cortex and reward-guided learning and decision-making. *Neuron* 70, 1054–1069.
- Saalmann, Y.B., and Kastner, S. (2011). Cognitive and perceptual functions of the visual thalamus. *Neuron* 71, 209–223.
- Schall, J.D., Purcell, B.A., Heitz, R.P., Logan, G.D., and Palmeri, T.J. (2011). Neural mechanisms of saccade target selection: gated accumulator model of the visual-motor cascade. *Eur. J. Neurosci.* 33, 1991–2002.
- Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol.* 57, 87–115.
- Schultz, W., Preusschoff, K., Camerer, C., Hsu, M., Fiorillo, C.D., Tobler, P.N., and Bossaerts, P. (2008). Explicit neural signals reflecting reward uncertainty. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 3801–3811.
- Snyder, L.H., Batista, A.P., and Andersen, R.A. (2000). Intention-related activity in the posterior parietal cortex: a review. *Vision Res.* 40, 1433–1441.
- So, N., and Stuphorn, V. (2012). Supplementary eye field encodes reward prediction error. *J. Neurosci.* 32, 2950–2963.
- Sprague, N., and Ballard, D.H. (2005). Modeling embodied visual behaviors. *ACM Trans. Appl. Percept.* 1, 1–26.
- Stanford, T.R., Shankar, S., Massoglia, D.P., Costello, M.G., and Salinas, E. (2010). Perceptual decision making in less than 30 milliseconds. *Nat. Neurosci.* 13, 379–385.
- Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787.
- Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2005). Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat. Rev. Neurosci.* 6, 363–375.
- Sutton, R.S., and Barto, A.G. (1998). *Reinforcement Learning: An Introduction* (Cambridge, MA: MIT Press).
- Takahashi, Y.K., Roesch, M.R., Wilson, R.C., Toreson, K., O'Donnell, P., Niv, Y., and Schoenbaum, G. (2011). Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nat. Neurosci.* 14, 1590–1597.
- Tatler, B.W., Hayhoe, M.M., Land, M.F., and Ballard, D.H. (2011). Eye guidance in natural vision: reinterpreting salience. *J. Vis.* 11, 5–25.
- Thompson, K.G., and Bichot, N.P. (2005). A visual salience map in the primate frontal eye field. *Prog. Brain Res.* 147, 251–262.
- Thompson, K.G., Biscoe, K.L., and Sato, T.R. (2005). Neuronal basis of covert spatial attention in the frontal eye field. *J. Neurosci.* 25, 9479–9487.
- Tobler, P.N., Christopoulos, G.I., O'Doherty, J.P., Dolan, R.J., and Schultz, W. (2009). Risk-dependent reward value signal in human prefrontal cortex. *Proc. Natl. Acad. Sci. USA* 106, 7185–7190.
- Vilares, I., and Kording, K. (2011). Bayesian models: the structure of the world, uncertainty, behavior, and the brain. *Ann. N Y Acad. Sci.* 1224, 22–39.
- Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends Cogn. Sci.* 9, 585–594.
- Waelti, P., Dickinson, A., and Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412, 43–48.
- Walton, M.E., Behrens, T.E., Noonan, M.P., and Rushworth, M.F. (2011). Giving credit where credit is due: orbitofrontal cortex and valuation in an uncertain world. *Ann. N Y Acad. Sci.* 1239, 14–24.
- Wardak, C., Olivier, E., and Duhamel, J.R. (2004). A deficit in covert attention after parietal cortex inactivation in the monkey. *Neuron* 42, 501–508.
- Wardak, C., Ibos, G., Duhamel, J.R., and Olivier, E. (2006). Contribution of the monkey frontal eye field to covert visual attention. *J. Neurosci.* 26, 4228–4235.

Wilson, R.C., and Niv, Y. (2011). Inferring relevance in a changing world. *Front. Hum. Neurosci.* 5, 189.

Wittmann, B.C., Daw, N.D., Seymour, B., and Dolan, R.J. (2008). Striatal activity underlies novelty-based choice in humans. *Neuron* 58, 967–973.

Yakushijin, R., and Jacobs, R.A. (2011). Are people successful at learning sequences of actions on a perceptual matching task? *Cogn. Sci.* 35, 939–962.

Yang, T., and Shadlen, M.N. (2007). Probabilistic reasoning by neurons. *Nature* 447, 1075–1080.

Yang, H., Chen, X., and Zelinsky, G.J. (2009). A new look at novelty effects: guiding search away from old distractors. *Atten. Percept. Psychophys.* 71, 554–564.

Yu, A.J., and Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron* 46, 681–692.