

Review

Curiosity-driven exploration: foundations in neuroscience and computational modeling

Alireza Modirshanechi,¹ Kacper Kondrakiewicz,^{2,3,4} Wulfram Gerstner,^{1,*} and Sebastian Haesler ^{2,3,4,5,*}

Curiosity refers to the intrinsic desire of humans and animals to explore the unknown, even when there is no apparent reason to do so. Thus far, no single, widely accepted definition or framework for curiosity has emerged, but there is growing consensus that curious behavior is not goal-directed but related to seeking or reacting to information. In this review, we take a phenomenological approach and group behavioral and neurophysiological studies which meet these criteria into three categories according to the type of information seeking observed. We then review recent computational models of curiosity from the field of machine learning and discuss how they enable integrating different types of information seeking into one theoretical framework. Combinations of behavioral and neurophysiological studies along with computational modeling will be instrumental in demystifying the notion of curiosity.

An urge to explore the unknown

The scientific study of curiosity emerged in the late 19th and early 20th centuries with the advent of experimental psychology which recognized curiosity early on as a major factor influencing human and animal behavior [1,2]. Experimental observations – such as rats foregoing food in favor of spontaneously exploring a novel environment – were interpreted as evidence for a motivational drive which needs to be satisfied by acquiring information [3].

A prominent conceptual framework developed by Daniel Berlyne further discriminates between ‘perceptual’ and ‘epistemic’ curiosity [4]. Perceptual curiosity describes the phenomenon observed in all animals, including humans, that novel sensory stimuli elicit **arousal** (see [Glossary](#)), sensory inspection, and exploration. Epistemic curiosity, by contrast, refers to curiosity that creates ‘manifestations of knowledge’, considered to be exclusive to humans. However, the boundaries between perceptual and epistemic curiosity appear rather blurry, since long-lasting stimulus memories caused by perceptual curiosity may be considered knowledge, and knowledge acquisition is also observed in animals other than humans [5]. Parallel conceptual frameworks introduced the notion of violated expectations as the cause of curiosity [6,7]. Building on this idea, the information gap framework describes the main cause of curiosity as the difference between ‘what one knows and what one wants to know’ [8]. Collectively, past conceptual frameworks have been instrumental in highlighting different aspects of curiosity, yet no single, widely accepted definition or framework has emerged. In this review, we therefore take an approach that is agnostic about specific conceptual frameworks.

There is a growing consensus that curious behavior is (i) related to seeking or reacting to information, and (ii) not goal-directed (i.e., exploratory actions are not aimed toward a goal or completion of a task). Hence, we focus on behavioral and neurophysiological studies which satisfy these two conditions and group them into three categories according to the type of information-seeking observed ([Figure 1](#)). The first category includes curious behavior indicated by **reactive orienting**.

Highlights

Exploratory behaviors are central to the study of curiosity, since they allow us to infer an otherwise hidden state of being curious from quantitative, experimental observations.

Reactive orienting behaviors reflect a fundamental form of curiosity and are driven by the immediate novelty and/or surprise of stimuli.

Operant tasks probe curiosity by offering experimental subjects explicit action choices to collect information, in a manner that is either neutral or costly/useful with regards to overall gains/losses.

Recently developed computational models of curiosity, based on intrinsically motivated reinforcement learning, describe different expressions of curiosity in a unified theoretical framework.

¹École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

²Neuroelectronics Research Flanders (NERF), Leuven, Belgium

³VIB, Leuven, Belgium

⁴Department of Neuroscience, KU Leuven, Leuven, Belgium

⁵Leuven Brain Institute, Leuven, Belgium

*Correspondence: wulfram.gerstner@epfl.ch (W. Gerstner) and sebastian.haesler@nerf.be (S. Haesler).



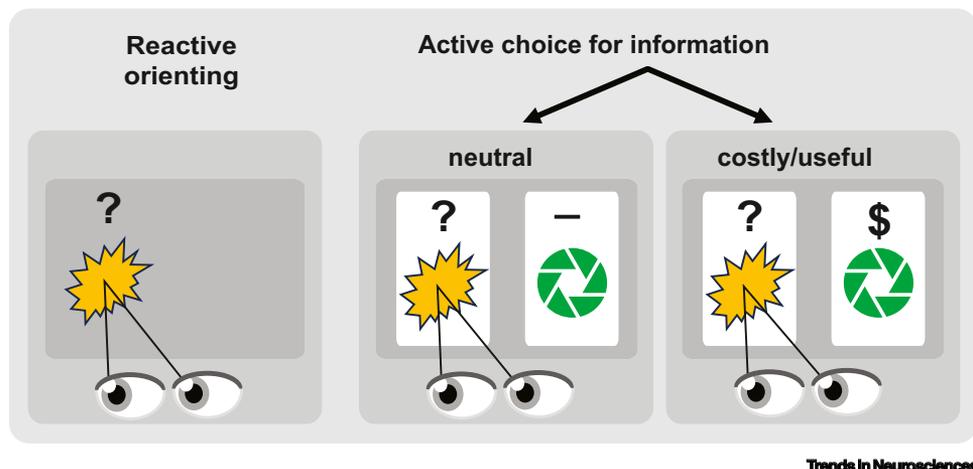


Figure 1. Categorization of behavioral paradigms to study curiosity. Left: reactive orienting occurs spontaneously in response to curiosity-evoking stimuli. Examples of exploratory orienting behaviors include gaze shifts towards novel or informative stimuli in humans and non-human primates. Right: curiosity revealed through active choices for information. In an example task, an agent prefers fixating on an image that has been associated with the possibility of receiving new information over a second image with no such associations. The second image can be reward-neutral ('-') leading to a paradigm of reward-neutral information gathering; or it can potentially give access to a positive reward ('\$') or loss leading to a paradigm where curiosity is costly/useful in terms of overall gain.

In the second category, we include specific tasks involving an active choice for information, where, similar to reactive orienting, information seeking does not confer any extrinsic gains or losses (neutral information). In the third category, we include tasks with active choices for costly or useful information, that is, tasks where exploratory actions affect the amount of reward or punishment received. We finally review recent computational models emerging from the field of artificial intelligence and discuss how they help link different experimental studies in psychology and neuroscience.

Reactive orienting

The most fundamental form of curiosity may be found among orienting behaviors. Across animal species, novel or surprising stimuli spontaneously elicit arousal and evoke sensory inspection and exploration with very short latency [9–13] (Figure 2). The ubiquity of orienting behaviors in the animal kingdom suggests that behavioral reactions are hardwired by evolution to induce exploration independently of other goals. Orienting responses show **habituation** after a few exposures, suggesting a rapid form of non-associative learning. When encountering multiple stimuli, animals and humans have the innate tendency to investigate novel or surprising stimuli longer than familiar, expected ones [14]. This preference for novel stimuli, also referred to as **novelty-seeking**, has been widely used as a measure of curiosity [15]. However, avoidance of novel stimuli has also been well described, indicating considerable variability among individuals and experimental paradigms with regard to novelty preference [16].

Reactive orienting requires neural pathways that process sensory stimuli to detect novelty or **surprise** and initiate orienting movements. Moreover, orienting has been linked to the activation of **neuromodulatory systems**. In the following, we review findings related to these three aspects of the neurophysiology of reactive orienting.

Evaluating surprise and novelty signals

Both novelty and surprise broadly impact neural activity in a wide variety of neural circuits [17, 18]. We distinguish two different ways in which this can happen (Figure 3). First, novelty and surprise

Glossary

Active choice for information: active exploratory actions to collect anticipated information or intrinsic rewards.

Depending on the experimental design, the association between curiosity-driven actions and the tasks' extrinsic rewards can be averted (neutral curiosity) or be positive/negative (useful/costly curiosity).

Arousal: physiological and neuronal state of heightened activity or alertness. Arousal is evoked by surprise and novelty, but it is not restricted to that; for example, arousal can be evoked by fully expected large rewards.

Attention: allocation of neural processing resources to relevant stimuli. The extent to which a stimulus is processed depends on the amount of attentional capacity available, which is among other things determined by levels of arousal. Surprising and novel stimuli attract processing resources.

Empowerment: quantifies the potential an agent perceives about how much it can control the environment.

Extrinsic rewards: in this review we use the term 'extrinsic rewards' to refer to gains or penalties in experimental setups.

Habituation: the decrement of behavioral or neuronal responses due to repeated stimulation and not caused by motor fatigue or nonspecific sensory adaptation.

Information gain: refers to the impact a stimulus has on an agent's expectation about upcoming stimuli. It is a type of progress rate and is sometimes considered as a form of surprise.

Intrinsic rewards: refers to appealing aspects of a stimulus that are different from the stimulus association with the task's extrinsic rewards. Examples of intrinsic rewards are novelty, surprise, progress rate, information gain, and empowerment. Curiosity-driven actions are often interpreted as actions towards collecting intrinsic rewards.

Multistep tasks with sparse reward: tasks where extrinsic rewards (and punishments) are delivered only after several actions of the agents, as opposed to simple tasks with immediate frequent extrinsic rewards (e.g., bandit tasks). These tasks are particularly important to study useful/costly curiosity.

Neuromodulatory systems: clusters of neurons which have widespread projection pathways and regulate cellular excitability and synaptic plasticity

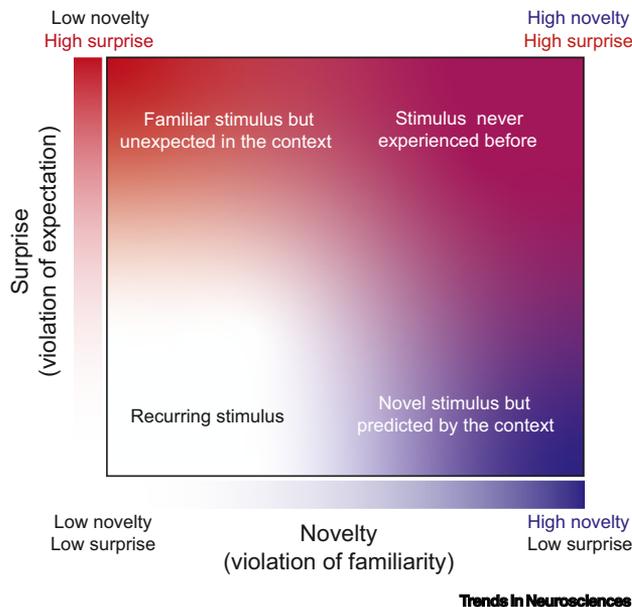


Figure 2. Surprise and novelty evoke curiosity. The novelty of a stimulus depends on whether it has been encountered before, how frequently it has been encountered, and how similar it is to previously encountered stimuli (familiarity dimension) [17,18,22,49,78]. The surprise of a stimulus, however, depends on how predictable the stimulus is given contextual cues (e.g., a conditioned stimulus) and the extent of its influence on an agent's expectation of upcoming stimuli (expectation dimension) [122,123]. The two dimensions are highly correlated in typical experiments (e.g., oddball tasks [124]), leading to maximal orienting responses. However, novelty and surprise can be dissociated by specific experiment designs (see [17] for a review). For example, consider two sequences of binary stimuli: (i) the repeating sequence of 'AAABAAAB...' and (ii) a

purely random sequence with occurrence probability of 75% for 'A' and 25% for 'B'. Stimulus 'A' is more familiar/frequent than 'B' in both cases. However, while 'A' is more expected than 'B' in the random sequence, the two stimuli are equally expected in the repeating sequence (because of the deterministic regularity). Several formal definitions have been suggested for how expectation and familiarity can be quantified [122,125], and the neural signatures of many of these definitions have been found in the brain [17]. While novelty and surprise are dissociable concepts, we do not exclude the possibility that some neuronal populations are activated by both. We further note that progress rate and information gain, as two other main proposed drives of curiosity [87,92,102], can also be categorized as measures of surprise [122].

can modulate sensory responses in structures along sensory pathways of different modalities, including the thalamus, subcortical nuclei, and sensory cortices [19,20]. As with behavioral orienting reactions, neuronal responses habituate with recurring experience. The modulation of sensory responses can occur already very early during sensory processing, as has been demonstrated for mitral cells in the mouse olfactory bulb, only one synapse away from the chemical receptors [21]. In a recent study using large-scale electrical recordings in macaque monkeys, all of 22 investigated brain regions contained at least a few per cent of neurons which showed larger visual responses when stimuli were either novel and surprising or just surprising [22]. Frequently, the same neurons were modulated in both conditions. Consistent with the magnitude of behavioral responses, the combination of novelty and surprise evoked the largest response modulation (Figure 2).

Aside from having a modulatory influence on sensory responses, novelty and surprise can also be explicitly encoded in a manner which is invariant to stimulus identity or the sensory modality (Figure 3). This has been observed less widely (outside neuromodulatory systems) (see next subsection). Selective responses to novelty have been identified in primates late in the ventral visual stream, in the anterior ventral medial temporal cortex, in the perirhinal cortex [23,24], as well as in the hippocampus and amygdala [25,26]. While lesions of the neocortex typically disrupt orienting behaviors [27], lesions of perirhinal [28,29] and hippocampal [30] areas leave orienting intact but disrupt recognition memory [31–33]. This suggests that the medial temporal lobe structures evaluate novelty signals in the context of memory formation rather than for mediating orienting behaviors.

in their projection targets by modulating rather than directly exciting or inhibiting neural activity. Major neuromodulators include noradrenaline, serotonin, dopamine, and acetylcholine.

Novelty: subjective stimulus quality related to the frequency with which a stimulus has been experienced in the past. Forgetting increases the novelty of a stimulus, whereas similarity to other frequently encountered stimuli lowers it. The term 'familiarity' refers to the opposite situations.

Progress rate: the influence that a stimulus has on the predictions and policy adjustments of an agent upon observation.

Reactive orienting: spontaneous reaction to curiosity-evoking stimuli which includes arousal, sensory inspection, and orienting of sensory organs or the whole body to the stimulus. Reactive orienting can be modeled by a fixed action policy to seek immediately available intrinsic rewards.

Surprise: measures how unexpected a stimulus is based on contextual cues such as conditioned stimuli or specific environmental circumstances. The term 'expectedness' refers to the opposite effect.

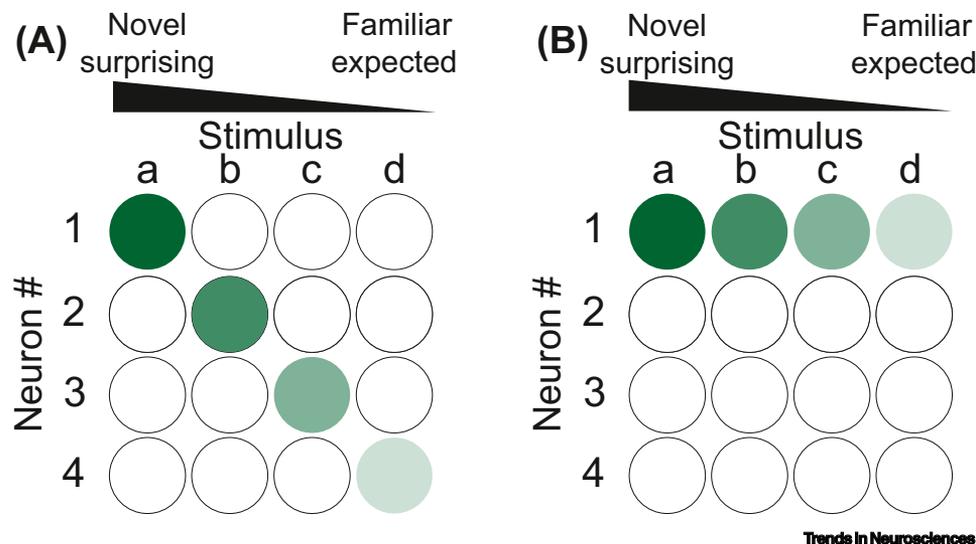


Figure 3. Hypothetical neuronal encoding schemes related to novelty and surprise. Schematic representation of four neurons (1–4) encoding four different stimuli (a–d), which have different degrees of novelty or surprise. While stimulus ‘a’ is highly novel and/or surprising, stimulus ‘d’ is expected and/or familiar. (A) Neurons encode sensory stimulus identity but are also modulated by familiarity and/or expectation. This has been observed, for example, in visual cortical areas [126]. While in the illustrated simplified case, each neuron responds to only one stimulus, in reality neurons typically respond to multiple stimuli in a combinatorial manner. (B) Neuron 1 responds to novelty and/or surprise but is invariant to the sensory stimulus identity. Example neurons reside in the locus coeruleus (LC) [41].

Initiating orienting movements

The successful detection of a novel and/or surprising stimulus evokes a cascade of orienting reactions. These commonly include reactions associated with arousal – such as transient changes in skin conductance, respiration, pupil diameter, and heart rate [34] – but also exploratory movements of sensory organs such as the eyes or the whiskers, and even the whole body [9]. Neural effector circuits for initiating orienting and exploratory movements have been localized to the superior colliculus in the midbrain [35], and some of its major inputs including the subthalamic nucleus zona incerta [14] and the substantia nigra pars reticulata [36,37], as well as various brainstem nuclei [38]. Importantly, stimuli of different sensory modalities evoke broadly similar orienting responses: for example, mice start sniffing in response to not only novel odorants but also novel sounds or lights [39,40]. This suggests that the premotor brain areas involved in orienting behaviors are multimodal, that is, they are shared between different sensory modalities.

Activation of neuromodulatory systems

A third aspect of the neurophysiology of reactive orienting involves catecholaminergic neuromodulation, which can be functionally linked to both evaluating surprise/novelty and initiating movements. Catecholaminergic neurons exhibit selective responses to novel stimuli with latencies of <150 ms [41,42], which is consistent with their involvement in early processing of curiosity-evoking stimuli. Specifically, neurons in the pontine locus coeruleus (LC) – which is the source of noradrenaline in the neocortex and hippocampus and has been implicated in the control of arousal [43] – respond to novel stimuli with bursts of action potentials, which habituate rapidly [44]. LC neurons also co-release dopamine in novel environments, leading to memory enhancement [45]. Pharmacological evidence further suggests that noradrenaline actively promotes orienting-related exploratory behaviors [46]. Activation of the LC has also been linked to the initiation of the distinct cortical electroencephalographic potential elicited by novel stimuli, called

'novelty P3' [34]. The novelty P3 habituates with subsequent presentations of novel items and is markedly reduced for expected compared with unexpected novel stimuli, suggesting an enhancement by surprise [18,47,48].

Novel environments and novel stimuli have further been shown to stimulate dopamine neuron activity in rodents, non-human primates, and humans [42,49–53]. Dopamine responses to novelty facilitate associative learning of novel stimuli [49], underlie novelty-induced exploration [54], and may be related more generally to movement initiation [55]. However, dopamine release specifically in the tail of striatum during assessment of a novel object has been related to avoidance of novel objects, which was interpreted to be based on an individual animal's estimation of threat [56].

Interestingly, there is little direct evidence that surprise alone activates dopamine neurons. Pharmacological manipulations of dopamine in rats have been shown to influence surprise processing in the auditory midbrain [57]. However, in an olfactory oddball experiment in mice, rare stimuli, which did cause a behavioral orienting reaction, failed to activate dopamine neurons [49]. Further studies are needed to characterize how dopamine neurons respond to surprise and novelty.

Active choice for neutral information

In the second category of information-seeking behaviors, we review studies using behavioral tasks, also referred to as non-instrumental information-gathering tasks [15], in which experimental subjects actively choose to collect anticipated information: for example, monkeys actively take actions that are followed by novel visual stimuli [58]. Contrary to innate reactive orienting, such information-gathering behaviors require action policies which are based on the learned associations between possible actions and expected future information. To keep subjects engaged, **extrinsic rewards** (often food in the context of animal studies or monetary compensation in humans) are commonly used, but any association between curiosity-driven exploratory actions and the tasks' extrinsic rewards are averted by either delivering extrinsic rewards randomly or as one-offs, independently of experimental subjects' action choices [59–61]. Even though final extrinsic rewards are thus the same for all action-selection strategies, and seeking information is neutral with respect to extrinsic rewards, animals (including humans) consistently choose to collect information about unknown stimuli [58], unfamiliar facts [62], or upcoming extrinsic rewards [61,63].

Active choices for neutral information are often accompanied by reactive orienting reactions: for example, human eye movements are modulated by interest in symbolic, semantic information [64]. This highlights the close interplay between reactive orienting and active information seeking, also present at the level of neural circuits. Many brain areas involved in orienting (e.g., the medial temporal cortex, amygdala, or zona incerta) are similarly activated when selecting task-related actions to seek novel information [22].

Even though, by design, extrinsic rewards play no role in tasks involving active choices for neutral information, evidence suggests an important role for the extrinsic reward system. Dopamine neurons respond to information-predicting cues [63] and neurons in the lateral habenula, a major input to dopamine neurons, are activated by sensory cues that are repeatedly presented before novel stimuli [65]. This anticipatory activity has also been referred to as 'information prediction error' [65], akin to the concept of reward prediction errors [66]. Consistent with this, neuroimaging revealed an activation of the substantia nigra pars compacta/ventral tegmental area (SNc/VTA) in anticipation of novel stimuli [67]. Novelty further enhances reward responses in the ventral

striatum, a major projection target of dopamine neurons [68], which has also been implicated in the choice of novel over familiar stimuli [69]. Finally, activation of the SNc/VTA and striatum in humans has been shown to be correlated with the degree to which human participants feel curious [62,70].

While information and extrinsic rewards thus seem to be processed in a similar manner, there are also important differences in the neuronal populations and mechanisms underlying active choices for neutral information and reward-seeking. While all SNc/VTA dopamine neurons tend to respond to unpredicted reward or reward-predicting cues, only a subset of them is activated by novel or novelty-predicting stimuli [49,58,71]. Moreover, in the dorsal striatum of macaque monkeys, and its likely afferent projections from the anterior cingulate cortex, a group of neurons displayed information-anticipatory activity which was largely unaffected by extrinsic reward value [72]. Consistent with this, pharmacological manipulations of this cortico-basal ganglia network affect gaze reaction time specifically for informative cues but not non-informative cues, without affecting overall motivation to perform the task [72].

Distinct processing of information and extrinsic reward has also been demonstrated in the lateral intraparietal cortex (LIP) of macaque monkeys. The LIP is involved in identifying significant locations in visual space and shifting gaze towards them. Stimulus novelty or the expected information gain associated with visual cues reliably enhance the activity of LIP neurons, but independently of the value of upcoming rewards [73–76]. Similarly, human neuroimaging reveals a modulation of parietal cortex by the uncertainty about outcomes but not by expected reward [77].

Active choice for costly/useful information

In the third category of information-seeking behaviors, we include studies in which the choice to collect information has a consequence for the overall gains and losses of the experimental subject. These tasks are motivated by the fact that in the real world, curious exploration often comes at a price. Seeking new information takes time and metabolic energy and might distract curious agents from pursuing extrinsic rewards. To study how agents balance their desires for both extrinsic rewards and information, specific tasks have been developed which introduce positive or negative associations between curiosity-driven actions and the tasks' extrinsic rewards. These experiments demonstrate that humans and animals robustly engage in information seeking while searching for extrinsic rewards [73,78–80]. They even exhibit curiosity in experimental paradigms where curiosity-driven behavior is explicitly penalized by, for example, adding physical punishment [81], delay periods [77], or reward reduction [82]. In a particularly striking example, humans take the risk of receiving a mild electrical shock to get the chance of knowing the secret of a magic trick [81]. In tasks with active choices for costly/useful curiosity, the curiosity level is indeed balanced by both the strength and the direction of the association between curiosity-driven actions and the tasks' extrinsic rewards: for example, the preference for information or novelty decreases by increasing the chance of punishments [81,82] or by decreasing the agents' level of optimism about novel objects [83,84].

The systematic preference for novelty is particularly advantageous in complex environments where rewards are sparse and delivered only after long sequences of actions. In such environments, involving multistep decision-making, human explorers reach rewarding states up to two orders of magnitude faster than artificial agents searching randomly [78]. A preference for exploring novel actions has also been observed in simple tasks with immediate frequent extrinsic rewards (e.g., bandit tasks), which enable humans to efficiently explore different actions and find the one with the highest average reward [79,85,86]. However, whether the principles underlying

curious exploration in **multistep tasks with sparse rewards** also apply to exploration in simpler tasks with immediate frequent rewards needs further investigation.

In general, it has been debated whether useful/costly curiosity is ‘purely’ information-driven or, at least in part, motivated by external incentives [15]. Agents may simply use their curiosity as a heuristic tool to reach task-related goals [87]. For example, in tasks where a reward is associated with a specific location, just searching for new objects and locations would eventually lead to discovering the reward location. Accordingly, a single ‘curiosity mechanism’ might be sufficient to explain both active choices for costly/useful information [78] and ‘purely’ information-driven choices for neutral information [88]. This argument is supported by the observation that similar exploration strategies are used by humans and animals even when such strategies are not optimal or useful for collecting rewards. For example, humans seek novelty to search for rewards even when novelty-seeking leads to distraction by reward-independent stochasticity [83] (also known as the ‘noisy-TV’ problem [89]), and monkeys keep looking for reward-informative cues even when there is no reward-uncertainty [73]. These results suggest that humans and animals employ at least two separate action policies in parallel, one for curiosity and one for seeking reward. In other words, information gets assigned a value on its own, rather than merely modulating reward value.

Consistent with this, reward and information can also be distinguished at the neural level for costly or useful curiosity [90]. Neurons in the orbitofrontal cortex of monkeys trading reward for information about gamble outcomes signaled both the amount of extrinsic reward and informativeness, but in an independent manner [82]. Importantly, since animals can weigh information against or in addition to extrinsic rewards in experimental tasks with costly or useful curiosity, these tasks enable us to study how and where in the brain preferences for extrinsic rewards and information are integrated to influence the behavior. A recent preprint reports a possible site for such integration in the macaque brain. In a decision-making task involving a trade-off between information and reward, neurons in the lateral habenula combined the value of information and extrinsic reward to signal the subjective value of possible choices using a common code [91]. Moreover, perturbation of neural activity in the lateral habenula biased choice behavior in a manner consistent with reducing subjective value.

Computational models of curiosity

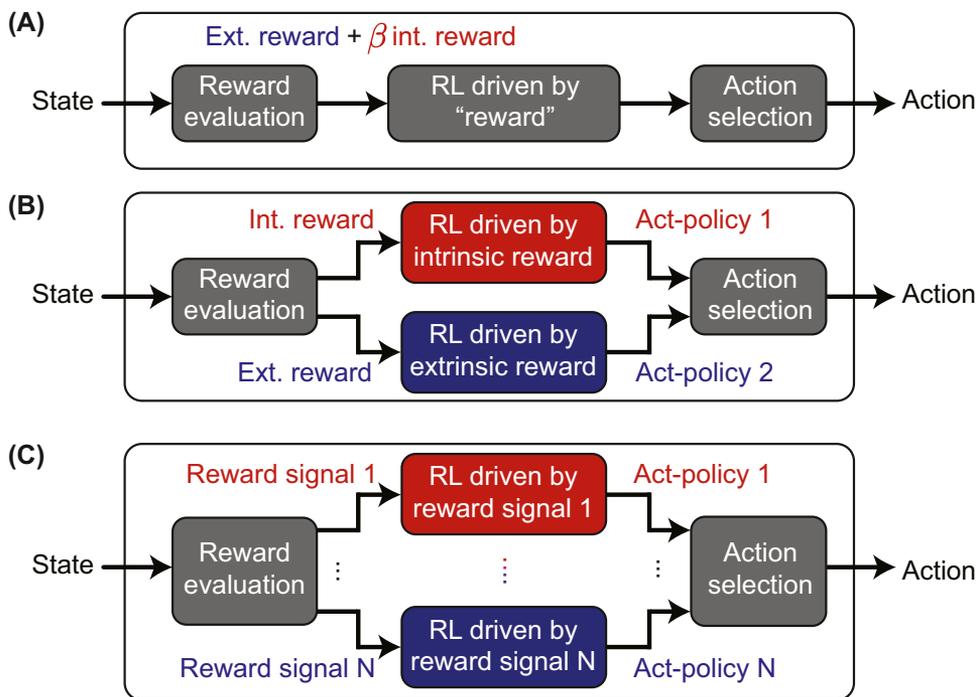
In the final section of this review we discuss recent computational models emerging from the field of artificial intelligence and discuss how they can help organize the richness and diversity of experimental findings in a unified theoretical framework. Consistent with brain imaging and neurophysiology findings, theories of curiosity often interpret curiosity-driven behaviors as actions driven by an **intrinsic reward**-like signal (i.e., a signal generated inside the brain) [4, 15, 92]. Many modern theories rely on this idea and use intrinsically motivated reinforcement learning (RL) algorithms as computational models of curiosity [87, 93]. RL describes the behavior of artificial agents that maximize some ‘primary’ extrinsic rewards (e.g., nutrition or money) by interacting with their environments [94]. Intrinsically motivated RL algorithms [95] assume that there exists one or more additional ‘secondary’ (intrinsically generated) reward signals (e.g., novelty) that are independent of the primary (extrinsic) reward but help exploration for collecting primary rewards [96, 97] and learning complex skills [98, 99]. In this section we discuss how intrinsically motivated RL can describe different expressions of curiosity (reviewed earlier) and what challenges they face in doing so.

Integrating different expressions of curiosity in one theoretical framework

All three categories of curiosity-driven behavior that we have reviewed can be formally described in the intrinsically motivated RL framework: innate, reactive orienting does not need the learning

component of RL algorithms and can be modeled by a fixed action policy to seek immediately available intrinsic rewards defined as surprise, novelty, or a combination of the two (e.g., the surprise-seeking models of visual **attention** [100,101]). However, active choices for neutral information have to be modeled by a learned action policy that maximizes future intrinsic rewards by interacting with the environment: for example, models of seeking progress rate in experiments where subjects freely choose which task to complete [59,102]. Hence, models of reactive orienting and active choices for neutral information differ with respect to the flexibility of their action policies (i.e., fixed versus learned throughout a task) and the time-scales of intrinsic reward-seeking (i.e., one-step versus multistep reward-seeking), but both can be modeled by action policies that are purely driven by intrinsic rewards.

By contrast, modeling exploration in tasks where curiosity is useful or costly needs action policies that are jointly driven by both intrinsic and extrinsic rewards: for example, novelty-seeking models of exploration in environments with sparse rewards [78,83]. Even in the simplistic case of having only one intrinsic reward signal, there are multiple possibilities for how the relationship between intrinsic and extrinsic reward signals can be described in the RL framework (Figure 4). The traditional approach in machine learning is to use a weighted sum of the intrinsic and extrinsic rewards as a single scalar reward signal that drives action selection [89] (Figure 4A). A successful example of using this approach in neuroscience has been in explaining the joint dopamine response to



Trends in Neurosciences

Figure 4. Modern computational models of curiosity use reinforcement learning (RL) and characterize curiosity as a desire for an intrinsic reward signal. An assumption of many models of curiosity is that the reward signal is composed of an intrinsic part and an extrinsic part. Models of the joint influence of extrinsic and intrinsic rewards on action policy either (A) use a weighted sum (with weighting factor β) of the two reward signals as a single scalar reward signal that drives action selection, or (B) assume that different reward signals are processed in parallel and result in separate action policies which are only then combined to drive action selection. (C) The latter can be generalized to include N parallel reward signals (e.g., nutrition, sex, novelty, surprise, etc.) that result in different action policies. Actions are taken based on competition between these policies.

extrinsic rewards and novelty, in a single ‘reward’-seeking framework [103]. Although the exceptional simplicity of this approach makes it attractive for computational modeling, it is not consistent with the experimental evidence reviewed herein for separate neural pathways of reward- and information-driven behavior in the brain [58,73,78,82,91,104], and does not allow for rapid and flexible arbitration between seeking different reward signals observed in human experimental subjects [78,83]. Hence, despite its initial success, this approach cannot explain all different expressions of curiosity reviewed in the previous sections.

An alternative approach is to consider two parallel processing pathways for the two reward signals, resulting in two separate action policies that are only in later processing steps combined for action-selection [78,104,105] (Figure 4B). From the perspective of viewing the brain as a ‘mixture of experts’ [106], each action policy can be seen as an expert, where one advocates for curiosity-driven exploration (seeking intrinsic reward) and the other for seeking extrinsic rewards. The competition between the two experts and their weights on decision-making can be controlled by multiple factors such as optimism about availability of rewards [83], reliability of different experts [107], and cognitive load [108]. Importantly, this approach can be generalized to include more than two parallel action policies and to model action-selection in tasks where behavior is driven by multiple parallel intrinsic and extrinsic reward signals [14,61] (Figure 4C). In principle, each parallel action policy can be driven by a complex reward signal that is a different weighted combination of multiple simple intrinsic and extrinsic reward signals. The joint dopamine response to extrinsic reward and novelty [103] is an example mechanism for evaluating such complex reward signals.

What is the ‘intrinsic reward’?

The central challenge in modeling any of the three categories of curiosity-driven behavior is (i) to characterize the intrinsic reward signals in different tasks, to answer the question ‘what are we curious about?’, and (ii) to identify the benefits of seeking intrinsic rewards in various tasks, to answer the question ‘why are we curious?’. Depending on which question is prioritized, different computational models of curiosity can be classified into bottom-up and top-down models.

Bottom-up models of curiosity start with the question ‘what are we curious about?’ and characterize the intrinsic reward signal based on inspirations from experimental observations [93]. These models have been successful in explaining and predicting many aspects of curiosity-driven behavior in humans and animals by considering novelty [78], surprise [61], **information gain** [109,110], **progress rate** [59,102], or **empowerment** [111,112] as the intrinsic reward signal. However, a drawback of the existing bottom-up models has been their limited focus on individual experimental tasks: for example, showing that human exploration is best explained by novelty-seeking models in some experiments [78,83] but by models driven by information gain or progress rate in some other tasks [59,109]. Importantly, no single unifying bottom-up model has been proposed to explain the rich and diverse experimental findings reviewed herein, and it has remained unclear why and how intrinsic rewards driving exploration differ from one task to another. Simultaneous modeling of multiple experiments with potentially many drives of curiosity is necessary to address these questions. A potential candidate model of multiple experiments includes parallel action policies (similar to Figure 4C), where each action policy is driven by a different intrinsic reward signal (or a different mixture of them).

The question ‘why are we curious?’ in bottom-up models is often approached by heuristic reasoning. For example, it has been argued that seeking *novelty* allows agents to explore their environment efficiently and discover otherwise unknown states with high primary reward values [78,83], seeking *progress rate* enables agents to develop skills that can be used for planning in

future [87], and seeking *empowerment* enables agents to gain control over their environment [112]. While these arguments are supported by the success of modern machine-learning algorithms [96,113,114] and have some theoretical foundations [115], bottom-up models do not directly address why curiosity has evolved in humans and animals.

Top-down models, however, start with the question ‘why are we curious?’, define an overarching long-term goal for agents (i.e., assuming the answer to the ‘why’ question), and derive curiosity as a near-optimal mechanism to achieve this goal. Top-down models do not directly describe curiosity-driven behavior but rather describe the process through which such behaviors would emerge.

Typical examples of top-down models start with the assumption that an agent’s goal is to gain knowledge of how to predict ‘the correct action’ (e.g., the action with the highest average extrinsic reward) from incomplete knowledge of the environment [84,116]. Even if there is an optimal strategy for achieving this goal, it is *a priori* unclear how this strategy is linked to curiosity. For example, seeking novelty and seeking information gain can be seen as special cases of a more general optimal exploration strategy, but with different assumptions about the task’s structure [116]. If this can be generalized further, then the fact that humans seek novelty in some tasks but information gain in others can be explained by the observation that the optimal exploration strategy is different for different tasks. In a model with parallel action policies driven by different intrinsic rewards (Figure 4C), such a variable strategy can be implemented as a task-dependent modulation of how much different policies influence action selection. The main challenge in top-down models of curiosity is to formulate assumptions on the task structure and the agents’ prior knowledge thereof, which is non-trivial. Moreover, since humans have been shown to not always choose the optimal exploration strategy [83], top-down models of curiosity also face a challenge for explaining the reasons for suboptimality (see [117] for an example).

Conceptual evolution-based examples of top-down models consider curiosity as the outcome of an evolutionary process for maximizing population-level performance [95] (e.g., survival rate of the species). Measures of evolutionary performance are evaluated for an entire population of agents and only after several generations. Therefore, the actions of an individual cannot simply be optimized with respect to such a performance measure. However, an individual agent’s internal reward signal can serve as an approximate indicator of the evolutionary performance measure and guide the agent’s actions, throughout its life, such that the population’s cumulative behavior maximizes the population-level performance over generations. Accordingly, evolution-based models of curiosity assume that evolution optimizes the agents’ internal intrinsic reward signals over generations [95]. While recent studies in machine learning have shown that such approaches can lead to intrinsic reward signals that can be interpreted as drives of curiosity (e.g., signals similar to surprise [118,119]), it is not trivial whether evolution-based reward signals are necessarily linked to curiosity. Overall, evolution-based models should be seen at this stage as conceptual theories, and it is difficult to see experimental predictions of these models that can be definitively tested using existing tools.

Concluding remarks

Exploratory behaviors are central to the study of curiosity, since they allow us to infer an otherwise hidden state of being curious from quantitative, experimental observations. Much progress has been made in identifying neural circuits related to curiosity. However, many fundamental questions remain to be addressed (see [Outstanding questions](#)). What is the specific contribution of a particular brain region or cellular ensemble to measurable outcomes such as orienting? Is an area just relaying information from input to output structures? Or is it involved in a critical

Outstanding questions

How do sensory areas detect novelty and surprise?

How do dynamic signals in neural circuits implement computations related to curiosity?

Where in the brain are different neuronal networks shared and where are they distinct between reactive orienting and active choices for information?

Is there a single ‘curiosity mechanism’ to explain both active choices for costly/useful information and active choices for neutral information?

Are mechanisms underlying curiosity the same or different between multistep decision-making tasks with sparse rewards, on the one hand, and simple exploration tasks with immediate, frequent rewards on the other hand?

How many principally different types of intrinsic reward are involved in different behavioral patterns of curiosity? How do these intrinsic reward signals interact to drive action-selection?

How can computational models be used to design hypothesis-driven experiments and potentially link otherwise separated experimental phenomena?

computation such as detecting novelty, detecting surprise, or integrating information and reward? Causal neuroscience tools do not necessarily resolve these questions since any manipulation along the critical path from sensory input to motor output will have disruptive consequences. It is thus crucial to collect high-resolution neural data [120] to begin to understand how dynamic signals in neural circuits implement computations related to curiosity. Complementary brain-wide measurements [121] provide the opportunity to differentiate neuronal networks that are specifically involved in processing novelty, surprise, and neutral, costly, or useful information from those that are shared across different types of curiosity-driven exploration.

Recently developed computational models of curiosity which are based on intrinsically motivated reinforcement learning promise to inform our understanding of behavioral performance and neural signals. These models describe curiosity-driven behavior as actions driven by some intrinsic reward signals that are either defined based on experimental observations (bottom-up theories) or derived as optimal solutions to some optimization problems (top-down theories). Both approaches face multiple challenges (see Outstanding questions). Importantly, is there a single intrinsic reward signal that drives all different patterns of behavior? Or are there multiple reward signals, one for each behavioral pattern? If the latter is true, then how do different reward signals interact? Different answers to these questions lead to different experimental predictions concerning both behavioral expressions and neural mechanisms of curiosity. Computational models may thus help design hypothesis-driven experiments and potentially link otherwise separated experimental phenomena.

In conclusion, we suggest that a phenomenological approach centered on behavioral expressions of curiosity might be useful to inform future studies of curiosity, aimed at connecting behavior, neural circuit investigations, and computational modeling.

Acknowledgments

A.M. thanks Johanni Brea and Sophia Becker for many useful discussions on related topics. This work was supported by the Swiss National Science Foundation No. 200020_207426 (A.M. and W.G.), KU Leuven C14/21/111 (S.H.), Fonds Wetenschappelijk Onderzoek-Vlaanderen (FWO-Flanders) postdoctoral fellowship 1276122N (K.K.) and grant G097022N (S.H.).

Declaration of interests

The authors declare no competing interests in relation to this work.

References

- James, W. (1899) *Talks to Teachers on Psychology – and to Students on Some of Life's Ideals*, Metropolitan Books/Henry Holt
- Pavlov, I.P. and Anrep, G.V. (2015) *Conditioned Reflexes: an Investigation of the Physiological Activity of the Cerebral Cortex*, Martino Publishing
- Dashiell, J.F. (1925) A quantitative demonstration of animal drive. *J. Comp. Psychol.* 5, 205
- Berlyne, D.E. (1954) A theory of human curiosity. *Br. J. Psychol. Gen. Sect.* 45, 180–191
- Heyes, C. and Huber, L. (2000) *The Evolution of Cognition*, MIT Press
- Hebb, D.O. (1955) Drives and the C. N. S. (conceptual nervous system). *Psychol. Rev.* 62, 243–254
- Hunt, J. (1963) Motivation inherent in information processing and action. In *Motiv. Soc. Interaction* (1963) (Harvey, O.J., ed.), pp. 35–94, Ronald Press
- Loewenstein, G. (1994) The psychology of curiosity: a review and reinterpretation. *Psychol. Bull.* 116, 75–98
- Sokolov, E.N. (1963) Higher nervous functions: the orienting reflex. *Annu. Rev. Physiol.* 25, 545–580
- Deschênes, M. et al. (2012) Sniffing and whisking in rodents. *Curr. Opin. Neurobiol.* 22, 243–250
- Esquivelzeta Rabell, J. et al. (2017) Spontaneous rapid odor source localization behavior requires interhemispheric communication. *Curr. Biol.* 27, 1542–1548.e4
- Fantz, R.L. (1964) Visual experience in infants: decreased attention to familiar patterns relative to novel ones. *Science* 146, 668–670
- Bunzeck, N. et al. (2009) Reward motivation accelerates the onset of neural novelty signals in humans to 85 milliseconds. *Curr. Biol. CB* 19, 1294–1300
- Ahmadlou, M. et al. (2021) A cell type-specific cortico-subcortical brain circuit for investigatory and novelty-seeking behavior. *Science* 372, eabe9681
- Gottlieb, J. and Oudeyer, P.-Y. (2018) Towards a neuroscience of active sampling and curiosity. *Nat. Rev. Neurosci.* 19, 758–770
- Crane, A.L. et al. (2020) An ecological framework of neophobia: from cells to organisms to populations. *Biol. Rev.* 95, 218–231
- Modirshanechi, A. et al. (2023) Surprise and novelty in the brain. *Curr. Opin. Neurobiol.* 82, 102758
- Schomaker, J. and Meeter, M. (2015) Short- and long-lasting consequences of novelty, deviance and surprise on brain and cognition. *Neurosci. Biobehav. Rev.* 55, 268–279

19. Parras, G.G. *et al.* (2017) Neurons along the auditory pathway exhibit a hierarchical organization of prediction error. *Nat. Commun.* 8, 2148
20. Li, L. *et al.* (1993) The representation of stimulus familiarity in anterior inferior temporal cortex. *J. Neurophysiol.* 69, 1918–1929
21. Kato, H.K. *et al.* (2012) Dynamic sensory representations in the olfactory bulb: modulation by wakefulness and experience. *Neuron* 76, 962–975
22. Zhang, K. *et al.* (2022) Surprise and recency in novelty detection in the primate brain. *Curr. Biol.* 32, 2160–2173.e6
23. Fahy, F.L. *et al.* (1993) Neuronal activity related to visual recognition memory: long-term memory and the encoding of recency and familiarity information in the primate anterior and medial inferior temporal and rhinal cortex. *Exp. Brain Res.* 96, 457–472
24. Xiang, J.Z. and Brown, M.W. (1998) Differential neuronal encoding of novelty, familiarity and recency in regions of the anterior temporal lobe. *Neuropharmacology* 37, 657–676
25. Rutishauser, U. *et al.* (2006) Single-trial learning of novel stimuli by individual neurons of the human hippocampus–amygdala complex. *Neuron* 49, 805–813
26. Wilson, F.A.W. and Rolls, E.T. (1993) The effects of stimulus novelty and familiarity on neuronal activity in the amygdala of monkeys performing recognition memory tasks. *Exp. Brain Res.* 93, 367–382
27. Thompson, R.F. and Welker, W.I. (1963) Role of auditory cortex in reflex head orientation by cats to auditory stimuli. *J. Comp. Physiol. Psychol.* 56, 996–1002
28. Albasser, M.M. *et al.* (2015) Perirhinal cortex lesions in rats: novelty detection and sensitivity to interference. *Behav. Neurosci.* 129, 227–243
29. Olarte-Sánchez, C.M. *et al.* (2015) Perirhinal cortex lesions impair tests of object recognition memory but spare novelty detection. *Eur. J. Neurosci.* 42, 3117–3127
30. Honey, R.C. *et al.* (1998) Hippocampal lesions disrupt an associative mismatch process. *J. Neurosci.* 18, 2226–2230
31. Levy, D.A. *et al.* (2004) Impaired odor recognition memory in patients with hippocampal lesions. *Learn. Mem.* 11, 794–796
32. Prusky, G.T. *et al.* (2004) Visual memory task for rats reveals an essential role for hippocampus and perirhinal cortex. *Proc. Natl. Acad. Sci.* 101, 5064–5068
33. Zola, S.M. *et al.* (2000) Impaired recognition memory in monkeys after damage limited to the hippocampal region. *J. Neurosci.* 20, 451–463
34. Nieuwenhuis, S. *et al.* (2011) The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology* 48, 162–175
35. Allen, K.M. *et al.* (2021) Orienting our view of the superior colliculus: specializations and general functions. *Curr. Opin. Neurobiol.* 71, 119–126
36. Hikosaka, O. and Wurtz, R.H. (1983) Visual and oculomotor functions of monkey substantia nigra pars reticulata. II. Visual responses related to fixation of gaze. *J. Neurophysiol.* 49, 1254–1267
37. Hikosaka, O. and Wurtz, R.H. (1983) Visual and oculomotor functions of monkey substantia nigra pars reticulata. IV. Relation of substantia nigra to superior colliculus. *J. Neurophysiol.* 49, 1285–1301
38. McEvain, L.E. *et al.* (2018) Circuits in the rodent brainstem that control whisking in concert with other orofacial motor actions. *Neuroscience* 368, 152–170
39. Inglis, F.M. and Fibiger, H.C. (1995) Increases in hippocampal and frontal cortical acetylcholine release associated with presentation of sensory stimuli. *Neuroscience* 66, 81–86
40. Scourse, N.J.S. and Hinde, R.A. (1973) Habituation to auditory stimuli in mice. *Behaviour* 47, 1–13
41. Bouret, S. and Richmond, B.J. (2008) Relation of locus coeruleus neurons in monkeys to Pavlovian and operant behaviors. *J. Neurophysiol.* 101, 898–911
42. Ljungberg, T. *et al.* (1992) Responses of monkey dopamine neurons during learning of behavioral reactions. *J. Neurophysiol.* 67, 145–163
43. Carter, M.E. *et al.* (2010) Tuning arousal with optogenetic modulation of locus coeruleus neurons. *Nat. Neurosci.* 13, 1526–1533
44. Vankov, A. *et al.* (1995) Response to novelty and its rapid habituation in locus coeruleus neurons of the freely exploring rat. *Eur. J. Neurosci.* 7, 1180–1187
45. Takeuchi, T. *et al.* (2016) Locus coeruleus and dopaminergic consolidation of everyday memory. *Nature* 537, 357–362
46. Sara, S.J. *et al.* (1995) Novelty seeking behavior in the rat is dependent upon the integrity of the noradrenergic system. *Cogn. Brain Res.* 2, 181–187
47. Schomaker, J. and Meeter, M. (2018) Predicting the unknown: novelty processing depends on expectations. *Brain Res.* 1694, 140–148
48. Modirshanechi, A. *et al.* (2019) Trial-by-trial surprise-decoding model for visual and auditory binary oddball tasks. *NeuroImage* 196, 302–317
49. Morrens, J. *et al.* (2020) Cue-evoked dopamine promotes conditioned responding during learning. *Neuron* 106, 142–153.e7
50. Rebec, G.V.G.V. *et al.* (1997) Regional and temporal differences in real-time dopamine efflux in the nucleus accumbens during free-choice novelty. *Brain Res.* 776, 61–67
51. Bunzeck, N. and Düzel, E. (2006) Absolute coding of stimulus novelty in the human substantia nigra/VTA. *Neuron* 51, 369–379
52. Lak, A. *et al.* (2016) Dopamine neurons learn relative chosen value from probabilistic rewards. *eLife* 5, e18044
53. Kamiński, J. *et al.* (2018) Novelty-sensitive dopaminergic neurons in the human substantia nigra predict success of declarative memory formation. *Curr. Biol.* 28, 1333–1343.e4
54. Schiemann, J. *et al.* (2012) K-ATP channels in dopamine substantia nigra neurons control bursting and novelty-induced exploration. *Nat. Neurosci.* 15, 1272–1280
55. Da Silva, J.A. *et al.* (2018) Dopamine neuron activity before action initiation gates and invigorates future movements. *Nature* 554, 244–248
56. Akiti, K. *et al.* (2022) Striatal dopamine explains novelty-induced behavioral dynamics and individual variability in threat prediction. *Neuron* 110, 3789–3804.e9
57. Valdés-Baizabal, C. *et al.* (2020) Dopamine modulates subcortical responses to surprising sounds. *PLoS Biol.* 18, e3000744
58. Ogasawara, T. *et al.* (2022) A primate temporal cortex–zona incerta pathway for novelty seeking. *Nat. Neurosci.* 25, 50–60
59. Ten, A. *et al.* (2021) Humans monitor learning progress in curiosity-driven exploration. *Nat. Commun.* 12, 5972
60. Cen, D. *et al.* (2021) Anticipation of novel environments enhances memory for incidental information. *Learn. Mem.* 28, 254–259
61. Kobayashi, K. *et al.* (2019) Diverse motives for human curiosity. *Nat. Hum. Behav.* 3, 587–595
62. Gruber, M.J. *et al.* (2014) States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron* 84, 486–496
63. Bromberg-Martin, E.S. and Hikosaka, O. (2009) Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron* 63, 119–126
64. Baranes, A. *et al.* (2015) Eye movements reveal epistemic curiosity in human observers. *Vis. Res.* 117, 81–90
65. Bromberg-Martin, E.S. and Hikosaka, O. (2011) Lateral habenula neurons signal errors in the prediction of reward information. *Nat. Neurosci.* 14, 1209–1216
66. Schultz, W. *et al.* (1997) A neural substrate of prediction and reward. *Science* 275, 1593–1599
67. Wittmann, B.C. *et al.* (2007) Anticipation of novelty recruits reward system and hippocampus while promoting recollection. *NeuroImage* 38, 194–202
68. Guitart-Masip, M. *et al.* (2010) Contextual novelty changes reward representations in the striatum. *J. Neurosci.* 30, 1721–1726
69. Wittmann, B.C. *et al.* (2008) Striatal activity underlies novelty-based choice in humans. *Neuron* 58, 967–973
70. Kang, M.J. *et al.* (2009) The wick in the candle of learning: epistemic curiosity activates reward circuitry and enhances memory. *Psychol. Sci.* 20, 963–973
71. Menegas, W. *et al.* (2017) Opposite initialization to novel cues in dopamine signaling in ventral and posterior striatum in mice. *eLife* 6, e21886

72. White, J.K. *et al.* (2019) A neural network for information seeking. *Nat. Commun.* 10, 5168
73. Foley, N.C. *et al.* (2017) Parietal neurons encode expected gains in instrumental information. *Proc. Natl. Acad. Sci.* 114, E3315–E3323
74. Foley, N.C. *et al.* (2014) Novelty enhances visual salience independently of reward in the parietal lobe. *J. Neurosci.* 34, 7947–7957
75. Horan, M. *et al.* (2019) Parietal neurons encode information sampling based on decision uncertainty. *Nat. Neurosci.* 22, 1327–1335
76. Li, Y. *et al.* (2022) Uncertainty modulates visual maps during noninstrumental information demand. *Nat. Commun.* 13, 5911
77. van Lieshout, L.L.F. *et al.* (2018) Induction and relief of curiosity elicit parietal and frontal activity. *J. Neurosci.* 38, 2579–2588
78. Xu, H.A. *et al.* (2021) Novelty is not surprise: human exploratory and adaptive behavior in sequential decision-making. *PLoS Comput. Biol.* 17, e1009070
79. Wilson, R.C. *et al.* (2014) Humans use directed and random exploration to solve the explore–exploit dilemma. *J. Exp. Psychol. Gen.* 143, 2074–2081
80. Dubois, M. *et al.* (2021) Human complex exploration strategies are enriched by noradrenaline-modulated heuristics. *eLife* 10, e59907
81. Lau, J.K.L. *et al.* (2020) Shared striatal activity in decisions to satisfy curiosity and hunger at the risk of electric shocks. *Nat. Hum. Behav.* 4, 531–543
82. Blanchard, T.C. *et al.* (2015) Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. *Neuron* 85, 602–614
83. Modirshanechi, A. *et al.* (2022) The curse of optimism: a persistent distraction by novelty. *bioRxiv* Published online July 5, 2022. <https://doi.org/10.1101/2022.07.05.498835>
84. Gershman, S.J. and Niv, Y. (2015) Novelty and inductive generalization in human reinforcement learning. *Top. Cogn. Sci.* 7, 391–415
85. Schulz, E. and Gershman, S.J. (2019) The algorithmic architecture of exploration in the human brain. *Curr. Opin. Neurobiol.* 55, 7–14
86. Wu, C.M. *et al.* (2018) Generalization guides human exploration in vast decision spaces. *Nat. Hum. Behav.* 2, 915–924
87. Oudeyer, P.-Y. (2018) Computational theories of curiosity-driven learning. *arXiv* 1802.10546v2
88. Montgomery, K.C. (1954) The role of the exploratory drive in learning. *J. Comp. Physiol. Psychol.* 47, 60–64
89. Aubret, A. *et al.* (2023) An information-theoretic perspective on intrinsic motivation in reinforcement learning: a survey. *Entropy* 25, 327
90. Filimon, F. *et al.* (2020) The ventral striatum dissociates information expectation, reward anticipation, and reward receipt. *Proc. Natl. Acad. Sci.* 117, 15200–15208
91. Bromberg-Martin, E.S. *et al.* (2022) A neural mechanism for conserved value computations integrating information and rewards. *bioRxiv* Published online August 15, 2022. <https://doi.org/10.1101/2022.08.14.503903>
92. Schmidhuber, J. (2010) Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE Trans. Auton. Ment. Dev.* 2, 230–247
93. Murayama, K. (2022) A reward-learning framework of knowledge acquisition: an integrated account of curiosity, interest, and intrinsic–extrinsic rewards. *Psychol. Rev.* 129, 175–198
94. Sutton, R.S. and Barto, A.G. (2018) *Reinforcement Learning: An Introduction*, MIT press
95. Singh, S. *et al.* (2010) Intrinsically motivated reinforcement learning: an evolutionary perspective. *IEEE Trans. Auton. Ment. Dev.* 2, 70–82
96. Bellemare, M. *et al.* (2016) Unifying count-based exploration and intrinsic motivation. *Adv. Neural Inf. Proces. Syst.* 29, 1471–1479
97. Ostrovski, G. *et al.* (2017) Count-based exploration with neural density models. *Proceedings of the 34th International Conference on Machine Learning* PMLR 70:2721–2730
98. Sekar, R. *et al.* (2020) Planning to Explore via Self-Supervised World Models. In *Proceedings of Machine Learning Research* (119), pp. 8583–8592
99. Mendonca, R. *et al.* (2021) Discovering and achieving goals via world models. *Adv. Neural Inf. Proces. Syst.* 34, 24379–24391
100. Itti, L. and Baldi, P. (2009) Bayesian surprise attracts human attention. *Vis. Res.* 49, 1295–1306
101. Baldi, P. and Itti, L. (2010) Of bits and wows: a Bayesian theory of surprise with applications to attention. *Neural Netw.* 23, 649–666
102. Poli, F. *et al.* (2022) Contributions of expected learning progress and perceptual novelty to curiosity-driven exploration. *Cognition* 225, 105119
103. Kakade, S. and Dayan, P. (2002) Dopamine: generalization and bonuses. *Neural Netw.* 15, 549–559
104. Cogliati Dezza, I. *et al.* (2022) Independent and interacting value systems for reward and information in the human brain. *eLife* 11, e66358
105. Kim, Y. *et al.* (2019) *Curiosity-bottleneck: exploration by distilling task-specific novelty*. Proceedings of the 36th International Conference on Machine Learning 97 pp. 3379–3388
106. O’Doherty, J.P. *et al.* (2021) Why and how the brain weights contributions from a mixture of experts. *Neurosci. Biobehav. Rev.* 123, 14–23
107. Daw, N. *et al.* (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711
108. Piray, P. and Daw, N.D. (2021) Linear reinforcement learning in planning, grid fields, and cognitive control. *Nat. Commun.* 12, 4942
109. Horvath, L. *et al.* (2021) Human belief state-based exploration and exploitation in an information-selective symmetric reversal bandit task. *Comput. Brain Behav.* 4, 442–462
110. Nelson, J.D. (2005) Finding useful questions: on Bayesian diagnosticity, probability, impact, and information gain. *Psychol. Rev.* 112, 979–999
111. Brändle, F. *et al.* (2023) Empowerment contributes to exploration behaviour in a creative video game. *Nat. Hum. Behav.* 7, 1481–1489
112. Klyubin, A.S. *et al.* (2005) *Empowerment: a universal agent-centric measure of control*. 2005 IEEE Congress on Evolutionary Computation 1 pp. 128–135
113. Pathak, D. *et al.* (2019) *Self-supervised exploration via disagreement*. Proceedings of the 36th International Conference on Machine Learning 97 pp. 5062–5071
114. Mohamed, S. and Jimenez Rezende, D. (2015) Variational information maximisation for intrinsically motivated reinforcement learning. *Adv. Neural Inf. Proces. Syst.* 28, 2125–2133
115. Strehl, A.L. and Littman, M.L. (2008) An analysis of model-based interval estimation for Markov decision processes. *J. Comput. Syst. Sci.* 74, 1309–1331
116. Dube, R. and Griffiths, T.L. (2019) Reconciling novelty and complexity through a rational analysis of curiosity. *Psychol. Rev.* 127, 455–476
117. Binz, M. and Schulz, E. (2022) Modeling human exploration through resource-rational reinforcement learning. *Advances in Neural Information Processing Systems* 35, 31755–31768
118. Zheng, Z. *et al.* (2020) *What can learned intrinsic rewards capture?*. Proceedings of the 37th International Conference on Machine Learning 119 pp. 11436–11446
119. Alet, F. *et al.* (2020) Meta-learning curiosity algorithms. *International Conference on Learning Representations* <https://openreview.net/forum?id=BygdixHFDS>
120. Steinmetz, N.A. *et al.* (2021) Neuropixels 2.0: a miniaturized high-density probe for stable, long-term brain recordings. *Science* 372, eabf4588
121. Brunner, C. *et al.* (2020) A platform for brain-wide volumetric functional ultrasound imaging and analysis of circuit dynamics in awake mice. *Neuron* 108, 861–875.e7
122. Modirshanechi, A. *et al.* (2022) A taxonomy of surprise definitions. *J. Math. Psychol.* 110, 102712
123. Antony, J.W. *et al.* (2021) Behavioral, physiological, and neural signatures of surprise during naturalistic sports viewing. *Neuron* 109, 377–390.e7
124. Squires, K.C. *et al.* (1976) The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science* 193, 1142–1146
125. Barto, A. *et al.* (2013) Novelty or Surprise? *Front. Psychol.* 4, 907
126. Jaegle, A. *et al.* (2019) Visual novelty, curiosity, and intrinsic reward in machine learning and the brain. *Curr. Opin. Neurobiol.* 58, 167–174