

Towards a neuroscience of active sampling and curiosity

Jacqueline Gottlieb^{1,2,3*} and Pierre-Yves Oudeyer^{4,5}

Abstract | In natural behaviour, animals actively interrogate their environments using endogenously generated ‘question-and-answer’ strategies. However, in laboratory settings participants typically engage with externally imposed stimuli and tasks, and the mechanisms of active sampling remain poorly understood. We review a nascent neuroscientific literature that examines active-sampling policies and their relation to attention and curiosity. We distinguish between information sampling, in which organisms reduce uncertainty relevant to a familiar task, and information search, in which they investigate in an open-ended fashion to discover new tasks. We review evidence that both sampling and search depend on individual preferences over cognitive states, including attitudes towards uncertainty, learning progress and types of information. We propose that, although these preferences are non-instrumental and can on occasion interfere with external goals, they are important heuristics that allow organisms to cope with the high complexity of both sampling and search, and generate curiosity-driven investigations in large, open environments in which rewards are sparse and ex ante unknown.

Since the earliest days of psychology and neuroscience, it has been recognized that the stream of evidence impinging on sensory receptors is ambiguous and incomplete, and animals must use active inference to make sense of the world. In vision, which is a dominant sensory modality in humans and non-human primates, the brain must use a retinal input that is 2D, constantly moving and ambiguous to infer the true state of a world that is stable, 3D and populated by meaningful entities. The relative insufficiency of the raw sensory input and the consequent need for active interpretation extend to all sensory modalities and all types of decision makers and behavioural situations. The efficiency with which biological nervous systems satisfy this goal is arguably a crowning achievement of evolution; its magnitude is made fully apparent by modern artificial intelligence applications such as drones or self-driving cars, in which it remains a considerable challenge to interpret rich, naturalistic sensory streams.

Among the most striking manifestations of active interpretation is the fact that, rather than building complete representations of all the information available to them, intelligent beings sparsely sample the rich, incoming sensory streams. Sparse sampling is a necessity for any limited-capacity organism that can sense much more information than it can fully process. Sampling is routinely manifested in attention and active-sensing behaviours, whereby animals inspect — that is, touch, listen, whisk or look at — selected sensory cues. In addition, it is expressed in intrinsically motivated behaviours such

as curiosity that reflect animals’ interest in specific topics or questions.

Despite the ubiquity and importance of sampling strategies, the organization and neural substrates of these strategies remain oddly unexplored. Studies of curiosity are relative newcomers to the neuroscience field^{1,2}. Similarly, although attention and active sensing have been investigated in voluminous literatures, these literatures focus on the ways in which attention and active sensing modulate other systems after they are deployed, rather than on the mechanisms that direct attention and generate sampling policies. Therefore, very little is known regarding the motives that drive attention and curiosity³. How do animals deem some sources of information to be more attention-worthy than others? How do they decide which stimuli or questions warrant investigation and which ones can be safely ignored?

Here, we review a nascent neuroscientific literature that examines these questions relying on novel active-sampling tasks inspired by earlier studies in cognitive psychology and the animal-learning literature (for examples, see REFS^{4–6}). We take an unusually integrative approach and focus on the commonalities between attention and curiosity and their relationship with decision-making, in particular in the learning and exploration–exploitation literature. Although attention and curiosity each encompass distinct and heterogeneous mechanisms and have been discussed in separate literatures, we propose that an integrative approach is appropriate at this stage because it highlights a core question

¹Department of Neuroscience, Columbia University, New York, NY, USA.

²Kavli Institute for Brain Science, Columbia University, New York, NY, USA.

³Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, USA.

⁴Inria, Bordeaux, France.

⁵Ensta ParisTech, Paris, France.

*e-mail: jg2141@columbia.edu

<https://doi.org/10.1038/s41583-018-0078-0>

that is relevant to both processes: which factor or factors motivate animals to engage with a stimulus or a question?

To organize the discussion, we introduce an important distinction between information sampling and information search. Information sampling involves gathering information relevant for a familiar task, such as looking at relevant stimuli while driving or asking for the answer to a trivia question. Information search, by contrast, refers to situations in which agents explore without prior knowledge of the task or goal. We emphasize the fact that both sampling and, especially, search entail high levels of complexity that are not fully recognized by current normative learning and decision theories⁷. Finally, we propose that animals cope with this complexity using systems of belief-based utility⁸, which confer value to information as a good in itself and motivate them to explore under conditions in which they must consider many alternatives and the relevant states and potential rewards are *ex ante* unknown.

Insofar as the questions we consider are related to fundamental ontological constraints of uncertainty and capacity limitations, our discussion draws on many strands of literature to which we cannot do full justice in this brief Review. For further enquiry, the interested reader is referred to excellent studies of active inference and information demand in cognitive psychology and computer science^{4,9–11}, studies of curiosity in personality and affective psychology^{12–15}, work on child development and education^{2,16–20}, experimental and theoretical considerations of information demand in economics^{21–23}, and discussions in philosophy²⁴ and the popular literature²⁵.

Active interrogation

In both neuroscience and psychology, the prevailing approach to studying brain and behaviour has been to provide participants with a preselected source of information — typically a sensory cue — and to require them to attend to, memorize or otherwise act on that cue. Studies of active sampling, by contrast, extend this approach by allowing participants to determine which source of information to consult before choosing an action. This apparently simple methodological change prompts a significant conceptual shift. Rather than being solely concerned with reactive processes by which agents process given sensory cues, this approach begs the question of how agents proactively determine which stimulus or question they wish to explore. As we will see in the following sections, this exposes multiple unexplored questions at the interface of traditional studies of cognition, decision-making and motivation.

Because the commodity that is sought by active-sampling behaviours is information, understanding the principles behind these behaviours requires a discussion of an individual's informational (or epistemic) states. Indeed, it is the extent of an agent's knowledge at the start of an investigation that defines the key distinction between information sampling and information search.

In information-sampling scenarios, animals gather information in the context of a familiar task that is oriented towards a known goal. These scenarios describe most instrumental tasks that are tested in the laboratory, in which participants seek to maximize an

external reward such as money or food and which emulate goal-directed natural behaviours, such as walking, driving or preparing tea^{26,27}. In all these behaviours, the decision maker knows the task structure, and this knowledge allows them to focus on a relatively small set of task-relevant actions and cues. For instance, a driver knows that they are likely to experience uncertainty when reaching an intersection and that specific stimuli (such as a traffic light) will help them resolve that uncertainty²⁸. As we discuss in the following sections, information sampling in these familiar instrumental settings is closely related to the exploration–exploitation literature and can be modelled as a strategy of reducing momentary uncertainty to maximize long-term operant gains²⁹.

In the case of information search, by contrast, agents investigate under conditions of much higher uncertainty, before knowing whether a useful pattern exists or what it may be. Consider a primitive human who notices that sparks fly when striking two stones. Although the human may notice and be surprised by the spark, they have few bases on which to decide whether and for how long to investigate this observation. Because the human knows next to nothing about the potential uses of fire and sparks, their decision cannot be motivated by reward maximization. Moreover, because they have very little knowledge of the possible explanation of what may give rise to the spark, they must consider a very large set of potentially relevant stimuli and hypotheses. Unlike the driver in the previous example, who can restrict their sampling to a small set of relevant cues, an agent motivated by curiosity must consider a much larger set of potentially relevant hypotheses and invest considerable time and effort into learning and discovery before knowing whether they can reap any benefits from their investigations.

Nevertheless, humans and other animals become curious about specific questions in what seems to be a targeted, non-random fashion, suggesting that they make well-defined choices even in conditions of ignorance and ambiguity. A central argument we make in this Review is that current normative theories (including learning and decision theories) fail to provide adequate descriptions of these choices because they do not take into account their computational complexity^{7,30}. We propose instead the alternative view that animals cope with complexity using systems of intrinsic motivations, including curiosity, by which they assign value to specific types of information gain or (changes in) cognitive states, independently of external rewards or environment structure.

We start by reviewing studies of instrumental information sampling and their relation to the attention and exploration–exploitation literatures. We continue by discussing non-instrumental sampling (or information seeking) tasks that operationalize intrinsically motivated sampling and open the door to studies of information search and curiosity.

Attention and decisions

Many studies of information sampling have been carried out in the domain of eye movements and visual attention, which are our key means of sampling visual information. Studies of eye movements and attention,

Agents

Any entities that are capable of learning and decision-making, including humans, other animals and artificial intelligence applications such as robots and self-driving cars.

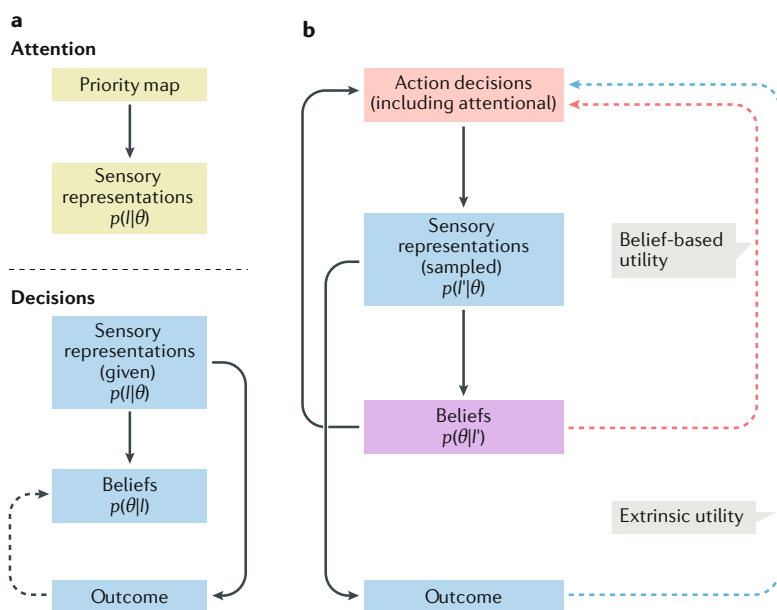


Fig. 1 | Proposed architectures of attention and decision-making in current research. **a** | Current views of attention and decision-making. Attention (top panel), according to the prevailing conception³¹, is envisioned as starting with a priority map (or an attention field) that specifies the to-be-attended feature or location and modulates the sensory representations in earlier visual areas. The genesis of the priority map is not specified: it has an output but no inputs. Decision-making (bottom panel), according to a common conception¹³⁰, starts with a given source of information, based on which the agent infers the state (model) of the world, selects the appropriate actions and experiences an outcome. Feedback from the outcome (dashed arrow) is used to learn the values of actions. The mathematical notation refers to the Bayesian description of these computations. Sensory responses provide a likelihood function, $p(l|\theta)$, which describes how likely is the current sensory response, l , given a hypothesis about the state of the world, θ . The decision maker's inference consists in inverting the likelihood to obtain a posterior belief, $p(\theta|l)$, that specifies how likely is a given state of the world, θ , given the sensory observation, l . **b** | Integration of attention and decision-making in information-sampling decisions. As in the standard view, the agent makes inferences, takes actions and experiences outcomes. However, in contrast to the standard view, attention is integrated into the decision chain as an action that selects sensory cues (or, more generally, selects information). Thus, the evidence that supports the decision is no longer given but is endogenously selected (indicated by l'). This selection, in turn, depends on the agent's beliefs about the relevant states (top black arrow on the left). For instance, a driver looks at a traffic light on the basis of their belief that the light will provide relevant information. Moreover, and particularly important, this allows the attentional selection (and not only the final actions) to be optimized for the current situation. This optimization may involve two types of utility-based feedback. One type of feedback describes the value of the eventual outcome — the material gains that the agent expects to obtain (extrinsic utility; dashed light blue arrow). Another type of feedback describes the value of cognitive states — the agent's preferences for what they wish to believe, know or observe (belief-based utility; dashed red arrow). Belief-based utility is particularly useful and indeed required in open-ended contexts in which the agent may lack advance knowledge of the material outcome (or the outcome is very delayed or uncertain). In these settings, intrinsic preferences for cognitive states allow agents to generate intrinsically motivated, curiosity-driven learning and discovery.

however, have remained largely separate from the decision literature. Let us start by reviewing this conceptual separation and the ways in which active sampling can help to bridge this gap.

On one side of this great divide, studies of selective attention examine the mechanisms by which the brain modulates the representation of sensory cues. Most extensively developed in the visual systems of humans

and non-human primates, this literature documents a range of attentional modulations that shape early and mid-level visual representations, and proposes that these modulations are driven by top-down input from higher-order control networks distributed in the frontal, parietal and temporal lobes³¹ (FIG. 1a). In monkeys, neurons in parts of this network, most notably the frontal eye fields (FEFs) and the lateral intraparietal (LIP) area, have visual responses that are spatially tuned and highly selective for task-relevant cues, and are thought to encode sparse priority maps in which only attention-worthy items are strongly represented and which can direct spatial attention or gaze^{32,33}. On the other side of the attention–decision divide, decision research explores how animals select between alternative actions^{34,35}. These studies have traced the encoding of sensory cues, and the subsequent mechanisms, which are distributed throughout the basal ganglia, the frontal cortex and the parietal lobes that read out the sensory representations, map them on relevant actions and mediate the learning of action–outcome associations^{34–36} (FIG. 1a).

However, although these studies provide our foundational knowledge of cognitive and brain function, they adopt a simplifying assumption that limits their explanatory power and enforces their conceptual separation. This is the assumption that animals act with near-complete knowledge of a decision situation.

In studies of decision-making, participants select actions from a set of predefined, clearly presented options. Consistent with this empirical practice, decision models that are applied to the data assume that the brain compares the representations of the pre-specified sources of 'signal' and 'noise'. These theories (most notably, signal detection theory and its dynamical variants, sequential sampling and drift diffusion models) allow the possibility that the decision maker has uncertainty about the specific values of decision-relevant states and some degree of control over how much information to sample to reduce that uncertainty and, in some variants, can incorporate attention-like modulations of decision thresholds or the rate of evidence accumulation (for examples, see REFS^{37,38}). However, these models universally assume that the agent has ex ante knowledge about the identity of the relevant states — that is, which portions of the environment constitute a signal and noise; they make no attempt to explain how the agent makes that determination.

Studies of selective attention allow the possibility that the brain differentially weights sensory cues, but they also start from the simplifying assumption that the decision maker knows to what to attend. In tasks of selective attention, humans are explicitly instructed, for instance, to "look for the T among the Ls", and monkeys are extensively trained to attend to specific features or locations³⁹. Likewise, neurocomputational models assume that the frontoparietal network has a priority map and can use it to orient gaze and attention, but they do not explain how the map is computed.

In their existing incarnation, therefore, studies of attention and decision-making adopt the simplistic assumption that decision makers have near-perfect knowledge of the relevant aspects of a decision

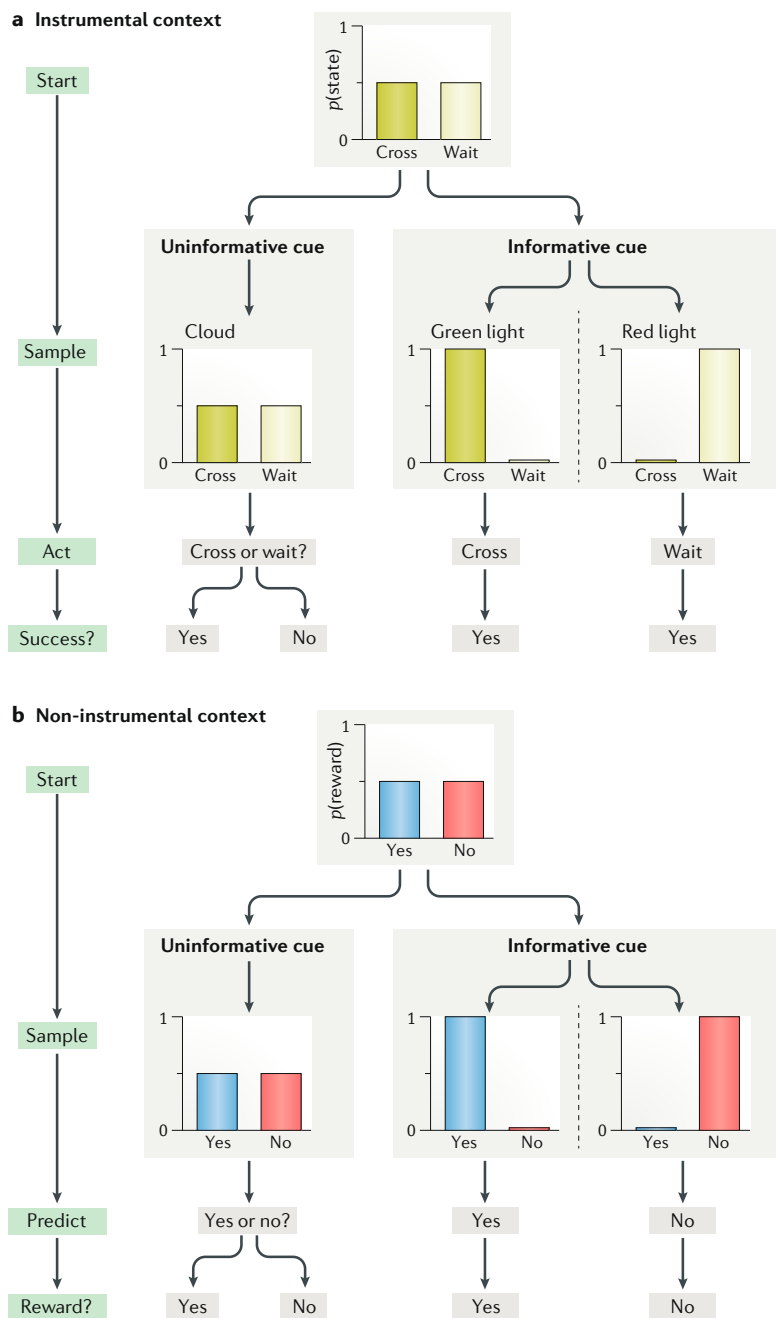


Fig. 2 | Evolution of beliefs when sampling information in different contexts. **a** | The panel provides an example of belief evolution in an instrumental context. A decision maker (for instance, a driver) arrives at an intersection ('Start') and decides whether to sample an informative cue (for example, a traffic light) or an uninformative stimulus (for example, a cloud). The driver may initially have uncertainty about the state of the world and estimate that crossing and waiting are equally likely to be the better actions. After viewing the traffic light, their beliefs will change and become strongly biased towards crossing or stopping, after which they are likely to take the appropriate action and receive a reward ('Success?'). After viewing an uninformative cue, uncertainty remains high, and the driver has a higher chance of taking the wrong action. **b** | The panel provides an example of belief evolution in a non-instrumental context. In this case, the states and belief changes are similar to those above, except that the agent can no longer take a decision based on the information. Therefore, both the informative and uninformative cues are associated with equivalent reward probabilities and are distinguished only by the agent's certainty regarding these probabilities.

situation^{3,27,40}. They are thus ill-equipped to capture realistic scenarios in which decision makers must consider multiple potentially relevant attributes and determine which attribute to attend when choosing an action.

Instrumental sampling

In contrast to the traditional approach described in the previous section, studies of active sampling directly address the nature of information-sampling policies. In these studies, participants are allowed to determine not only which action to take but also which one of several cues to consult before choosing that action. By examining participants' information demand (as expressed through an eye movement or the press of a button), this approach naturally bridges the divide between attention and decision research. Specifically, it reframes attention as one of several actions that an agent can take that has the role of reducing uncertainty and can be optimized to best serve a given situation (FIG. 1b).

In instrumental-sampling scenarios, agents are presumed to seek information that is relevant to a task and can decide which stimulus to sample on the basis of their familiarity with the task-relevant actions and cues. Consider again a driving scenario in which you reach an intersection and must decide what to do (FIG. 2). When arriving at the intersection, you may have uncertainty about which action to take — “should I step on the accelerator or the brake?” — and you may expect that the traffic light will reduce that uncertainty. This can be formally described as a distribution of beliefs about the relevant actions, which is initially uncertain (uniform) but is expected to become more heavily skewed towards one of the options if you look at the appropriate cue (FIG. 2). In a framework of probabilistic inference, changes in belief distributions can be measured as the difference between the dispersions (uncertainties) of the prior and posterior distributions, using metrics such as Shannon entropy, the Kullback–Leibler divergence or probability gains^{5,28}. These measures of expected changes in belief states can, in turn, serve as decision variables for sampling policies. An estimate of expected information gain would motivate the driver to look at the traffic light rather than a cloud, as the light is expected to reduce their uncertainty about the relevant actions to a greater extent than the cloud.

It is critical to note that, because instrumental information by definition pertains to a desirable goal, it has value only insofar as it furthers that goal. Technically, the value of instrumental information depends on the value of the outcome one seeks to achieve and the marginal increase in the probability of obtaining that outcome when acting with, relative to without, the information^{41,42}. In the preceding example, looking at the traffic light has no value per se; it has value only if it helps a driver safely cross an intersection.

Perhaps relying on this direct link between rewards and a reduction in decision uncertainty, many studies to date have focused on each factor individually. At one end of the spectrum, a recent study modelled attention on the basis of simple reward associative rules. The authors trained participants to choose between several options that differed in their reward probabilities

while simultaneously learning which stimulus feature was associated with the highest probability⁴³. They proposed that attention is allocated to features on the basis of their recent reward history and that this is mediated by changes in connectivity between networks of cognitive control and reward valuation (specifically, the dorsal frontoparietal network and the ventromedial prefrontal cortex). It must be noted, however, that the simple setting that the authors used for that task (where a visual feature had a one-to-one mapping with rewarded actions) cannot be applied directly to most natural conditions, in which cues bear arbitrary relations with future states and actions⁴⁴.

Several studies addressed these scenarios by modeling visual search as a mechanism for reducing uncertainty in a belief-updating framework similar to that described in FIG. 2 (REFS^{45–48}). In a series of functional MRI studies, it was proposed that the brain learns the reliability of alternative sensory cues (the potential of a cue to reduce uncertainty) by dynamically tracking visual prediction errors — that is, the extent to which the predictions made by a cue are confirmed or violated — and that this learning depends on the functional connectivity between several areas, including the temporal-parietal junction, the putamen, the FEF and the intraparietal sulcus^{49–51}. Finally, a neurophysiological study in monkeys showed that LIP neurons encoded the relative reliability of competing visual cues and could guide the monkeys' strategy for sampling the more informative cue⁵² (FIG. 3a,b). The evidence for reliability-based attention control is consistent with reports of reliability-based cue integration⁵³ and with studies of explicit information demand during categorization⁵⁴, suggesting that it is a widespread cognitive strategy.

Importantly, studies of instrumental information demand are closely related to the exploration–exploitation literature²⁹. This literature documents several strategies through which animals and humans trade off between gathering rewards from well-known, familiar options versus exploring less familiar options in order to potentially enhance the reward probability on longer time scales. These studies show that humans use targeted exploration, preferentially sampling options with higher uncertainty when this maximizes long-term operant gains^{55–57}. A related proposal is that humans rapidly adjust the weights (leverage) that they afford to different sensory features in ways that compensate for decision variability^{58–60}. Additional studies suggest that humans detect points in time in which the structure of the environment changes and upregulate arousal and learning rates specifically at these points⁵¹. Analogous findings in monkeys show that saccade-related activity in the FEF differs for exploratory and exploitative saccades⁶². Finally, LIP neurons have enhanced reward learning at the informative step in a two-step task in which only one of the decisions was consequential for the final reward, implying that the cells reflected the correct temporal credit assignment independently of the delay between a choice and its outcome⁶³. While the tasks used in these studies are substantially different from those used in attention research, the results are closely related. Both lines of research support the conclusion that humans and

monkeys can detect task junctures with high uncertainty and take actions to reduce that uncertainty, including initiating attentional sampling and upregulating arousal and learning rates.

Many open questions remain regarding the mechanisms of instrumental-sampling policies. One important set of questions involves the role of cognitive effort. As we noted above, sampling policies depend on hierarchical, top-down mechanisms that maintain a memory of the task-relevant cues. Acquiring information is associated with belief updating and longer fixation duration⁵², and changing the attentional set entails cognitive effort^{64,65}, potentially reducing performance in tasks that require high flexibility^{66,67}. Thus, a full understanding of sampling requires a better characterization of the neural mechanisms underlying cognitive effort and top-down attention control^{68–70}. A final consideration is that humans have imperfect metacognitive accuracy^{71,72}, and the control of behaviour based on uncertainty develops slowly with age⁵⁶, suggesting that behaviour based on decision uncertainty has important limitations.

In closing, let us return to the point we made at the outset — that, in instrumental scenarios, a reduction in decision uncertainty is closely related with an increase in reward gains. Because of this correlation, most studies did not attempt to determine whether the two quantities can be dissociated. Strikingly, however, a recent in monkeys shows that the brain honours this dissociation. The study showed that LIP neurons encode the reliability of visual cues independently of the rewards expected from acting based on these cues⁵². Specifically, these neurons differentiated between cues that provided decision-relevant information of different levels of reliability versus uninformative cues that had equivalent reward probability (FIG. 3c). This finding suggests that the brain encodes the expected reduction in uncertainty independently of expected gains in reward, an important idea that is further supported by experiments on non-instrumental sampling and curiosity.

Intrinsic motivation

Converging evidence shows that, in addition to gathering information to serve behavioural goals, animals are motivated to obtain information as a good in itself. Pure information preferences have been demonstrated in species as diverse as pigeons, humans and monkeys using so-called non-instrumental tasks, in which animals can observe predictive cues but cannot take actions based on these cues^{73–77}.

In a task of this kind used in a neurophysiological study, monkeys had, in each trial, a 50:50 chance of receiving a large or small reward and, before receiving the outcome, could choose to inspect one of two cues that provided early reward information⁷⁵. Similar to an instrumental context, the cues differed in their reliability, allowing the monkeys to form more-or-less accurate predictions regarding the outcome (FIG. 2b). However, critically different from instrumental contexts, the monkeys could not take actions based on the information — that is, they had no instrumental incentive to observe

Instrumental context

A context in which agents are motivated by the desire to obtain a known goal, which is operationalized in the laboratory as maximizing a material reward (such as money, points, food or safety).

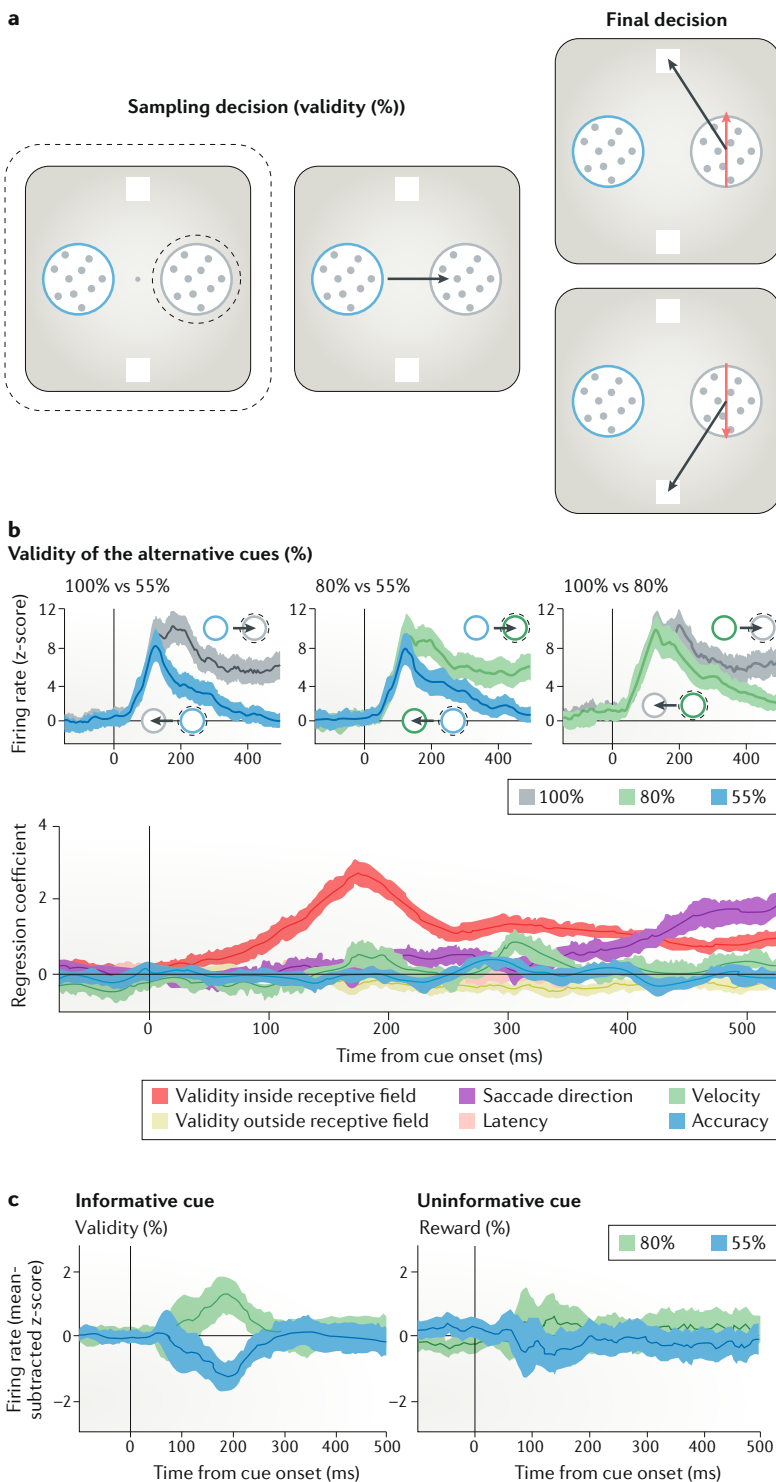


Fig. 3 | Neurons in lateral intraparietal area encode expected gains in information during instrumental sampling. a | In an instrumental-sampling task, monkeys viewed a display containing two cues (round apertures containing small dots) and two targets (small white squares). After a brief delay period, they made a first saccade to a cue ('Sampling decision') and a second saccade to a target ('Final decision'). The initial saccade gathered information relevant to the second saccade. The information consisted of 100% coherent dot motion directed towards one of the targets, which was only delivered after the sampling saccade (right panel, red arrows). During the epoch preceding the first saccade (left panel, dashed line frame), the cues were stationary) but were identified by a border whose colour indicated their reliability (percentage validity), defined as the probability that the motion signal that would be delivered by the cue correctly designated the rewarded target. **b** | The top row shows the firing rates (FRs) of a population of lateral intraparietal (LIP) cells during the sampling decision period, sorted according to saccade direction and the percentage validity of the cue. As shown by the inset cartoons, the monkeys almost always chose to sample the cue that had the higher percentage validity. The LIP neurons had higher FRs if the monkeys chose the cue inside rather than outside the receptive field (dashed circle), but this direction coding (the difference in FRs between the two saccade directions) scaled with the validity (compare the left, middle and centre panels). In the bottom row, the time-resolved regression analysis showed that the cells encoded percentage validity independently of confounding variables, including the percentage validity of the opposite cue and saccade metrics. **c** | Contrary to the predictions of a standard reinforcement learning account, LIP neurons encoded the differences in percentage validity between the informative cues but not equivalent differences in reward probability for uninformative cues. The traces show the mean-subtracted FR to highlight the cue-related modulations above and beyond the standard visual response. Figure adapted with permission from REF⁵², Proceedings of the National Academy of Sciences USA.

a particular cue (compare FIG. 2a and FIG. 2b). Despite this lack of incentive, monkeys showed robust preferences for viewing informative (rather than uninformative) cues. Moreover, these preferences were encoded by neurons implicated in reward and motivation in the orbitofrontal cortex⁷⁸ and midbrain dopamine cells⁷⁵, supporting the view that animals assign intrinsic value to engaging with reliable cues.

Understanding the logic of such pure information demand poses major conundrums for decision theories. In the absence of instrumental incentives, what motivates an individual to know or observe certain items? Although our forays into this question are in their infancy, two main hypotheses have been advanced in the theoretical literature. These views propose, alternatively, that individuals have an intrinsic desire for the early resolution of uncertainty²¹ or that they are simply motivated to engage with positive items^{23,79,80}.

These hypotheses predict substantially different sampling strategies. An individual who is intrinsically motivated to reduce uncertainty (independently of material gain) will attempt to maximize the accuracy of their beliefs, that is, to gather more precise information. By contrast, an individual who is intrinsically motivated to engage with positive cues will gather information in a biased fashion and may preferentially seek out pleasant but less accurate cues. Such a person may eagerly enquire about an upcoming vacation because this information makes them feel good, but they may avoid enquiring about a medical diagnosis because this produces dread and anxiety.

Emerging evidence suggests that reward-dependent attentional biases are pervasive in animals and humans and depend on dopaminergic mechanisms. One hypothesis, formalized in the context of reinforcement learning theory, is that dopamine neurons confer value to stimuli that allow animals to anticipate (savour) a positive outcome⁸¹. According to this view, stimuli that signal an increase in reward trigger a dopamine prediction error response, which, in addition to promoting learning, assigns value to engaging with the predictor itself. If the predictor is separate from the instrumental goal, this can interfere with the appropriate actions. For instance, some rats exhibit so-called sign tracking behaviours, in which they orient themselves towards a light that predicts a reward rather than to the magazine that delivers the reward⁸². Similarly, if monkeys view a reward-predictive cue at one visual location, they have difficulty making saccades to a separate location, and the neural correlates of these attentional biases are expressed in saccade-related activity in the superior colliculus and LIP area^{83–85}. These paradoxical effects of rewards reflect the fact that animals assign value to reward predictors above and beyond their valuation of the rewarded actions.

A second mechanism by which dopamine can produce attentional biases is by producing ‘reward-based salience’ — plasticity that confers visual salience independently of reward anticipation. This idea is supported by a large body of evidence showing that humans automatically orient to stimuli that have past reward associations even when the stimuli appear as irrelevant distractors (and presumably do not affect reward anticipation)^{86,87}. Moreover, monkeys proactively search for redundant reward cues (FIG. 4a), suggesting that reward-based salience motivates not only reactive orienting but also information demand in the absence of predictive utility⁸⁸. Finally, studies in humans show that reward-based distraction is associated with activation of the substantia nigra–ventral tegmental areas and that the ability to suppress such distraction depends on frontal and parietal areas associated with cognitive control^{87,89,90}.

In addition to inducing motivational conflict between task-irrelevant and task-relevant cues, reward-based attentional biases can impair the identification of informative items⁹¹. In a particularly elegant demonstration, researchers trained human participants to search for targets in natural scenes such that different targets were associated with monetary gains or the evasion of loss⁹². Although the targets had equal informativeness and operant value, object-selective cortical areas more faithfully encoded targets associated with gains rather than the evasion of loss, whereas the intraparietal sulcus had a less biased (strictly uncertainty-driven) target-related response. In a similar vein, the human N2pc response — a reliable electroencephalographic signature of spatial attention — was enhanced by reward probability independently of predictive value in a gambling task⁹³.

Together, these results underscore the multifaceted effects of rewards on attention and show that rewards can bias attention in a potentially maladaptive fashion by conferring intrinsic value to sensory cues. Such

attentional biases coexist with uncertainty reduction mechanisms, making it imperative to understand their neural substrates.

Curiosity

The fact that animals and humans seek information even when it serves no obvious purpose seems closely related to curiosity, and studies of information seeking have set the stage for recent investigations of the neural mechanisms of curiosity^{78,94}.

Curiosity, the intrinsic desire to know, has long been recognized as an important motive that influences human behaviour throughout the lifespan^{8,13,95}. Investigators as early as Berlyne proposed that curiosity is of several kinds, including perceptual curiosity (interest in specific stimuli), diversive curiosity (novelty or sensation seeking) and epistemic curiosity (interest in specific topics)¹³. As we discuss below, recent work in artificial intelligence suggests that the term ‘curiosity’ also applies to sensorimotor actions, as expressed, for example, by a child intent on learning how to manipulate a toy^{17,96}. All these behaviours can be computationally characterized as the autonomously generated motivation to answer a question in the absence of instrumental incentives. The question may be about the state of the environment (“what is out there in the universe?”) or how the environment can be manipulated (“how can I make the toy light up and play tunes?”).

The finding that individuals assign utility to non-instrumental information provides an obvious starting point for laboratory investigations of curiosity^{1,2}. Recent experiments extended this approach to study epistemic curiosity, using tasks in which participants rate their interest in a trivia question, such as “how many tons of steel are in the Eiffel tower?”^{97–100}. An additional study examined perceptual curiosity by exposing people to ambiguous (blurry) images, which were followed, in a fraction of trials, by visual disambiguation¹⁰¹. These investigations take us a step beyond the interest in material outcomes and raise the question of how people become interested in natural knowledge domains.

Extending the results from information-seeking tasks, these studies have linked curiosity with the widespread activation of systems of motivation, memory and attention. Higher curiosity ratings are associated with increased functional MRI activation in midbrain reward structures, including the caudate nucleus and the substantia nigra–ventral tegmental areas^{94,97,98,101}, supporting the idea that information is intrinsically motivating. In addition, questions that evoke higher curiosity are associated with better memory for the answer and enhanced connectivity between the ventral tegmental area and the hippocampus^{97,98}. Finally, perceptual curiosity engages frontal and parietal areas implicated in attention and cognitive control¹⁰¹. Participants with higher trait curiosity (assessed by personality questionnaires) engage in more widespread saccadic exploration of visual scenes¹⁰²; additionally, in trivia tasks, higher curiosity is associated with faster anticipatory shifts of gaze to the expected location of the answer (FIG. 4b), and curiosity levels can be read out by machine learning algorithms using only gaze patterns⁹⁹. These nascent

REVIEWS

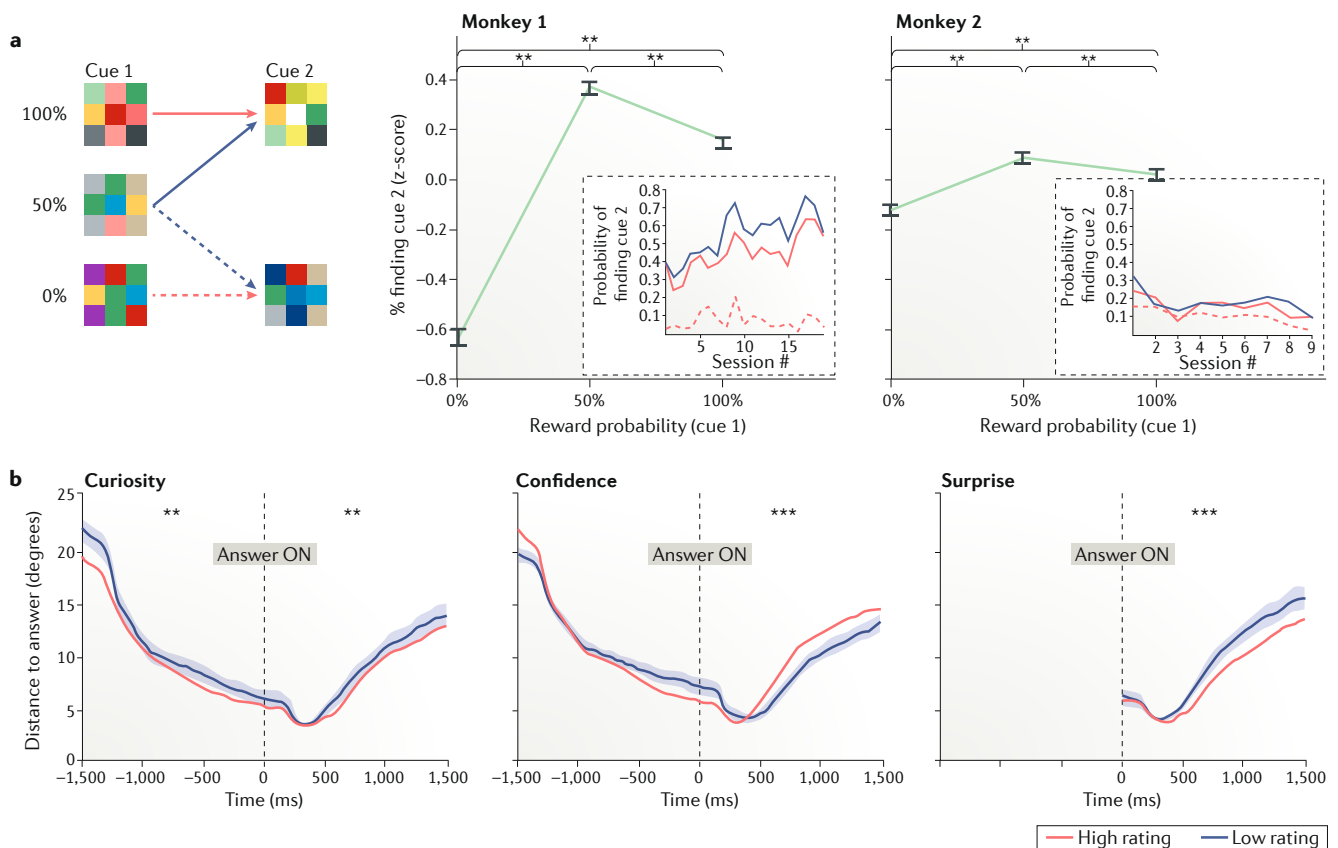


Fig. 4 | Eye movements are impacted by non-instrumental search and curiosity. **a** | In a visual information-seeking task, monkeys viewed an initial cue (cue 1) indicating the trial's reward probability (0%, 50% or 100%) and then had the opportunity to search for an additional cue (cue 2) that delivered perfect reward information (0% or 100% probability). The monkeys were motivated to search for cue 2 when they had reward uncertainty, that is, after viewing a 50% initial cue. However, they also searched vigorously in trials in which cue 2 was redundant, that is, it merely confirmed the previously signalled high reward probability (after a 100% but not a 0% initial cue). This preference was spontaneous (not trained) and remained consistent across sessions regardless of the monkeys' overall willingness to search (insets in each panel show the average of the raw data in chronological order per session). Stars indicate $P < 0.025$ (Wilcoxon test). **b** | This part provides evidence that human anticipatory gaze is influenced by curiosity. Participants received trivia questions and rated their curiosity about it and their confidence in knowing the answer. After a 2-second anticipatory period, the answer appeared in the upper-left corner of the monitor, and participants were asked to rate their surprise regarding the answer. The most reliable signature of curiosity was a faster deployment of gaze at the anticipated answer location (left panel). By contrast, higher surprise and lower confidence were associated with longer durations of dwelling on the answer after its presentation (right and centre panels, respectively). **, $P < 10^{-45}$; ***, $P < 10^{-75}$. Part **a** is adapted from REF.⁸⁸, Springer Nature Limited. Part **b** is adapted with permission from REF.⁹⁹, Elsevier.

studies of curiosity may shed light on complex aspects of human behaviour such as aesthetic appreciation^{12,13}. The longstanding idea that aesthetic pleasure is evoked by stimuli with intermediate levels of predictability, novelty or complexity^{103,104} suggests a link between aesthetics and informational demands, which is becoming amenable to investigation on the basis of recent studies that probe novelty–familiarity preferences^{105,106} and the role of sensory complexity¹⁰⁷. Moreover, the surprising finding that approximately 5% of humans do not experience pleasure from music despite having normal music perception and normal responses to monetary rewards raises the additional possibility that, in addition to activating the general purpose motivational and reward systems, aesthetic appreciation depends on domain-specific rewards¹⁰⁸.

Studies of epistemic curiosity, in turn, touch on the question of how humans develop lifelong interests and skills. The information gap theory suggests that curiosity arises when an individual encounters a question and generates a set of possible answers to it based on their previous knowledge, which in turn define a degree of uncertainty that a person may wish to resolve⁷⁹. Consistent with this proposed reliance on prior knowledge and memory, curiosity peaks if an individual has intermediate confidence that they know the answer to a question but declines if their confidence is too high or too low, indicating too little familiarity with the topic^{2,98,99}. The involvement of memory may be critically important for generating specific information search rather than a non-specific search for novelty (diversive curiosity). Particularly, it may allow individuals to

systematically build on knowledge domains they have some familiarity with and thereby develop lifelong interests and skills^{15,19}.

Curiosity as a tool for discovery

In the previous sections, we described evidence that humans and other animals assign intrinsic value to knowledge and information, but we have yet to discuss what such intrinsic utility may be useful for. What advantage do organisms derive from systems of intrinsic motivation and curiosity? In this final section, we propose that to answer this question, we must look beyond the ‘small-world’ information-sampling scenarios we have considered thus far and consider instead the full scope of curiosity-driven information search, in which decision makers cope with very high complexity on extended timescales, as is the case during scientific research or long-term learning (developmental or educational) trajectories.

In the computational modelling literature, a common approach^{17,109} to describing information search is to propose that humans and other animals are intrinsically motivated to learn the hidden structure of their environment (or, in technical terms, a ‘world model’), on the basis of the default assumption that this knowledge is useful for solving new problems that were not previously known or suspected^{110,111}. This perspective underlies several recent theories of curiosity, including normative (for example, the free-energy principle^{109,112}) and heuristic (for example, the learning progress hypothesis¹⁷) approaches. For instance, the free-energy principle proposes that learners select actions that optimize their beliefs — or equivalently, minimize surprise — over all possible states and hypotheses regarding the world^{109,112}.

However, as noted by other investigators^{7,30}, such a normative account does not take into account the complexity of the search process vis a vis the biological reality

of limited capacity. To quote Bossaerts and Murawski, in normative theories, “a decision problem with two alternatives is not distinguished from one with 2¹⁰⁰ alternatives”⁷. As such, these theories are incompatible with the abundant evidence that humans, in even moderately complex laboratory paradigms (for example, requiring advance planning over several steps), produce inconsistent solutions¹¹³ and adopt frugal heuristics that rely only on a very limited set of the available cues³⁰. In the domain of artificial intelligence, it is likewise well appreciated that normative optimization schemes such as active Bayesian inference scale poorly with problem complexity¹¹⁴; even non-parametric sampling-based models that reduce computational costs¹¹⁵ do not allow the systems to scale up to realistic problems of embodied control¹¹⁶.

In artificial intelligence, heuristic optimization mechanisms are the standard solution to computational complexity. Converging evidence shows that heuristics that assign intrinsic value to information gains can guide efficient information search and discovery (BOX 1). Robot-learning experiments that examined which heuristic mechanisms could scale to real-world, real-time learning in embodied agents fuelled the development of new theories of human epistemic curiosity, such as the learning progress hypothesis^{17,117}.

As a concrete example, consider a study in which a humanoid robot interacts with a set of objects with hidden interdependencies — for instance, a cylindrical object that is too far away to be moved by hand but can be moved with a stick, which can itself be moved only by appropriate movements of the robot’s arm and hand gripper — as well as with objects that cannot be controlled (although the robot does not initially know what is learnable)¹¹⁸ (FIG. 5a). In a classic reinforcement learning approach, the robot learns only if it receives an external reward (for example, if it can move the cylindrical object), which in this case would be exceedingly rare, producing learning that is prohibitively slow. By contrast, an algorithm^{116–118} based on intrinsic information rewards allows an agent to define its own goals, focusing on those where learning progress is possible in practice, and to remain motivated to learn to achieve these self-generated goals even when the material rewards are rare or unknown. This process leads the agent to discover a variety of effects that it can produce in the environment, acting as stepping stones for further discoveries, while avoiding spending too much time on goals that are either too easy or too difficult.

The learning progress models have been used to study these heuristics^{17,116,117}, using a low-level learning module that incrementally learns a predictive world model as new observations are collected and a meta-cognitive module that uses unsupervised learning to build representations of discrete tasks, estimate the prediction errors that a robot has in a task and provide intrinsic rewards if it detects an improvement in predicting or controlling the task. In the previous robotic example, this architecture motivates the robot to start by moving its hands, self-generating goals to reach with the hand, which initially provides maximal learning progress. In so doing, the robot serendipitously discovers how to move

Box 1 | Curiosity-driven learning in artificial intelligence and machine learning

There has been a growing interest in the field of artificial intelligence, machine learning and robotics for curiosity-driven learning algorithms (also called intrinsically motivated learning)^{96,117,126,127}. Beyond the use of neuroscience-inspired concepts to solve difficult engineering problems, these algorithms provide insights into understanding curiosity in the living.

Several algorithmic mechanisms have been systematically included in machines to implement efficient curiosity-driven learning. These mechanisms form an ecosystem of mechanisms with multi-way feedback loops^{117,128}. Learning mechanisms learn regularities, as well as their representations, from multimodal streams of perception and action. Meta-learning mechanisms learn to represent, track and predict the quality of a learning mechanism’s prediction or competence. An action generation system taps into the evolving memory abstractions of learning and meta-learning mechanisms to imagine potential actions to be tried^{117,127,128} or goals to be targeted¹¹⁶. Then, an action selection system chooses what to do by computing measures of intrinsic rewards derived from the expected impact on learning estimated by meta-learning mechanisms. As actions or goals are explored, new data are gathered, leading to updates of predictions and values and new representations of these predictions and values. In such a perspective, it becomes obvious that curiosity-driven behaviour cannot be explained solely by focusing on peculiar aspects of the architecture (for example, how intrinsic rewards are computed or maximized) but requires a systemic understanding in which, for example, latent representation learning may strongly influence (and be influenced by) the kind of situations, activities or self-generated goals that are targeted in curiosity-driven exploration^{116,128,129}.

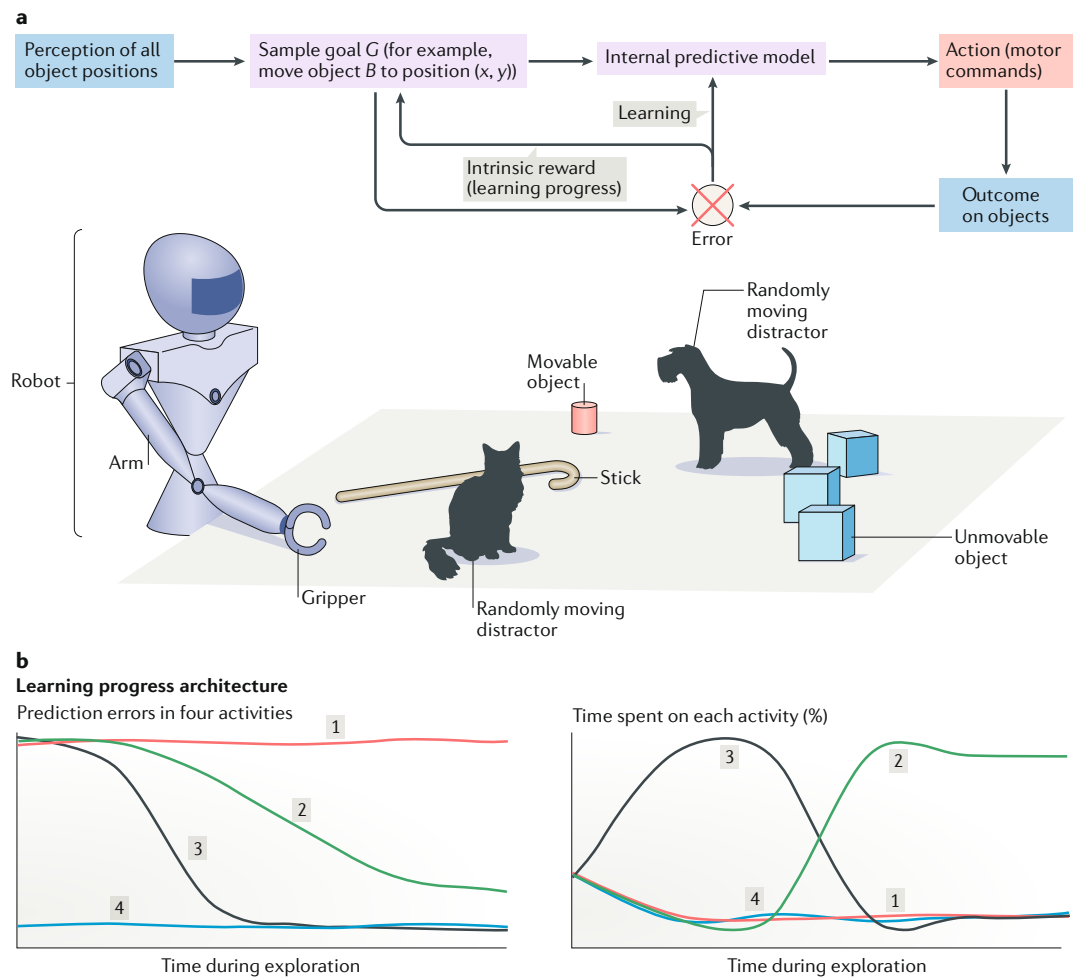


Fig. 5 | Artificial curiosity based on learning progress and autonomous goal sampling. **a** | The image shows an experimental set-up for testing a curiosity-driven exploration algorithm^{116,118} in (simulated) robots. A robot is placed in an environment with several objects, including distractors (some moving randomly and some unmovable). It receives rewards only if it manages to move the cylindrical object, which it can do by moving its arm to grasp and move the stick. Such a rare reward makes this problem very hard to solve with a reinforcement learning approach that optimizes for the reward. Ignoring the reward and exploring self-generated goals (for example, move object *B* to position (x, y)) that improve predictions and control results in a self-organized learning curriculum, in which the robot first learns to move the hand and then the stick and ultimately ends up discovering rather quickly how to move the cylindrical object. **b** | The charts show the results of a simulated experiment and reveal that a learning-progress-guided agent focuses on tasks that produce learning in practice (activities for which prediction errors decline) and avoids activities that show no change in prediction errors, as they are either already predictable (curve 4) or too difficult to learn (curve 1). Part **b** is adapted with permission from REF.¹⁷, Frontiers.

the stick, which it then identifies as a new niche of learning progress and focus. In turn, owing to the physical couplings in the environment, the robot discovers how to move the cylindrical object when moving the stick, creating a new niche of progress. Meanwhile, the robot learns to avoid interacting with distractor objects, from which little can be learned. Motivated by its own learning progress, the robot thus generates autonomous goals and progressively discovers new skills, and importantly, it does so without requiring advance knowledge of the environmental structure or the external rewards^{17,117,118}.

The efficiency of these intrinsically motivated architectures has provided support to the learning progress hypothesis¹⁷ and raises the question of whether

biological organisms might implement similar mechanisms. Although this remains an open question for future research, it is noteworthy that the predictions of learning progress models are consistent with psychological observations, in particular accounting for major phase transitions in infant development of vocalizations and tool use^{119,120}, and that software providing learning-progress-based personalized sequences of exercises improves childrens' learning in educational settings¹²¹.

A particularly important parallel is the fact that, consistent with behaviours reported in children and adults^{3,16,122}, a learning progress architecture generates a self-organized exploration curriculum that progresses from easier to more difficult tasks and favours tasks

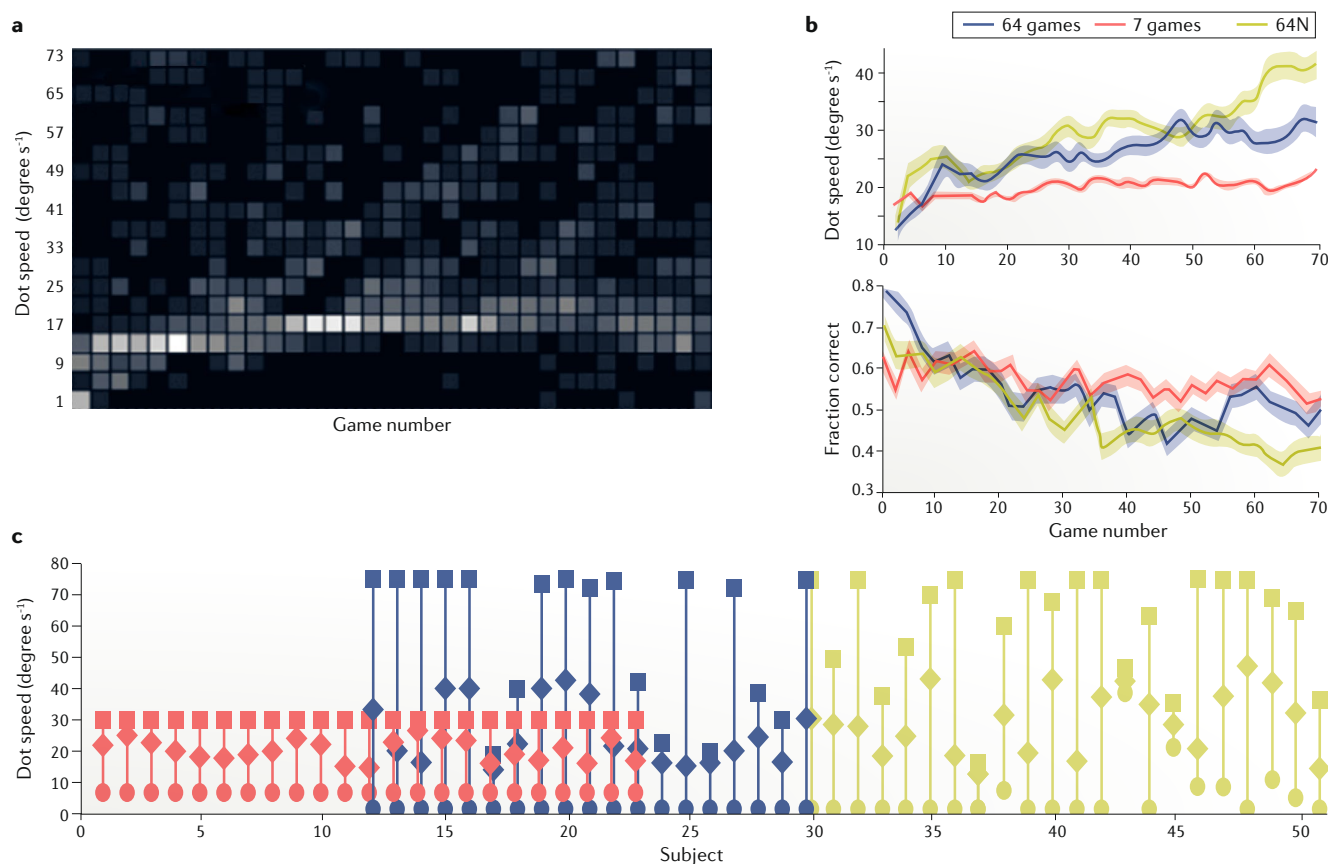


Fig. 6 | Self-organized play in a laboratory game. The figure shows the probability of selecting a game as a function of difficulty during a free-play session. **a** | Participants played a series of ~30-second games in which they pressed a key to intercept dots in an irregularly spaced stream as they crossed the centre of the screen. After playing an individual game, participants saw a choice display from which they freely chose the next game they played, with no additional constraints or instructions. The greyscale indicates the probability of selection of a given speed, across all participants as a function of time over the session. Despite the lack of explicit instructions, participants endogenously progressed from the easier to the more difficult games and focused their play on intermediate levels of difficulty. **b** | During a session, participants chose progressively more difficult games, indicated by increasing speeds (top panel) and lower success rates (bottom panel), regardless of whether the choice display had 7 or 64 games and whether a game provided the same (64) or a novel (64N) sequence each time it was selected. Numbers are averages and s.e.m. across players. **c** | This panel shows the range of speeds selected by each player. Each vertical bar shows the minimum (circle), maximum (square) and mean (diamond) dot speed selected by one player. Players are grouped according to the version they chose and are arranged in chronological order of testing. The vast majority of players sampled the most difficult game, whose speed was much faster than they could master. Figure adapted with permission from REF.¹²⁵, Frontiers.

of intermediate difficulty^{123,124} (FIG. 5b). Direct laboratory evidence supporting this view comes from a task in which adult humans freely interacted with a set of games of varying difficulty¹²⁵. Despite the fact that the participants received no specific instructions, they spontaneously organized their exploration in increasing order of difficulty and, after surveying the entire set of available tasks, settled on intermediate games in which their performance was 70–80% correct for the bulk of the session and gradually progressed to games of higher difficulty, consistent with the predictions of a learning progress mechanism¹²⁵ (FIG. 6). Continued development of strategic student tasks of this kind, in which participants freely explore a set of learning problems of varying difficulty, will help test the biological plausibility of the learning-progress-based process and clarify the behavioural and neural mechanisms by which

the brain autonomously organizes information search and generates useful discoveries.

Conclusions

We reviewed emerging neuroscientific evidence of the mechanisms generating information demand. We emphasized the fact that, although animals can often learn sophisticated goal-directed sampling policies, they also have intrinsic information preferences that are independent of instrumental demands, and they endogenously bias their attention and learning towards specific types of information and specific levels of challenge or accuracy.

We also proposed that these intrinsic drives are double-edged swords. On the one hand, in instrumental settings, intrinsic preferences may generate sampling biases and suboptimal learning strategies⁴⁴. On the other

hand, in novel or exploratory contexts, they may represent critically important heuristics for generating intermediate goals, organizing curiosity-driven investigations and making discoveries that would otherwise require implausibly complex optimization strategies.

The scant neuroscientific evidence available to date suggests that intrinsic information preferences depend on systems of reward and motivation and interact antagonistically or synergistically with systems of cognitive control that generate goal-directed sampling policies.

Although our understanding of these mechanisms is in its infancy, their continued investigation promises to shed light on important topics that fall at the intersection of traditional studies on cognition and decision-making, including aesthetic preferences, active learning and intrinsic motivation, preferences for self-challenge and the engagement of control, which have thus far remained poorly explored.

Published online: 05 November 2018

1. Gottlieb, J., Oudeyer, P. Y., Lopes, M. & Baranes, A. Information seeking, curiosity and attention: computational and empirical mechanisms. *Trends Cogn. Sci.* **17**, 585–593 (2013).
2. Kidd, C. & Hayden, B. Y. The psychology and neuroscience of curiosity. *Neuron* **88**, 449–460 (2015).
3. Gottlieb, J., Hayhoe, M., Hikosaka, O. & Rangel, A. Attention, reward and information seeking. *J. Neurosci.* **34**, 15497–154504 (2014).
4. Rehder, B. & Hoffman, A. B. Eye tracking and selective attention in category learning. *Cogn. Psychol.* **51**, 1–41 (2005).
5. Nelson, J. Finding useful questions: on Bayesian diagnosticity, probability, impact and information gain. *Psychol. Rev.* **112**, 979–999 (2005).
6. Coenen, A., Nelson, J. & Gureckis, T. Asking the right questions about the psychology of human inquiry: nine open challenges. *Psychon Bull. Rev.* <https://doi.org/10.3758/s13423-018-1470-5> (2018).
7. Bossaerts, P. & Murawski, C. Computational complexity and human decision-making. *Trends Cogn. Sci.* **21**, 917–929 (2017).
8. Loewenstein, G. & Molnar, A. The renaissance of belief-based utility in economics. *Nat. Hum. Behav.* **2**, 166–167 (2018).
9. Chater, N. & Loewenstein, G. The under-appreciated drive for sense-making. *J. Econ. Behav. Organiz.* **126**, 137–154 (2016).
10. Wu, C. M., Meder, B., Filimon, F. & Nelson, J. D. Asking better questions: how presentation formats influence information search. *J. Exp. Psychol. Learn. Mem. Cogn.* **43**, 1274–1297 (2017).
11. Markant, D. B. & Gureckis, T. M. Is it better to select or to receive? Learning via active and passive hypothesis testing. *J. Exp. Psychol. Gen.* **143**, 94–122 (2014).
12. Berlyne, D. *Conflict, Arousal and Curiosity* (McGraw-Hill, 1960).
13. Berlyne, D. E. A theory of human curiosity. *Br. J. Psychol.* **45**, 180–191 (1954).
14. Litman, J. A. in *Issues in the Psychology of Motivation* (ed. Zelik, P. R.) (Nova Science Publishers, 2007).
15. Silvia, P. J. *Exploring the Psychology of Interest* (Oxford Univ. Press, 2006).
16. Di Domenico, S. I. & Ryan, R. M. The emerging neuroscience of intrinsic motivation: a new frontier in self-determination research. *Front. Hum. Neurosci.* <https://doi.org/10.3389/fnhum.2017.00145> (2017).
17. Kaplan, F. & Oudeyer, P.-Y. In search of the neural circuits of intrinsic motivation. *Frontiers Neurosci.* **1**, 225–225 (2007).
This is a clear and succinct review of the concepts and computational models of intrinsic motivation and their importance to artificial intelligence.
18. Gopnik, A. Scientific thinking in young children: theoretical advances, empirical research, and policy implications. *Science* **337**, 1623–1627 (2012).
19. Renninger, K. A. & Hidi, S. E. *The Power of Interest for Motivation and Engagement* (Routledge, NY, 2016).
20. Begus, K., Gliga, T. & Southgate, V. Infants' preferences for native speakers are associated with an expectation of information. *Proc. Natl Acad. Sci. USA* **113**, 12397–12402 (2016).
21. Kreps, D. M. & Porteus, E. L. Temporal resolution of uncertainty and dynamic choice theory. *Econometrica* **46**, 185–200 (1978).
22. Caplin, A. & Dean, M. Revealed preference, rational inattention and costly information acquisition. *Am. Econ. Rev.* **105**, 2183–2203 (2015).
23. Caplin, A. & Leahy, J. Psychological expected utility theory and anticipatory feelings. *Q. J. Econ.* **116**, 55–79 (2001).
24. Clark, A. *Surfing Uncertainty: Prediction, Action and the Embodied Mind*. (Oxford Univ. Press, 2015).
25. Livio, M. *Why? What Makes Us Curious?*. (Simon and Schuster, 2017).
26. Hayhoe, M. & Ballard, D. Modeling task control of eye movements. *Curr. Biol.* **24**, 622–628 (2014).
This paper provides an excellent overview of empirical and modelling studies of eye movement control in natural tasks.
27. Tatler, B. W., Hayhoe, M. N., Land, M. F. & Ballard, D. H. Eye guidance in natural vision: reinterpreting salience. *J. Vis.* **11**, 5–25 (2011).
28. Bach, D. R. & Dolan, R. J. Knowing how much you don't know: a neural organization of uncertainty estimates. *Nat. Rev. Neurosci.* **13**, 572–586 (2012).
29. Cohen, J. D., McClure, S. M. & Yu, A. J. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Phil. Trans. R. Soc. B* **362**, 933–942 (2007).
30. Todd, P. M. & Gigerenzer, G. Précis of simple heuristics that make us smart. *Behav. Brain Sci.* **23**, 727–780 (2000).
31. Reynolds, J. H. & Heeger, D. J. The normalization model of attention. *Neuron* **61**, 168–185 (2009).
32. Thompson, K. G. & Bichot, N. P. A visual salience map in the primate frontal eye field. *Prog. Brain Res.* **147**, 251–262 (2005).
33. Bisley, J. W. & Goldberg, M. E. Attention, intention, and priority in the parietal lobe. *Annu. Rev. Neurosci.* **33**, 1–21 (2010).
34. Hanks, T. D. & Summerfield, C. Perceptual decision making in rodents, monkeys, and humans. *Neuron* **93**, 15–31 (2017).
35. Kable, J. W. & Glimcher, P. W. The neurobiology of decision: consensus and controversy. *Neuron* **63**, 733–745 (2009).
36. Lee, D., Seo, H. & Jung, M. W. Neural basis of reinforcement learning and decision making. *Annu. Rev. Neurosci.* **35**, 287–308 (2012).
37. Krajbich, I., Armel, C. & Rangel, A. Visual fixations and the computation and comparison of value in simple choice. *Nat. Neurosci.* **13**, 1292–1298 (2010).
38. Krajbich, I., Lu, D., Camerer, C. & Rangel, A. The attentional drift-diffusion model extends to simple purchasing decisions. *Front. Psychol.* **3**, 193 (2012).
39. Gottlieb, J. Attention, learning, and the value of information. *Neuron* **76**, 281–295 (2012).
40. Gottlieb, J. Understanding active sampling strategies: empirical approaches and implications for attention and decision research. *Cortex* **102**, 150–160 (2018).
This is an overview of empirical approaches to information sampling in neurophysiology.
41. Johnson, L., Sullivan, B., Hayhoe, M. & Ballard, D. H. Predicting human visuomotor behavior in a driving task. *Phil. Trans. R. Soc. B* **369**, 20130044 (2014).
42. Sullivan, B. T., Johnson, L., Rothkopf, C. A., Ballard, D. & Hayhoe, M. The role of uncertainty and reward on eye movements in a virtual driving task. *J. Vis.* **12**, 19 (2012).
43. Leong, Y., Radulescu, A., Daniel, R., DeWoskin, V. & Niv, Y. Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron* **93**, 451–463 (2017).
44. Wilson, R. C. & Niv, Y. Inferring relevance in a changing world. *Front. Hum. Neurosci.* **5**, 189 (2011).
45. Najemnik, J. & Geisler, W. S. Eye movement statistics in humans are consistent with an optimal search strategy. *J. Vis.* **8**, 4 (2008).
46. Yang, S. C., Lengyel, M. & Wolpert, D. M. Active sensing in the categorization of visual patterns. *eLife* **5**, e12215 (2016).
This paper provides evidence for information-based eye movement strategies using behavioural analysis and Bayesian modelling in humans.
47. Najemnik, J. & Geisler, W. S. Optimal eye movement strategies in visual search. *Nature* **434**, 387–391 (2005).
48. Renninger, L. W., Verghese, P. & Coughlan, J. Where to look next? Eye movements reduce local uncertainty. *J. Vis.* **7**, 6 (2007).
49. Vossel, S., Vossel, S., Mathys, C., Stephan, K. E. & Friston, K. J. Cortical coupling reflects bayesian belief updating in the deployment of spatial attention. *J. Neurosci.* **35**, 11532–11542 (2015).
This is an analysis of attention in a Bayesian framework using functional MRI in humans.
50. Vossel, S. et al. Spatial attention, precision, and bayesian inference: a study of saccadic response speed. *Cereb. Cortex* **24**, 1436–1450 (2014).
51. Vossel, S., Thiel, C. M. & Fink, G. R. Cue validity modulates the neural correlates of covert endogenous orienting of attention in parietal and frontal cortex. *NeuroImage* **32**, 1257–1264 (2006).
52. Foley, N. C., Kelley, S. P., Mhatre, H., Lopes, M. & Gottlieb, J. Parietal neurons encode expected gains in instrumental information. *Proc. Natl Acad. Sci.* **114**, E3315–E3323 (2017).
This paper demonstrates that oculomotor neurons encode expected information gains in monkeys.
53. Ernst, M. O. & Banks, M. S. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**, 429–433 (2002).
54. Nelson, J., McKenzie, C., Cottrell, G. & Sejnowski, T. Experience matters: information acquisition optimizes probability gain. *Psychol. Sci.* **21**, 960–969 (2010).
55. Zajkowski, W. K., Kossut, M. & Wilson, R. C. A causal role for right frontopolar cortex in directed, but not random, exploration. *eLife* **6**, e27430 (2017).
56. Somerville, L. H. et al. Charting the expansion of strategic exploratory behavior during adolescence. *J. Exp. Psychol. Gen.* **146**, 155–164 (2017).
57. Manohar, S. G. & Husain, M. Attention as foraging for information and value. *Front. Hum. Neurosci.* **7**, 111 (2013).
58. Krishnamurthy, K., Nassar, M. R., Sarode, S. & Gold, J. I. Arousal-related adjustments of perceptual biases optimize perception in dynamic environments. *Nat. Hum. Behav.* **1**, 0107 (2017).
59. Li, V., Herce Castañón, S., Solomon, J. A., Vandormael, H. & Summerfield, C. Robust averaging protects decisions from noise in neural computations. *PLOS Comput. Biol.* **13**, e1005723 (2017).
60. Spitzer, B., Waschke, L. & Summerfield, C. Selective overweighing of larger magnitudes during noisy numerical comparison. *Nat. Hum. Behav.* **1**, 0145 (2017).
61. Gold, J. I. & Stocker, A. A. Visual decision-making in an uncertain and dynamic world. *Annu. Rev. Vis. Sci.* **3**, 227–250 (2017).
62. Ebitz, R. B., Albarran, E. & Moore, T. Exploration disrupts choice-predictive signals and alters dynamics in prefrontal cortex. *Neuron* **97**, 450–461 (2018).
63. Gersch, T. M., Foley, N. C., Eisenberg, I. & Gottlieb, J. Neural correlates of temporal credit assignment in the parietal lobe. *PLOS ONE* **9**, e88725 (2014).
64. Rossi, A. F., Pessoa, L., Desimone, R. & Ungerleider, L. G. The prefrontal cortex and the executive control of attention. *Exp. Brain Res.* **192**, 489–497 (2009).
65. Rossi, A. F., Bichot, N. P., Desimone, R. & Ungerleider, L. G. Top down attentional deficits in macaques with lesions of lateral prefrontal cortex. *J. Neurosci.* **27**, 11306–11314 (2007).
66. Morvan, C. & Maloney, L. Human visual search does not maximize the post-saccadic probability of identifying targets. *PLOS Comput. Biol.* **8**, e1002342 (2012).
This presents an intriguing demonstration that humans show suboptimal sampling strategies in a task requiring flexible adjustments based on estimates of visibility.

67. Ghahghaei, S. & Verghese, P. Efficient saccade planning requires time and clear choices. *Vision Res.* **113B**, 125–136 (2015).
68. Chong, T. T. et al. Neurocomputational mechanisms underlying subjective valuation of effort costs. *PLoS Biol.* **15**, e1002598 (2017).
69. Shenhav, A. et al. Toward a rational and mechanistic account of mental effort. *Annu. Rev. Neurosci.* **40**, 99–124 (2017).
70. Fan, J. An information theory account of cognitive control. *Front. Hum. Neurosci.* <https://doi.org/10.3389/fnhum.2014.00680> (2014). **This paper proposes a reframing of theories of cognitive control from the perspective of informational constraints.**
71. Fleming, S. & Daw, N. Self-evaluation of decision-making: a general Bayesian framework for metacognitive computation. *Psychol. Rev.* **124**, 91–114 (2017).
72. Zhang, H., Daw, N. D. & Maloney, L. T. Human representation of visuo-motor uncertainty as mixtures of orthogonal basis distributions. *Nat. Neurosci.* **18**, 1152–1158 (2015).
73. Vasconcelos, M., Monteiro, T. & Kacelnik, A. Irrational choice and the value of information. *Sci. Rep.* **5**, 13874 (2015).
74. Eliaz, K. & Schotter, A. Experimental testing of intrinsic preferences for noninstrumental information. *Am. Econ. Rev.* **97**, 166–169 (2007).
75. Bromberg-Martin, E. S. & Hikosaka, O. Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron* **63**, 119–126 (2009).
76. Bennett, D., Bode, S., Brydevall, M., Warren, H. & Murawski, C. Intrinsic valuation of information in decision making under uncertainty. *PLoS Comp. Biol.* **12**, e1005020 (2016).
77. Brydevall, M., Bennett, D., Murawski, C. & Bode, S. The neural encoding of information prediction errors during non-instrumental information seeking. *Sci. Rep.* **8**, 6134 (2018).
78. Blanchard, T. C., Hayden, B. Y. & Bromberg-Martin, E. S. Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. *Neuron* **85**, 602–614 (2015). **This paper demonstrates single-neuron encoding of non-instrumental information value in the monkey orbitofrontal cortex.**
79. Golman, R. & Loewenstein, G. Information gaps: a theory of preferences regarding the presence and absence of information. *Decision* **5**, 143–164 (2018).
80. Loewenstein, G. Anticipation and the valuation of delayed consumption. *Econ. J.* **97**, 666–684 (1987).
81. Igaya, K., Story, G. W., Kurth-Nelson, Z., Dolan, R. J. & Dayan, P. The modulation of savouring by prediction error and its effects on choice. *eLife* **5**, e13747 (2016). **This paper presents a reinforcement learning model of non-instrumental information demand, proposing that, in addition to producing learning, dopaminergic reward prediction errors confer value to predictor states.**
82. Flagel, S. B. & Robinson, T. E. Neurobiological basis of individual variation in stimulus-reward learning. *Curr. Opin. Behav. Sci.* **13**, 178–185 (2017).
83. Peck, C. J., Jangraw, D. C., Suzuki, M., Efem, R. & Gottlieb, J. Reward modulates attention independently of action value in posterior parietal cortex. *J. Neurosci.* **29**, 11182–11191 (2009).
84. Foley, N. C., Jangraw, D. C., Peck, C. & Gottlieb, J. Novelty enhances visual salience independently of reward in the parietal lobe. *J. Neurosci.* **34**, 7947–7957 (2014).
85. Isoda, M. & Hikosaka, O. A neural correlate of motivational conflict in the superior colliculus of the macaque. *J. Neurophysiol.* **100**, 1332–1342 (2008).
86. Anderson, B. The attention habit: how reward learning shapes attentional selection. *Ann. NY Acad. Sci.* **1369**, 24–39 (2016). **This is a comprehensive review of reward-related attention biases and their neural mechanisms and behavioural importance in humans.**
87. Hickey, C., Chelazzi, L. & Theeuwes, J. Reward guides vision when it's your thing: trait reward-seeking in reward-mediated visual priming. *PLoS ONE* **5**, e14087 (2010).
88. Daddaoua, N., Lopes, M. & Gottlieb, J. Intrinsically motivated oculomotor exploration guided by uncertainty reduction and conditioned reinforcement in non-human primates. *Sci. Rep.* **6**, 20202 (2016).
89. Hickey, C., Chelazzi, L. & Theeuwes, J. Reward changes salience in human vision via the anterior cingulate. *J. Neurosci.* **30**, 11096–11103 (2010).
90. Hickey, C. & Peelen, M. V. Neural mechanisms of incentive salience in naturalistic human vision. *Neuron* **85**, 512–518 (2015).
91. Hunt, L. T., Rutledge, R. B., Malalasekera, W. M., Kennerley, S. W. & Dolan, R. J. Approach-induced biases in human information sampling. *PLoS Biol.* **14**, e2000638 (2016).
92. Barbaro, L., Peelen, M. V. & Hickey, C. Valence, not utility, underlies reward-driven prioritization in human vision. *J. Neurosci.* **37**, 10438–10450 (2017). **This is among the first empirical demonstrations of reward-based and uncertainty-based modulations of visual representations in the human high-level cortex.**
93. San Martín, R., Appelbaum, L. G., Huettel, S. A. & Woldorff, M. G. Cortical brain activity reflecting attentional biasing toward reward-predicting cues covaries with economic decision-making performance. *Cereb. Cortex* **26**, 1–11 (2016).
94. van Lieshout, L. L. F., Vandenbroucke, A. R. E., Müller, N. C. J., Cools, R. & de Lange, F. P. Induction and relief of curiosity elicit parietal and frontal activity. *J. Neurosci.* **38**, 2579–2588 (2018). **This is a demonstration of non-instrumental information value and its neural correlates in humans.**
95. Loewenstein, G. The psychology of curiosity: a review and reinterpretation. *Psychol. Bull.* **116**, 75–98 (1994).
96. Baldassare, G., Mirolli, M. (eds) *Intrinsically Motivated Learning in Natural and Artificial Systems* (Springer-Verlag, Berlin, 2013).
97. Gruber, M. J., Gelman, B. D. & Ranganath, C. States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron* **84**, 486–496 (2014). **This paper demonstrates the effects of curiosity on memory and the hippocampus in humans.**
98. Kang, M. J. et al. The wick in the candle of learning: epistemic curiosity activates reward circuitry and enhances memory. *Psychol. Sci.* **20**, 963–973 (2009).
99. Baranes, A. F., Oudeyer, P. Y. & Gottlieb, J. Eye movements encode epistemic curiosity in human observers. *Vis. Res.* **117**, 81–90 (2015).
100. Marvin, C. B. & Shohamy, D. Curiosity and reward: valence predicts choice and information prediction errors enhance learning. *J. Exp. Psychol. Gen.* **145**, 266–272 (2016).
101. Jepma, M., Verdonck, R. G., van Steenbergen, H., Rombouts, S. A. & Nieuwenhuis, S. Neural mechanisms underlying the induction and relief of perceptual curiosity. *Front. Behav. Neurosci.* <https://doi.org/10.3389/fnbeh.2012.00005> (2012). **This is a study of perceptual curiosity using functional MRI in humans.**
102. Risko, E. F., Anderson, N. C., Lanthier, S. & Kingstone, A. Curious eyes: individual differences in personality predict eye movement behavior in scene-viewing. *Cognition* **122**, 86–90 (2012).
103. Salimpoor, V. N., Zald, D. H., Zatorre, R. J., Dagher, A. & McIntosh, A. R. Predictions and the brain: how musical sounds become rewarding. *Trends Cogn. Sci.* **19**, 86–91 (2015).
104. Huron, D. *Sweet Anticipation: Music and the Psychology of Expectation* (MIT Press, 2006).
105. Liao, H. I., Yeh, S. L. & Shimojo, S. Novelty versus familiarity principles in preference decisions: task-context of past experience matters. *Front. Psychol.* **2**, 43 (2011).
106. Park, J., Shimojo, E. & Shimojo, S. Roles of familiarity and novelty in visual preference judgments are segregated across object categories. *Proc. Natl Acad. Sci. USA* **107**, 14552–14555 (2010).
107. Güçlütürk, Y., Güçlü, U., van Gerven, M. & van Lier, R. Representations of naturalistic stimulus complexity in early and associative visual and auditory cortices. *Sci. Rep.* **8**, 3439 (2018).
108. Zatorre, R. J. Musical pleasure and reward: mechanisms and dysfunction. *Ann. NY Acad. Sci.* **1337**, 202–211 (2015).
109. Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P. & Pezzulo, G. Active inference: a process theory. *Neural Comput.* **29**, 1–49 (2016).
110. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction* (MIT Press, 1998).
111. Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P. & Dolan, R. J. Model-based influences on human choices and striatal prediction errors. *Neuron* **69**, 1204–1215 (2011).
112. Friston, K. J. et al. Active inference, curiosity and insight. *Neural Comput.* **29**, 2633–2683 (2017).
113. Morewedge, C. K. & Kahneman, D. Associative processes in intuitive judgment. *Trends Cogn. Sci.* **14**, 435–440 (2010).
114. Buckley, C., Kim, C. S., McGregor, S. & Seth, A. K. The free energy principle for action and perception: a mathematical review. *J. Math. Psychol.* **81**, 55–79 (2017).
115. Gershman, S. J. & Blei, D. M. A tutorial on Bayesian nonparametric models. *J. Math. Psychol.* **56**, 1–12 (2012).
116. Baranes, A. & Oudeyer, P. Y. Active learning of inverse models with intrinsically motivated goal exploration in robots. *Rob. Auton. Syst.* **61**, 49–73 (2013).
117. Oudeyer, P. Y., Kaplan, F. & Hafner, V. V. Intrinsic motivation systems for autonomous mental development. *IEEE Trans. Evol. Comput.* **11**, 265–286 (2007).
118. Forestier, S. & Oudeyer, P. Y. in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)* 3965–3972 (IEEE, 2016).
119. Moulin-Frier, C., Nguyen, S. M. & Oudeyer, P. Y. Self-organization of early vocal development in infants and machines: the role of intrinsic motivation. *Front. Psychol.* **4**, 1006 (2014).
120. Forestier, S. & Oudeyer, P. Y. in *Proc. 39th Annual Meeting of the Cognitive Science Soc.* 2013–2018 (Cogsci, 2017).
121. Clement, B., Roy, D., Oudeyer, P. Y. & Lopes, M. Multi-armed bandits for intelligent tutoring systems. *J. Educ. Data Mining* **7**, 2 (2015).
122. Metcalfe, J. Metacognitive judgments and control of study. *Curr. Dir. Psychol. Sci.* **18**, 159–163 (2009).
123. Lopes, M. & Oudeyer, P. Y. in *Proc. IEEE Int. Conf. on Development and Learning and Epigenetic Robotics (ICDL)* 1–8 (IEEE, 2012).
124. Son, L. & Sethi, R. Metacognitive control and optimal learning. *Cogn. Sci.* **30**, 759–774 (2006).
125. Baranes, A. F., Oudeyer, P. Y. & Gottlieb, J. The effects of task difficulty, novelty and the size of the search space on intrinsically motivated exploration. *Front. Neurosci.* **8**, 317 (2014). **This presents a novel laboratory task for examining intrinsically motivated exploration based on difficulty in humans.**
126. Barto, A., Singh, S. & Chenatez, N. in *Proc. 3rd Int. Conf. Dvp. Learn* 112–119 (San Diego, CA, 2004).
127. Schmidhuber, J. in *Proc. Int. Joint Conf. Neural Networks* **2**, 1458–1463 (IEEE, 1991).
128. Bellemare, M. et al. in *Proc. Advances in Neural Information Processing Systems 29 Conf.* 1471–1479 (NIPS, 2016).
129. Kulkarni, T. D., Narasimhan, K., Saedi, A. & Tenenbaum, J. B. in *Proc. Advances in Neural Information Processing Systems 29 Conf.* 3675–3683 (NIPS, 2016).
130. Pouget, A., Drugowitsch, J. & Kepecs, A. Confidence and certainty: distinct probabilistic quantities for different goals. *Nat. Neurosci.* **19**, 366–374 (2016).

Acknowledgements

The authors acknowledge support from the Human Frontiers Science Program (Collaborative Research Grant RGP0018/2016 to J.G. and P.-Y.O.), an Inria Neurocuriosity grant (to J.G. and P.-Y.O.), the National Eye Institute (RO1 grant to J.G.) and the National Institute of Mental Health (RO1 grant to J.G.).

Author contributions

The authors both researched data for the article, provided substantial contributions to discussion of content, wrote the article, and reviewed and edited the manuscript before submission.

Competing interests

The authors declare no competing interests.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Reviewer information

Nature Reviews Neuroscience thanks V. Stuphorn and the other anonymous reviewers for their contribution to the peer review of this work.