



**University of  
Zurich** UZH

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2024

---

## **Rationality, preferences, and emotions with biological constraints: it all starts from our senses**

Polanía, Rafael ; Burdakov, Denis ; Hare, Todd A

DOI: <https://doi.org/10.1016/j.tics.2024.01.003>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-259633>

Journal Article

Published Version



The following work is licensed under a Creative Commons: Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.

Originally published at:

Polanía, Rafael; Burdakov, Denis; Hare, Todd A (2024). Rationality, preferences, and emotions with biological constraints: it all starts from our senses. *Trends in Cognitive Sciences*, 28(3):264-277.

DOI: <https://doi.org/10.1016/j.tics.2024.01.003>

Review

# Rationality, preferences, and emotions with biological constraints: it all starts from our senses

Rafael Polanía,<sup>1,\*</sup> Denis Burdakov,<sup>2</sup> and Todd A. Hare<sup>3</sup>

**Is the role of our sensory systems to represent the physical world as accurately as possible? If so, are our preferences and emotions, often deemed irrational, decoupled from these 'ground-truth' sensory experiences? We show why the answer to both questions is 'no'. Brain function is metabolically costly, and the brain loses some fraction of the information that it encodes and transmits. Therefore, if brains maximize objective functions that increase the fitness of their species, they should adapt to the objective-maximizing rules of the environment at the earliest stages of sensory processing. Consequently, observed 'irrationalities', preferences, and emotions stem from the necessity for our early sensory systems to adapt and process information while considering the metabolic costs and internal states of the organism.**

## What is the role of our senses?

Our subjective experiences, which arise from perceiving the world through our senses, are generally thought to be grounded in a faithful representation of the physical world. These percepts are transmitted to downstream brain structures that interpret these signals and ultimately determine how we act and learn. According to this view, the formation of our subjective values, preferences, and emotional experiences should not be biased by our early sensory systems if they fulfill their role – to represent the world as accurately as biological limitations allow.

One may think that this line of argumentation makes sense, as a good measurement instrument (a sensor) must reliably measure the environmental variable that it was built for. For instance, if an engineer builds a voltmeter, she would like the device to faithfully map the physical input (i.e., the voltage) to the needle position indicating the measured value. The voltmeter does not know or care to what end the engineer uses the measured voltage. Indeed, the engineer wants the position of the needle on the measurement scale to reflect the physical unit as precisely as possible, with an irreducible error tolerance given that the sensor cannot be infinitely precise. This implies that the sensed measurements are only useful to the degree to which they are accurate over a given measurement range (Figure 1). However, contrary to this example, in which the engineer and sensor are considered as separate entities, sensory systems in living organisms are fully integrated into one deeply interconnected structure – the brain. Thus, we pose the following questions: (i) is it a reasonable goal of our sensory systems to invest their limited resources in representing the world as accurately as possible irrespective of the organism's goals? (ii) Do these early percepts play a role in determining our emotions, preferences, and behavior?

These questions have led to heated debates among scientists and philosophers across various fields including neuroscience, psychology, economics, and evolutionary biology [1–4]. Some

## Highlights

Recent theories and empirical evidence demonstrate that early sensory systems should encode abstract representations beyond 'realism'.

This flexibility significantly influences our preferences and supposedly irrational behavior, as well as our emotional categorization, aesthetics, pleasure, and risk attitudes.

However, these views leave it unclear at which level of the neural processing stream, and at which temporal scale, organisms should adopt efficient adaptations of the sensory information.

Survival and optimal behavior also hinge on a 'sixth sense' – sensing the internal state of the body. Emerging evidence suggests that early sensory encoding of these signals efficiently helps to extract arousal, emotions, and behavioral rules.

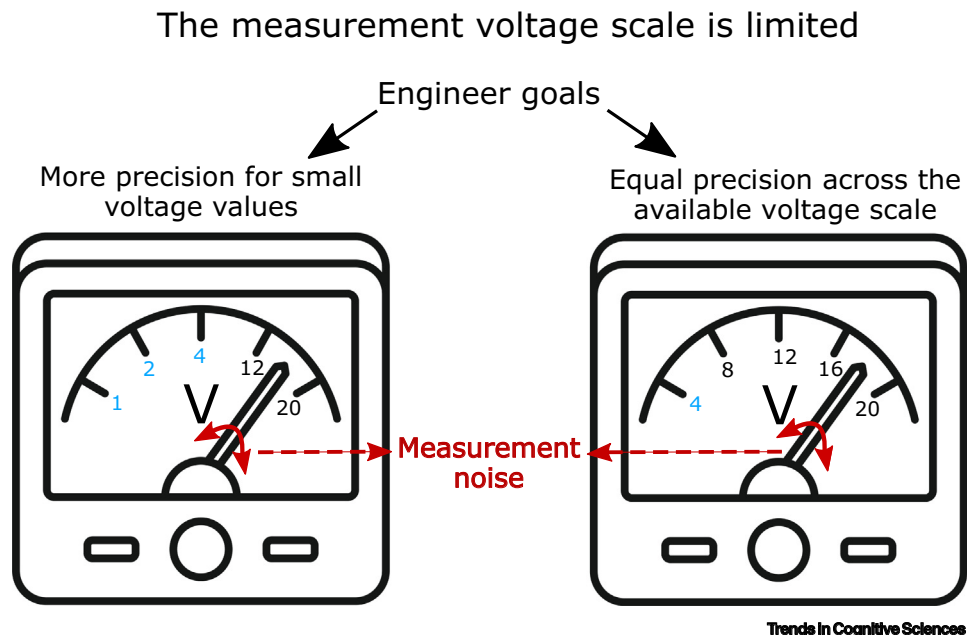
<sup>1</sup>Decision Neuroscience Laboratory, Department of Health Sciences and Technology, ETH, Zurich, Zurich, Switzerland

<sup>2</sup>Neurobehavioral Dynamics Laboratory, Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland

<sup>3</sup>Zurich Center for Neuroeconomics, Department of Economics, University of Zurich, Zurich, Switzerland

\*Correspondence: [rafael.polania@hest.ethz.ch](mailto:rafael.polania@hest.ethz.ch) (R. Polanía).





**Figure 1. Sensor design.** When designing a sensor, for instance, a voltmeter, the precision of the measurements made by the engineer depends on her goals given that the voltage scale is limited. This can be exemplified with the following two hypothetical scenarios, where we assume that the noise of the measurement needle is small but constant across the measurement scale. On the one hand, if the engineer needs to be more precise at low voltage values, then the optimal strategy is to design the voltmeter such that more of the available measurement space is dedicated to measuring small voltage values. However, this comes at the cost of losing precision at higher voltage values (left). On the other hand, if the engineer wishes to have equal precision across the voltage space, the measurement readouts should be equally spread across the measurement space (right). This comes at the cost of losing precision at small voltage values relative to the scenario depicted in the left panel. We argue here that similar optimization strategies are continuously at play in our sensory systems which are also limited in their operation range and the amount of information processing resources that can be dedicated to sensory encoding.

researchers support the idea that organisms should represent objects as they exist in the world [5,6]. Others argue that our perceptual representations should generally be different from the actual physical world and that these representations should be mapped to the utility they offer to the agents [7–9], their action affordances [10], and their interoceptive signals and emotional states [11].

Our goal is not to review this large and multidimensional debate. Instead, we focus on a premise that has been largely missing in the previous discussions, which often ignore constraints of information processing in biological systems from the formal point of view. Given that brain function is metabolically costly, and the brain will always lose some fraction of the information it encodes and transmits at each stage of information processing, sensory systems should adapt to the objective-maximizing rules of a particular environment and internal states of the organism at the earliest stages of sensory processing.

We start by highlighting advances in mathematical formalisms designed to study information processing constraints in biological systems and machines. These developments have played a key role in steering a research agenda across diverse disciplines, ultimately supporting our central thesis. Subsequently, we discuss recent research where such formalisms yield explicit predictions regarding the influence of adaptation of early sensory information processing on our goal-

oriented behavior. This, in turn, sheds light on what is often characterized as 'irrational' behavior during decisions guided by rewards. While acknowledging that the formalism may not yet be sufficiently tractable to be applied to all facets of our premise, we also demonstrate how recent advances in neuroimaging techniques and data analysis tools are starting to reveal profound connections between the neural representations of memories and emotions within our earliest sensory systems and the consequential impact on our actions.

### We are finite beings

When trying to understand how and why we perceive the world and behave in the way we do, two facts that hold for all living organisms must be considered. First, we are finite beings with a limited metabolic budget for interacting with the environment, and thus with a restricted capacity to process environmental and interoceptive signals. Second, sensing-output systems, such as the brain, typically act as noisy communication channels that always lose information during transmission along the way from the different stages of sensory processing to the generation of motor outputs and further interpretation of feedback signals that guide learning.

One attempt to study the interaction between perception and behavior proposes that, over the course of evolution, the nervous system has adopted computational strategies that incorporate environmental statistics and sensory signal uncertainty. Under this framework, the brain learns a statistical model of the world, and integrates this knowledge with imperfect sensory data for optimal computations amid the uncertainty and unreliability of our sensors [12,13]. Although this line of work has been instrumental in explaining many of the biases and variability of behaviors observed during sensory perception, this approach does not explicitly consider the biological limitations of information processing in the nervous system or the significant cognitive burden required to sense, learn, and act [14].

The incorporation of these concepts can be studied with the use of information-theoretical methods which can be used as a formal tool to understand our resource-limited neurobiology. The adoption of information theory in neurobiology led to the postulation of the efficient coding hypothesis introduced by Attneave and Barlow in the late 1950s [15,16]. This hypothesis can be summarized in the following sentence written by Horace Barlow: 'The hypothesis is that sensory relays recode sensory messages so that their redundancy is reduced but comparatively little information is lost' ([16], p. 225).

In other words, this entails that a sensor of the organism will try its best to transmit information from the environment as reliably as possible (also known as information maximization transmission), given that the sensor will never be perfect owing to physical and metabolic constraints. Interestingly, in the very same document, Horace Barlow also writes ([16], p. 219): 'The primary effect of the sensory messages an animal receives is not to enrich its subjective experience of the world but to modify its behavior in such a way that it and its species have a greater chance of survival'.

However, aspects of cognition and goal-oriented behavior were not directly considered in subsequent neurophysiological tests of the efficient coding hypothesis. This is critical, as it can be formally shown that information maximization transmission will not necessarily maximize the chances of survival of the organism [17,18]. This already suggests that representing the world as accurately as possible given our biological constraints is probably not the purpose of our sensory systems.

Nonetheless, the concepts and methods of information theory are not solely restricted to information maximization of sensory systems, as they can be extended to studying agents as processing systems that interact with the environment through sensing action loops [14]. The primary

advantage of this methodology lies in its capacity to generate universally applicable quantitative statements that are valid for all systems, whether biological or artificial. This is achieved with minimal assumptions regarding the specific realization and cognitive mechanisms underlying such systems.

Initially, some forms of these proposals were formally incorporated in the study of how to understand our resource-limited cognition which started with the influential work by Simon – interestingly also in the 1950s – who argued that rational decision strategies must be adapted to both the structure of the environment and the cognitive limitations of the mind [19,20]. A large body of evidence supports the idea that organisms optimize costs and benefits between the precision with which the brain represents value signals and the biological costs of that precision [21–23]. Increasing representational precision may reduce perceptual biases [24–26] and increase choice consistency [27–32], but the metabolic cost of increased precision is significant [33–37]. However, modern resource-rational analysis, known in economics as rational inattention (Box 1), often leaves it unclear at which level of the neural processing stream organisms adopt efficient adaptation and processing of the incoming sensory information to guide their bounded-rational behavior.

Assuming that the goal of organisms is to maximize some objective or utility function [38], we postulate the following hypotheses.

- (i) Following the efficient coding and the bounded-rationality theoretical frameworks, we argue that it is more efficient for the brain to adapt to the objective-maximizing rules of a particular environment at the earliest stages of sensory processing if the system has the ability to modify information-processing strategies at this point in the information processing pipeline. One reason is that biological transmission channels are imprecise, and information will be unavoidably lost in the transmission pipeline in the brain [39]. A second reason is the key role that the adaptability of sensory systems may play in evolution and the resulting behavior [40]. Although a sensor may be minimally optimal for a specific environment, it could still function effectively in different niches. This adaptability allows a somewhat independent evolution of sensor components, and increases the likelihood of individual evolutionary steps. Even if a sensor is minimal for achieving optimal performance in one scenario, it may still provide enough information for successful operation in different contexts, thereby enabling similar levels of effectiveness [14].
- (ii) When facing a series of decisions, organisms optimize their choice outcomes by limiting the information they encode [21,22], which may generate apparently irrational behavior [23,41] and idiosyncratic preferences due, at least in part, to early encoding limitations – for instance, aversion to risk. Thus, our second hypothesis is that some of the often-observed behaviors and statistical fallacies attributed to downstream processing might unavoidably originate from the way our early sensory codes must operate to guarantee that the fitness objectives of the organism are optimally achieved.

We provide here recent theoretical and empirical evidence that supports these hypotheses.

### Formally studying resource limitations and early sensory processing

The first question we address here is how efficient neural adaptations are formally related to neural limitations and cognitive/metabolic costs. This formalism is often studied in the resource-rational analysis framework [21,22]. In brief, this framework provides tools to study how cognitive and computation constraints can be incorporated into rational principles of decision behavior by determining the (constrained) optimal internal strategies that the organism must adopt to maximize the outcome of its choices. While the adoption of resource-rational frameworks is becoming popular across domains in neuroscience [33,42,43], psychology [29,44–47], and economics

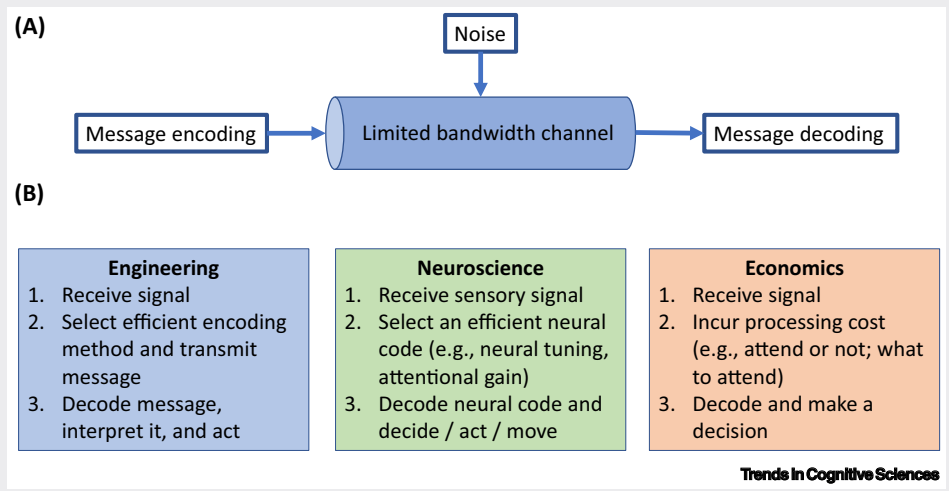
**Box 1. Rational inattention: a cross-domain foundation**

Rational inattention builds on the observation that, contrary to assumptions in classical approaches, decision-makers cannot pay attention to all available information but can choose (or develop learning strategies) to pay more attention to the most important pieces of information. There are well-established models that guide economists on how to deal with imperfect information; however, these models assume that the source of imperfection is of an exogenously given form. This is not what rational inattention deals with. Rational inattention acknowledges that (i) processing information is costly, and (ii) there is a limit to the amount of resources that can be dedicated to acquiring and processing information for decision-making purposes. Based on these constraints, the decision-maker chooses what costly signals to process. Introduced about 20 years ago, rational inattention is considered to be one of the most influential theories developed in economics in the 21st century [48].

Standard applications of rational inattention follow the large literature of information theory. This was originally established and developed by Claude Shannon in 1948 [91] to formally tackle the transmission of information over a communication channel with limited capacity, via a concept known as rate-distortion theory. This theory explores the trade-off between the rate at which information is sent and the fidelity with which it is reconstructed at the receiver. The relationship between rational inattention and rate-distortion theory is twofold.

- (i) Both rational inattention and rate-distortion theory involve optimizing the allocation of scarce resources. In the case of rational inattention the resource is cognitive attention, while in rate-distortion theory it is the capacity of the communication channel (Figure 1).
- (ii) In rational inattention, the decision-maker selectively processes information to maximize its utility given the costs of attention. This is analogous to the rate-distortion trade-off in which the sender aims to convey essential information within the constraints of limited transmission capacity without exceeding an acceptable level of distortion at the receiver.

In the same way as economists adopted the concepts of information theory to develop the formal foundations of rational inattention theory about two decades ago, neuroscientists adopted the same concepts and principles to study how neural systems process, generate, and evolve sensory coding signals to efficiently guide behavior considering the limited metabolic resources of biological systems. In fact, neuroscientists started to investigate and apply information theoretical concepts relatively early on, about a decade after information theory was formally introduced by Claude Shannon in 1948.



**Figure 1. Information transmission in communication systems.** (A) The communication channel is limited, in other words it does not have infinite bandwidth and is corrupted by noise. (B) Steps to solve the information transmission problem in engineering (left) were also adopted to study the efficiency of neural codes in neuroscience (middle) and also inspired the rational inattention theory in economics (right).

[23,31,48–50], in many cases it remains silent as to in which level of the information processing hierarchy in the brain the implementation of efficient information processing takes place, if at all, and whether such optimal information processing codes are related to the idiosyncratic preferences of an organism and/or its adaptation to changes in the environment.

A recent study employed a formal approach to predict from first principles how a (neural) system should allocate its limited information processing resources when agents face decisions in two of the most studied problems in decision-making research [51]. First, accuracy maximization in perceptual discrimination tasks, where organisms typically receive a fixed amount of reward for every correct discrimination they make based on sensory information. Second, reward maximization in economic problems where rewards are monotonically mapped to a given sensory feature (Figure 2). A fundamental aspect to consider when studying these problems is how the typicality of the sensory features that the agents encounter in each environment or context [i.e., the prior distribution  $f(s)$  over all possible sensory inputs  $s$ ] influences the way information processing resources must be allocated. For a two-alternative forced choice (2AFC) task with sensory inputs  $s_1$  and  $s_2$ , the expression that the agent intends to minimize (generally speaking and omitting details) is the penalty for every incorrect decision  $L(s_1, s_2)$ , which formally can be written as follows:

$$\min_h \iint f(s_1, s_2) \cdot P(\text{error}|h(s_1), h(s_2)) \cdot L(s_1, s_2) ds_1 ds_2, \quad [1]$$

where  $P(\text{error}|s_1, s_2)$  is the probability that the agent will be incorrect given that the choice problem occurs with probability  $f(s_1, s_2)$ . Critically, the agent is restricted to allocating its limited neural resources according to a biologically bounded encoding function  $h(s)$  (i.e., the mapping from the environmental signal to the sensor must occur with limited precision because of both physical constraints and metabolic costs, thus unavoidably leading to perceptual errors). Thus, the goal of the agent is to find the function  $h(s)$  that allocates resources optimally such that the average reward loss is minimized.

The results of the formal analyses applied to a sensory discrimination task reveal that, for both problems (accuracy and reward maximization), the allocation of information processing resources

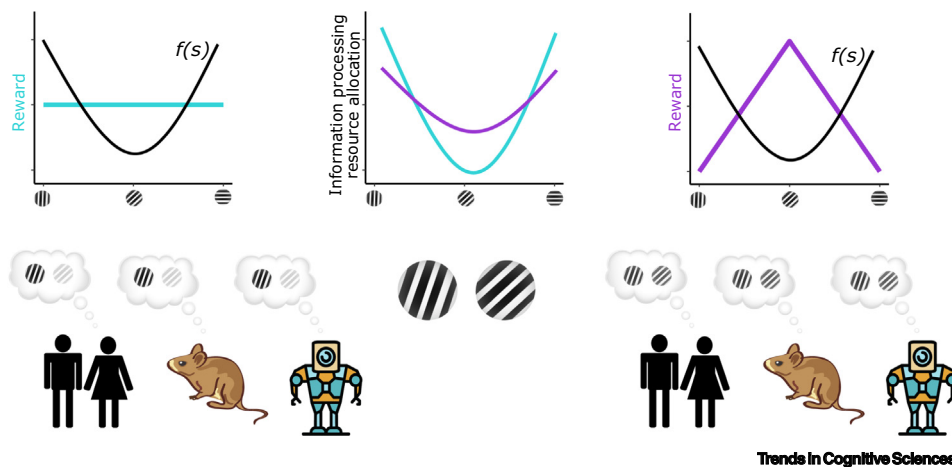


Figure 2. Sensory reliability as a function of environmental priors and stimulus–reward contexts. In the accuracy maximization context, organisms receive a fixed amount of reward for every correct discrimination that they make based on sensory information (left). In the reward maximization context (i.e., a typical economics problem context), reward levels are mapped to a given sensory feature (right). The typicality of the sensory features that the agents encounter in each environment or context [i.e., the prior distribution  $f(s)$  over all possible sensory inputs  $s$ ] influences the way in which information processing resources must be allocated. The optimal internal representation of sensory information when agents aim at maximizing reward in a two-alternative forced choice task reveals that, in the accuracy maximization problem, the agent dedicates more resources to sensory features that occur more often in the environment, which are the features that have lower value in this specific reward maximization problem (cardinal orientations in our example). Furthermore, in the reward maximization context, the agent flattens the resource allocation curve such that it dedicates more resources to features that promise more reward. Recent work has shown that these strategies are adopted by rodents [33], humans, and machines [51].



is higher for sensory features that occur more often in the environment. However, there are also differences in the allocation of such information processing resources. In the accuracy maximization problem, the agent dedicates resources to sensory features based purely on how often it expects them to occur in the environment. By contrast, when sensory features map directly onto levels of reward, the agent must take into account both the frequency with which a feature occurs and the level of reward it yields when dedicating sensory resources. This often means that the agent will dedicate more resources to rare sensory features in a reward maximizing compared with an accuracy maximizing encoding scheme (Figure 2). We note that the solutions described earlier hold for the case where the metabolic costs are nearly negligible; however, surprising results in the allocation of neural resources might be encountered when such costs are considered [33]. When human participants were asked to make decisions in this type of behavioral paradigm (where rewards were mapped to a low-level sensory feature: the orientation of a Gabor patch), the results showed that sensory discrimination performance matched the normative predictions of efficient sensory resource allocation in each context. Crucially, this efficient reorganization appears to occur at retinotopic-specific locations in humans, as inferred from psychophysical analyses, and therefore in the very early stages of sensory processing.

Further tests were carried out with artificial agents to test whether agents must recode their sensory representations to fitness-maximizing schemes to achieve the best decision-making performance at the earliest stages of sensory processing. Interestingly, after training on the same tasks that the humans performed, these neural networks revealed that the contents of sensory information at early stages match both the predictions of the formal mathematical models and the behavior observed in humans [51].

A similar approach was adopted to investigate whether other organisms, in this case mice, also conform to the same class of efficient adaptations [33]. The investigators found that mice adaptively allocate their sensory resources in a way that maximizes reward consumption in novel stimulus–reward association environments. A surprising result, and apparently irrational behavior observed in this study, is that the perception of mice of commonly occurring stimuli was relatively imprecise; however, this apparent statistical fallacy implies 'awareness' and efficient adaptation to their neurocognitive limitations. Taken together, the evidence provided by mice, humans, and machines strongly suggests that information processing resources are flexibly and efficiently reallocated to maximize context-specific objectives at the early stages of sensory processing (Figure 2). This is similar to the way the engineer would design a voltmeter measurement scale according to her goals (Figure 1).

Is it possible to move beyond the single attribute/channel case discussed previously, and introduce hierarchical structures into resource-constrained systems? The principles of rate-distortion theory, of which rational inattention is one application (Box 1), can be extended to network configurations. These networks can be configured to encompass nodes for perception, intermediate processing, and action. By optimizing information flow in such resource-constrained networks, we could gain insight into how information processing limitations affect various nodes, including early sensing nodes [52]. While these approaches have been primarily applied to machine learning [53], it is possible to extend them to brain processes, as has been shown in other applications such as distributional reinforcement learning [54].

### Sensory codes shape individual preferences

The next question we ask is whether our early sensory codes determine or at least partially influence our subjective preferences in terms of what we like. Recent neurocomputational accounts suggest that the perceived appeal of an object is, at least in part, determined by the principles of neural coding efficiency, possibly at the level of dedicated sensing processing structures [55,56].



For instance, a recent study found that social impressions of the attractiveness and trustworthiness of the facial features of a stranger are monotonically driven by a measure of the 'typicality' of the face with respect to the face feature distribution to which the observer has adapted (i.e., does the face look like the average person they have seen?) [56]. Strikingly, manipulation of the likelihood of a face feature being present in each context (e.g., gender type) largely explained the attractiveness ratings of human participants. However, why should the brain have an affect or attractiveness signal related to statistical typicality? The efficient coding hypothesis tells us that to attain and maintain coding efficiency in the brain, inefficiently coded stimuli should be aversive, on average, as dealing with unpredictability requires effort from the viewer. While this study did not include brain imaging data to support these conclusions, previous work has shown that energy expenditure across face-responsive (object-responsive) areas in the brain is negatively related to the face (object) typicality [57,58]. This means that the ensuing rewarding effect of observing an object can be derived from how rapidly the system transitions from a state of uncertainty to a state of increased predictability according to its contextual expectations [55,59].

Our attitudes towards risk are another form of preference. A large body of empirical research indicates that humans are generally risk-averse, and decision-makers tend to choose smaller but certain options over larger risky ones. Classical explanations posit that risk aversion emerges from how the brain assigns subjective values to objective monetary outcomes, which are typically related to downstream processes of reward processing [60,61]. A common observation is that there is considerable variability in the individual strength of this risk-aversiveness. However, a more fundamental observation is that the processing of both symbolic and non-symbolic numerosity values, which typically serve as value inputs during the evaluation of risky prospects, are imprecise and biased [62,63], and are influenced by the environmental statistics [29], attentional [64] and motivational levels [65], and time constraints [62,66]. Based on these observations the question one may ask here is – do early sensory numerosity codes influence our risk-aversion behavior? Recent theories on risky choices propose that risk attitudes may stem from imprecise mental representations of numerical magnitudes [67,68]. These models assume logarithmic noisy coding of numerosity values influenced by the prior beliefs of an individual. Consequently, estimations become more variable and underestimated for larger magnitudes, leading to risk-averse choices. Notably, recent findings link the degree of risk aversion of an individual to how precisely the brain represents numerical magnitude information – a stable trait that influences perceptual inferences and decision-making about risky prospects [69]. This suggests that individual economic behavior patterns might be partly shaped by neural processing limitations – or other coding strategies such as divisive normalization [70] – that are implemented in low-level sensory percepts, rather than by processes that assign subjective values to monetary rewards in downstream neural circuits.

### Early sensory coding of emotion schemas and disentanglement of abstract goal-oriented objectives

We have suggested that early sensory systems incorporate abstract structural information about the environment while considering contextual objective functions and affordances of the system [10,14,51], and that this might be simply because it could be beneficial to efficiently guide future behavior by mitigating information loss. Interestingly, however, recent evidence also shows that environmental feature coding in early sensory processing also occurs in the absence of action-driven learning (Box 2). We propose that such early sensory codes serve to convey schemas or situations that the organism must quickly react to.

To understand what we mean here, let us take the example of emotion processing which is thought to be a canonical response to situations ancestrally linked to survival [71,72]. Neuroscientific views

### Box 2. Environmental feature coding in early sensory processing in the absence of action-driven learning

Recent advances in neural population imaging have led to the discovery that early sensory systems in the cortex can form representations that go beyond physical sensory encoding [92], even in the absence of goal-oriented behavior. A good example comes from a study where the investigators recorded from neurons in the auditory cortex of mice as they implicitly learned sequences of sounds where the predictions of upcoming stimuli created interferences between the short-term memory of recent stimuli and the sensory representation of new stimuli [93]. Analyses of the neural signals revealed that the neural population represented both sensory inputs and the memory of recent stimuli in a contextual manner. Crucially, the investigators found that the auditory cortex mitigates interference by dynamically rotating sensory representations into an orthogonal memory representation. This evidence strongly suggests that, at the earliest sensory processing stages, the brain already has mechanisms for generating efficient representations that protect past experiences (memories) from sensory interference even in the absence of action-learning mechanisms – functions that are otherwise typically ascribed to higher-level processing areas such as the prefrontal cortex [94]. We are not suggesting that the implementation of such efficient coding strategies does not require the interaction between downstream and early sensory systems. Instead, we wish to emphasize that – because the brain is a noisy communication channel that always loses information during transmission – it is beneficial to make use of some of the limited neural and metabolic resources to incorporate abstract structural representations of the environment at the earliest stages of sensory processing.

have historically suggested that emotion is driven by specialized brain regions, often in the so-called limbic system [73,74]. It is also clear that sensory processing plays a prominent role in nearly every treatment of emotion [75]. However, evidence for emotion schemas in neural patterns of early sensory structures was lacking. A recent investigation addressed this, and revealed reliable mapping of category-specific visual features to distinct emotions within the earliest human visual cortical processing systems [76]. However, why do early sensory systems invest resources in encoding emotion schemas? Perhaps sensory cues alone effectively convey situations that demand rapid responses. If category-specific visual features reliably map to emotions in early visual structures, swift appraisals of abstract features could occur rapidly (even subconsciously), therefore aiding prompt decisions in critical situations. Thus, early sensory structures likely offer rapid and precise emotion predictions that are crucial for efficient responses to imminent events, particularly in high-stakes scenarios.

In real-life scenarios, the value of an action is closely linked to the behavioral goal of an agent [77]. However, it is uncertain whether survival-oriented actions prompt rapid reorganization of perceptual data at the earliest sensory processing stages. This would enable organisms to rapidly decipher behavioral goals tied to physical objects before losing critical information in the transmission pipeline of the brain. Conventionally, higher association areas such as the prefrontal cortex are thought to handle this task, in contrast to the view that this occurs in early sensory regions. Recent research explored this by having volunteers undergo fMRI while imagining that they were using objects for survival-oriented goals [78]. As expected, initial analyses revealed object identity representations in visual areas, whereas sensory information for novel goal-directed behavior was found in prefrontal structures. However, closer examination revealed that the usefulness of the objects in the survival scenario was not solely represented in downstream prefrontal regions but also in early visual areas [51,78]. These findings strengthen the idea that efficient representation of sensory information linked to context-specific, goal-driven behavioral rules – aimed at maximizing organism fitness – is present early in cortical sensory processing.

### Efficient sensing of internal body signals?

When we talk or write about sensory perception, we usually have in mind sensing the external world through vision, hearing, olfaction, gustation, or touch. However, there is a 'sixth sense' that is fundamental for survival and guiding optimal behavior: sensing the internal state of the body – known as interoception. This form of sensing is the ability to perceive information from within the body, such as hunger, nutrient levels, thirst, temperature, and pain. However, interoceptive signals and their corresponding sensing systems are less accessible and more difficult

to study relative to (external) environmental sensory signals. Therefore, it remains an open question whether neural coding of our internal world exhibits efficient neural representations, as has been established for the processing of sensory stimuli from the external world.

A key example of the fundamental importance of efficient and adaptive interoceptive processing comes from glucose sensing. Glucose is constantly released into the blood and taken up by tissues, resulting in complex minute-to-minute fluctuations of its concentration in the blood. This internal glucose variability is a fundamental determinant of brain function and disease states [79–82]. Interpretation of glucose dynamics by the brain thus attracted much interest, especially in relation to specialized 'glucose-sensing' neurons clustered in evolutionarily old brain areas such as the hypothalamus [83]. One population of such blood glucose-sensing neurons was recently identified as hypocretin/orexin neurons (HONs). These hypothalamic cells send efferent signals widely throughout the brain and act as a cornerstone of brain and behavioral control across mammals [84–87]. Interestingly, recent data indicate that, in relation to glucose, HONs operate as an early sensory system that performs efficient feature extraction and represents arousal, emotional, and behavioral rules, thus potentially conforming to our efficient coding postulates (Box 3).

### Concluding remarks

We wish to emphasize that, with the evidence and opinions presented here, it is not our intention to convey the radical idea that early sensory structures should completely discard veridical information and instead represent only reward or fitness values. The idea we want to convey is that given that noisy communication channels such as the brain always lose information during transmission, it is more efficient for sensory systems to allocate neural resources that promote fitness maximization at the earliest stages of sensory processing, although we support the idea that the veridicality of our percepts might be compromised to some extent, simply because our resources are finite.

#### Box 3. Early sensory coding of sugar levels

Hypothalamic orexin neurons (HONs) meet the definition of an early sensory system for glucose because small physiological changes in ambient extracellular glucose elicit direct electrical responses in isolated HONs *in vitro* [95]. Initial studies demonstrate that glucose – but not other nutrients – causes electrical inhibition of isolated HONs, originally leading to assumptions that HONs report absolute levels of extracellular glucose [95–97]. Recently, however, this was re-examined through direct measurements of blood glucose dynamics (using arterial glucose sensors) performed simultaneously with monitoring of HON activity (using genetically encoded fluorescent activity indicators) in freely behaving mice. These direct comparisons of concurrently measured HON activation and blood glucose dynamics indicated that HONs do not relay the 'simple reality' of glucose level. Instead of encoding the absolute blood glucose level, the HON population output transmits information about the speed of glucose change, specifically the inverted-sign first temporal derivative of blood glucose [98]. Interestingly, some biophysical features of isolated HONs in earlier studies are consistent with their intrinsic ability to implement such encoding [99]. Multivariate analysis of diverse behavioral and physiological variables confirmed specialized coding of the glucose derivative in the low-frequency activity regime of HONs. This early 'derivative encoding' of blood glucose dynamics by HONs enables efficient physiological and behavioral responses that anticipate peak glucose deviations by several minutes [98].

Importantly, this early coding of specific temporal features of blood glucose in the HON output may already contain arousal, emotional, and behavioral information. Selective optogenetic manipulations of HON indicate that changes in their electrical output are sufficient to alter arousal, anxiety, and movement, and such brain state and behavioral information can also be decoded from recordings of natural HON activity in behaving mice [100–105]. This implies that, in combination with the glucose trend information, HON activity already contains rules for arousal, emotion, and behavior, thus usefully avoiding expending brain resources on creating such rules in 'higher' brain structures. Anatomically, HONs are also well positioned to implement these codes because they directly (monosynaptically) connect to subcortical arousal and emotion regulators, as well as to primary motor systems of the spinal cord. The HONs can rapidly activate these regions by releasing multiple neurotransmitters such as glutamate, dynorphin, and the orexin/hypocretin peptides [106–108]. Overall, HONs can be viewed as a primary sensory system for our internal metabolic state that extracts temporal features about blood glucose into an efficient representation of emotional and behavioral rules at the earliest sensory stage.

### Outstanding questions

How do systems efficiently reallocate their limited coding resources via trial-to-trial experience? What neurocomputational mechanisms guarantee that such dynamic reallocation of neural resources in early sensory systems is efficient? Are these learning mechanisms universally applicable to all sensory modalities?

What are the timescales of fitness maximizing adaptation in early sensory systems at each level of the sensory hierarchy (e.g., in the retina or primary sensory cortices)? Seconds? Minutes? Days? Years? Generations? How, when, and where are such efficient sensory codes implemented?

How efficient is the sensing of our internal body signals? There are many unknowns regarding the efficiency of our interoceptive 'world', and many open questions are ahead.

Given the evidence that some of the often-observed irrationalities (typically attributed to downstream processing) unavoidably originate from the way that our early sensory systems adapt to and process information, how can this help us to guide policymaking? For instance, to improve the diagnoses and characterization of traits of various neuropsychiatric disorders that have been attributed to aberrant sensory processing and learning. In addition, how can this knowledge help to guide applications emerging from artificial intelligence, for example, by counteracting discrimination and systemic biases.

Another aspect deserving consideration is that we have mostly focused on the advantages of utilizing resources to adopt such early processing recoding strategies. However, it must be realized that this also imposes challenges on the computational processes to achieve such early sensory-stage adaptability. Continuously changing the encoding strategies via 'neural retuning' is not limited to only a subset of early sensory neurons. According to our premise, beyond well-controlled laboratory environments, such retuning must occur across sensory modalities, temporal and hierarchical scales, and a variety of contexts, which typically interact through complex covariance structures. How should downstream circuits keep track of all these potentially valid 'retuning' operations such that they know which one is relevant at any given point in time? This problem is known as 'the coding catastrophe' [88] (Box 4).

The set of concepts and theoretical frameworks discussed here may have important clinical and policymaking implications (see also [Outstanding questions](#)). In the clinical domain, recent investigations have developed theories that attempt to provide normative accounts of early sensory coding in complex neuropsychiatric conditions [89]. Unfortunately, some of these theories generally ignore the normative foundation that organisms must optimize behavioral processes because of biological restrictions on information processing. A recent study provided hints at how these resource-limitation concepts can be incorporated, where it was shown that autism spectrum disorder (ASD) can be reliably characterized by inflexible adaptation and allocation of sensory encoding of resources which cannot be attributed solely to learning deficits [90]. This might explain the origins of the aberrant perceptual flexibility that is typically observed in ASD.

Finally, we reported how statistical regularities alongside resource limitations in information processing systems can lead to biased preferences, for instance during appraisals of face features,

#### Box 4. We have a problem: the coding catastrophe

Constantly altering encoding strategies through 'neural retuning' extends beyond a specific subset of early sensory neurons. Our premise suggests that, outside controlled laboratory settings, this adaptation occurs across sensory modalities, various temporal and hierarchical scales, and diverse contexts that interact through complex covariance structures. How can downstream circuits manage these diverse 'retuning' operations to discern their relevance at any moment? This dilemma is known as 'the coding catastrophe'. We discuss ways in which this problem might be tackled by the brain.

We posit that the stability of early sensory neural codes relies on contextual timescales and hierarchical sensory organization. Optimal sensory codes might exhibit more stability under relatively fixed environmental and behavioral conditions over evolutionary timescales, as observed in efficient coding of low-level sensory features in the retina [109,110]. However, faster temporal scales allow rapid efficient filtering of sensory data via attention mechanisms and top-down contextual modulation in early sensory processing [33], which can be achieved, for instance, via mechanisms of reinforcement learning and top-down attentional normalization [111]. Prominent examples supporting this notion come from studies in rodents showing that stimulus-reward contingencies are rapidly implemented in primary cortical sensory structures [112,113]. In one such study, it was shown that these stimulus-reward mapping rules are initially learned by prefrontal structures which continuously engage in 'teaching' early sensory structures about these mapping rules [112]. Crucially, the implementation of these mappings in the early sensory structures is absent when the prefrontal cortex inputs are blocked. Hence, our thesis does not dismiss the development of efficient coding strategies in downstream structures such as prefrontal areas [43]. Instead, it underscores their fundamental role in learning and interpreting abstract environmental rules [112,114].

In dealing with the challenge of extracting crucial information encoded by early sensory systems amid numerous contextual rules, recent investigations are shedding light on how the brain copes with aspects of the coding catastrophe. Thanks to advances in large-scale neural population recordings, there is increasing evidence that our brains have developed strategies to ensure that veridical information is not completely suppressed, for instance by disentangling distinct aspects of the objective environmental information via orthogonalization [115,116]. Intriguingly, some studies indicate that environmental feature coding beyond sensory accuracy occurs using these strategies, even without action-driven learning [93]. In addition, it has recently been shown that early sensory structures adopt coding strategies that reflect a balance between the fraction of neural variance that is devoted to representing coarse and fine stimulus features via such 'orthogonalization' by using a property known as power-law eigenspectral representations [117]. The authors formally established that adhering to this principle enables the brain to generate highly efficient and flexible codes, thereby ensuring robust generalization at the earliest stages of sensory processing.

race, or gender. We argue that a formal study of how limitations in information processing lead to constrained-optimal solutions that have potentially harmful types of biases will have deep implications for policymaking in statistical and taste-based discrimination. A formal study of these constrained information processing phenomena can guide society in counteracting discrimination, for instance by guiding searches in the labor market of both firms and job applicants irrespective of whether the agents making the decisions are humans or machines.

### Acknowledgments

This work was supported by a European Research Council (ERC) starting grant (ENTRAINER) to R.P. This project has received funding from the ERC under the EU Horizon 2020 Research and Innovation Programme (grant 758604). T.A.H. received support from the Swiss National Science Foundation (SNSF; grant 32003B\_166566) and the Hochschulmedizin Zürich Flagship Program STRESS. D.B. thanks ETH Zurich for funding support. The funders had no role in study design, data collection, analysis, decision to publish, or preparation of the manuscript.

### Declaration of interests

The authors declare no conflicts of interest.

### References

- Descartes, R. (1999) *Discourse on Method and Meditations on First Philosophy*, Hackett Publishing
- Kant, I. (1908) *Critique of pure reason. 1781*, Modern Classical Philosophers/Houghton Mifflin
- Martin, J.M. et al. (2021) Useful misrepresentation: perception as embodied proactive inference. *Trends Neurosci.* 44, 619–628
- McKay, R.T. and Dennett, D.C. (2009) The evolution of misbelief. *Behav. Brain Sci.* 32, 493–510
- Searle, J.R. (2015) *Seeing Things as They Are: A Theory of Perception*, Oxford University Press
- Berke, M.D. et al. (2022) Flexible goals require that inflexible perceptual systems produce veridical representations: implications for realism as revealed by evolutionary simulations. *Cogn. Sci.* 46, e13195
- Jackson, F. (1977) *Perception: A Representative Theory*, Cambridge University Press
- Hoffman, D.D. et al. (2015) The interface theory of perception. *Psychon. Bull. Rev.* 22, 1480–1506
- Prakash, C. et al. (2020) Fitness beats truth in the evolution of perception. *Acta Biotheor.* 69, 319–341
- Gibson, J.J. (2014) *The Ecological Approach to Visual Perception*, Psychology Press
- Zadra, J.R. and Clore, G.L. (2011) Emotion and perception: the role of affective information. *Wiley Interdiscip. Rev. Cogn. Sci.* 2, 676–685
- Beck, J.M. et al. (2012) Not noisy, just wrong: the role of suboptimal inference in behavioral variability. *Neuron* 74, 30–39
- Pouget, A. et al. (2013) Probabilistic brains: knowns and unknowns. *Nat. Neurosci.* 16, 1170–1178
- Van Dijk, S.G. and Polani, D. (2012) Informational drives for sensor evolution. In *ALIFE 2012: Thirteenth International Conference on the Synthesis and Simulation of Living Systems*, pp. 333–340, ASME
- Attneave, F. (1954) Some informational aspects of visual perception. *Psychol. Rev.* 61, 183
- Barlow, H.B. (1961) Possible principles underlying the transformation of sensory messages. *Sensory Communication* 1, 217–233
- Robson, A.J. (2001) The biological basis of economic behavior. *J. Econ. Lit.* 39, 11–33
- Netzer, N. (2009) Evolution of time preferences and attitudes toward risk. *Am. Econ. Rev.* 99, 937–955
- Simon, H.A. (1956) Rational choice and the structure of the environment. *Psychol. Rev.* 63, 129–138
- Simon, H.A. (1955) A behavioral model of rational choice. *Q. J. Econ.* 69, 99–118
- Lieder, F. and Griffiths, T.L. (2019) Resource-rational analysis: understanding human cognition as the optimal use of limited computational resources. *Behav. Brain Sci.* 43, e1
- Bhui, R. et al. (2021) Resource-rational decision making. *Curr. Opin. Behav. Sci.* 41, 15–21
- Glimcher, P.W. (2022) Efficiently irrational: deciphering the riddle of human choice. *Trends Cogn. Sci.* 26, 669–687
- Wei, X.-X. and Stocker, A.A. (2015) A Bayesian observer model constrained by efficient coding can explain 'anti-Bayesian' percepts. *Nat. Neurosci.* 18, 1509–1517
- Sims, C.R. (2018) Efficient coding explains the universal law of generalization in human perception. *Science* 360, 652–656
- Wei, X.-X. and Stocker, A.A. (2017) Lawful relation between perceptual bias and discriminability. *Proc. Natl. Acad. Sci. U. S. A.* 114, 10244–10249
- Woodford, M. (2014) Stochastic choice: an optimizing neuroeconomic model. *Am. Econ. Rev.* 104, 495–500
- Woodford, M. (2020) Modeling imprecision in perception, valuation, and choice. *Annu. Rev. Econ.* 12, 579–601
- Heng, J.A. et al. (2020) Efficient sampling and noisy decisions. *Elife* 9, e54962
- Polania, R. et al. (2019) Efficient coding of subjective value. *Nat. Neurosci.* 22, 134–142
- Gershman, S.J. and Bhui, R. (2020) Rationally inattentive intertemporal choice. *Nat. Commun.* 11, 3365
- Prat-Carrabin, A. and Woodford, M. (2022) Efficient coding of numbers explains decision bias and noise. *Nat. Hum. Behav.* 6, 1142–1152
- Grujic, N. et al. (2022) Rational inattention in mice. *Sci. Adv.* 8, 8935
- Attwell, D. and Laughlin, S.B. (2001) An energy budget for signaling in the grey matter of the brain. *J. Cereb. Blood Flow Metab.* 21, 1133–1145
- Laughlin, S.B. (2001) Energy as a constraint on the coding and processing of sensory information. *Curr. Opin. Neurobiol.* 11, 475–480
- Laughlin, S.B. et al. (1998) The metabolic cost of neural information. *Nat. Neurosci.* 1, 36–41
- Padamsey, Z. et al. (2021) Neocortex saves energy by reducing coding precision during food scarcity. *Neuron* 110, 280–296
- Silver, D. et al. (2021) Reward is enough. *Artif. Intell.* 299, 103535
- Tkačik, G. and Bialek, W. (2016) Information processing in living systems. *Annu. Rev. Condens. Matter Phys.* 7, 89–117
- Kim, H. (2022) Evolution as explanation: the origins of neural codes and their efficiencies. *Arxiv*, Published online 18 March, 2022. <https://arxiv.org/abs/2203.09719v1>
- Page, L. (2022) *Optimally Irrational: The Good Reasons We Behave the Way We Do*, Cambridge University Press
- Kuo, M.F. et al. (2016) Physiology of transcranial direct and alternating current stimulation. In *Transcranial Direct Current Stimulation in Neuropsychiatric Disorders: Clinical Principles and Management* (Brunoni, A.R. et al., eds), pp. 29–46, Springer



43. Rustichini, A. *et al.* (2017) Optimal coding and neuronal adaptation in economic decisions. *Nat. Commun.* 8, 1208
44. Bhui, R. and Gershman, S.J. (2018) Decision by sampling implements efficient coding of psychoeconomic functions. *Psychol. Rev.* 125, 985–1001
45. Sims, C.R. (2016) Rate-distortion theory and human perception. *Cognition* 152, 181–198
46. Juechems, K. *et al.* (2021) Optimal utility and probability functions for agents with finite computational precision. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2002232118
47. Zaslavsky, N. *et al.* (2018) Efficient compression in color naming and its evolution. *Proc. Natl. Acad. Sci. U. S. A.* 115, 7937–7942
48. Sims, C.A. (2003) Implications of rational inattention. *J. Monet. Econ.* 50, 665–690
49. Padoa-Schioppa, C. and Rustichini, A. (2014) Rational attention and adaptive coding: a puzzle and a solution. *Annu. Rev. Econ.* 104, 507–513
50. Brus, J. *et al.* (2021) Sources of confidence in value-based choice. *Nat. Commun.* 12, 7337
51. Schaffner, J. *et al.* (2023) Sensory perception relies on fitness-maximizing codes. *Nat. Hum. Behav.* 7, 1135–1151
52. Peng, Z. *et al.* (2017) An information-theoretic on-line update principle for perception-action coupling. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 789–796, IEEE
53. Leibfried, F. *et al.* (2017) An information-theoretic optimality principle for deep reinforcement learning. *Arxiv*, Published online August 6, 2017. <https://arxiv.org/abs/1708.01867v5>
54. Dabney, W. *et al.* (2020) A distributional code for value in dopamine-based reinforcement learning. *Nature* 577, 671–675
55. Van de Cruys, S. and Wagemans, J. (2011) Putting reward in art: a tentative prediction error account of visual art. *Perception* 2, 1035–1062
56. Ryali, C.K. *et al.* (2020) From likely to likable: the role of statistical typicality in human social assessment of faces. *Proc. Natl. Acad. Sci. U. S. A.* 117, 29371–29380
57. Mattavelli, G. *et al.* (2012) Response of face-selective brain regions to trustworthiness and gender of faces. *Neuropsychologia* 50, 2205–2211
58. Said, C.P. *et al.* (2010) The amygdala and FFA track both social and non-social face dimensions. *Neuropsychologia* 48, 3596–3605
59. Brielmann, A.A. and Dayan, P. (2022) A computational model of aesthetic value. *Psychol. Rev.* 129, 1319–1337
60. Hsu, M. *et al.* (2009) Neural response to reward anticipation under risk is nonlinear in probabilities. *J. Neurosci.* 29, 2231–2237
61. Zalocusky, K.A. *et al.* (2016) Nucleus accumbens D2R cells signal prior outcomes and control risky decision-making. *Nature* 531, 642–646
62. Cheyette, S.J. and Piantadosi, S.T. (2020) A unified account of numerosity perception. *Nat. Hum. Behav.* 4, 1265–1272
63. Izard, V. and Dehaene, S. (2008) Calibrating the mental number line. *Cognition* 106, 1221–1247
64. Anobile, G. *et al.* (2012) Linear mapping of numbers onto space requires attention. *Cognition* 122, 454–459
65. Gershman, S.J. and Burke, T. (2022) Mental control of uncertainty. *Cogn. Affect. Behav. Neurosci.* 23, 465–475
66. Heng, J.A. *et al.* (2023) Efficient numerosity estimation under limited time. *BioRxiv*, Published online October 12, 2023. <https://doi.org/10.1101/2023.07.03.547493>
67. Khaw, M.W. *et al.* (2020) Cognitive imprecision and small-stakes risk aversion. *Rev. Econ. Stud.* 88, 1979–2013
68. Frydman, C. *et al.* (2021) Efficient coding and risky choice. *Q. J. Econ.* 137, 161–213
69. Barretto-Garcia, M. *et al.* (2023) Individual risk attitudes arise from noise in neurocognitive magnitude representations. *Nat. Hum. Behav.* 7, 1551–1567
70. Landry, P. and Webb, R. (2021) Pairwise normalization: a neuroeconomic theory of multi-attribute choice. *J. Econ. Theory* 193, 105221
71. LeDoux, J.E. (2012) Evolution of human emotion: a view through fear. *Prog. Brain Res.* 195, 431–442
72. Evans, D., Cruse, P., eds (2004) *Emotion, Evolution, and Rationality*, Oxford University Press
73. Lindquist, K.A. *et al.* (2012) The brain basis of emotion: a meta-analytic review. *Behav. Brain Sci.* 35, 121–143
74. Dalglish, T. (2004) The emotional brain. *Nat. Rev. Neurosci.* 57, 583–589
75. Mauss, I.B. and Robinson, M.D. (2009) Measures of emotion: a review. *Cogn. Emot.* 23, 209–237
76. Kragel, P.A. *et al.* (2019) Emotion schemas are embedded in the human visual system. *Sci. Adv.* 5, eaaw4358
77. De Martino, B. and Cortese, A. (2023) Goals, usefulness and abstraction in value-based choice. *Trends Cogn. Sci.* 27, 65–80
78. Castagnetti, G. *et al.* (2021) How usefulness shapes neural representations during goal-directed behavior. *Sci. Adv.* 7, eabd5363
79. Zhou, Z. *et al.* (2020) Glycemic variability: adverse clinical outcomes and how to improve it? *Cardiovasc. Diabetol.* 19, 102
80. Gailliot, M.T. and Baumeister, R.F. (2018) The physiology of willpower: linking blood glucose to self-control. *Self-regulation and Self-control* 129–172 Routledge
81. Messier, C. and Gagnon, M. (1996) Glucose regulation and cognitive functions: relation to Alzheimer's disease and diabetes. *Behav. Brain Res.* 75, 1–11
82. Gold, P.E. (1995) Role of glucose in regulating the brain and cognition. *Am. J. Clin. Nutr.* 61, 987–995
83. Burdakov, D. *et al.* (2005) Glucose-sensing neurons of the hypothalamus. *Philos. Trans. R. Soc. B Biol. Sci.* 360, 2227–2235
84. Sakurai, T. *et al.* (1998) Orexins and orexin receptors: a family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior. *Cell* 92, 573–585
85. De Lecea, L. *et al.* (1998) The hypocretins: hypothalamus-specific peptides with neuroexcitatory activity. *Proc. Natl. Acad. Sci. U. S. A.* 95, 322–327
86. Chemelli, R.M. *et al.* (1999) Narcolepsy in orexin knockout mice: molecular genetics of sleep regulation. *Cell* 98, 437–451
87. Burdakov, D. (2019) Reactive and predictive homeostasis: roles of orexin/hypocretin neurons. *Neuropharmacology* 154, 61–67
88. Schwartz, O. *et al.* (2007) Space and time in visual context. *Nat. Rev. Neurosci.* 8, 522–535
89. Stephan, K.E. and Mathys, C. (2014) Computational approaches to psychiatry. *Curr. Opin. Neurobiol.* 25, 85–92
90. Noei, J.P. *et al.* (2021) Individuals with autism spectrum disorder have altered visual encoding capacity. *PLoS Biol.* 19, e3001215
91. Shannon, C.E. (1948) A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423
92. Katzner, S. *et al.* (2019) V1 microcircuits underlying mouse visual behavior. *Curr. Opin. Neurobiol.* 58, 191–198
93. Libby, A. and Buschman, T.J. (2021) Rotational dynamics reduce interference between sensory and memory representations. *Nat. Neurosci.* 24, 715–726
94. Panichello, M.F. and Buschman, T.J. (2021) Shared mechanisms underlie the control of working memory and attention. *Nature* 592, 601–605
95. Burdakov, D. *et al.* (2005) Physiological changes in glucose differentially modulate the excitability of hypothalamic melanin-concentrating hormone and orexin neurons in situ. *J. Neurosci.* 25, 2429–2433
96. González, J.A. *et al.* (2008) Metabolism-independent sugar sensing in central orexin neurons. *Diabetes* 57, 2569–2576
97. Kamani, M.M. *et al.* (2011) Activation of central orexin/hypocretin neurons by dietary amino acids. *Neuron* 72, 616–629
98. Viskaitis, P. *et al.* (2022) Orexin cells efficiently decode blood glucose dynamics to drive adaptive behavior. *BioRxiv*, Published online November 9, 2022. <https://doi.org/10.1101/2022.04.14.488310>
99. Williams, R.H. *et al.* (2008) Adaptive sugar sensors in hypothalamic feeding circuits. *Proc. Natl. Acad. Sci. U. S. A.* 105, 11975–11980
100. Adamantidis, A.R. *et al.* (2007) Neural substrates of awakening probed with optogenetic control of hypocretin neurons. *Nature* 450, 420–424
101. Garau, C. *et al.* (2020) Orexin neurons and inhibitory Agrp→orexin circuits guide spatial exploration in mice. *J. Physiol.* 598, 4371–4383
102. Kamani, M.M. *et al.* (2020) Role of spontaneous and sensory orexin network dynamics in rapid locomotion initiation. *Prog. Neurobiol.* 187, 101771
103. Li, H.T. *et al.* (2022) Hypothalamic deep brain stimulation as a strategy to manage anxiety disorders. *Proc. Natl. Acad. Sci. U. S. A.* 119, 2113518119

104. Giardino, W.J. *et al.* (2018) Parallel circuits from the bed nuclei of stria terminalis to the lateral hypothalamus drive opposing emotional states. *Nat. Neurosci.* 21, 1084–1095
105. Grujic, N. *et al.* (2023) Control and coding of pupil size by hypothalamic orexin neurons. *Nat. Neurosci.* 26, 1160–1164
106. Kosse, C. *et al.* (2017) Orexin-driven GAD65 network of the lateral hypothalamus sets physical activity in mice. *Proc. Natl. Acad. Sci. U. S. A.* 114, 4525–4530
107. Blomeley, C. *et al.* (2018) Accumbal D2 cells orchestrate innate risk-avoidance according to orexin signals. *Nat. Neurosci.* 21, 29–34
108. Schöne, C. *et al.* (2014) Coreleased orexin and glutamate evoke nonredundant spike outputs and computations in histamine neurons. *Cell Rep.* 7, 697–704
109. Laughlin, S. (1981) A simple coding procedure enhances a neuron's information capacity. *Z. Naturforsch. C Biosci.* 36, 910–912
110. Laughlin, S.B. *et al.* (1987) Synaptic limitations to contrast coding in the retina of the blowfly *Calliphora*. *Proc. R. Soc. Lond. B Biol. Sci.* 231, 437–467
111. Reynolds, J.H. and Heeger, D.J. (2009) The normalization model of attention. *Neuron* 61, 168–185
112. Banerjee, A. *et al.* (2020) Value-guided remapping of sensory cortex by lateral orbitofrontal cortex. *Nature* 585, 245–250
113. Poort, J. *et al.* (2015) Learning enhances sensory and multiple non-sensory representations in primary visual cortex. *Neuron* 86, 1478–1490
114. Zhou, J. *et al.* (2020) Evolving schema representations in orbitofrontal ensembles during learning. *Nature* 590, 606–611
115. Avitan, L. and Stringer, C. (2022) Not so spontaneous: multi-dimensional representations of behaviors and context in sensory areas. *Neuron* 110, 3064–3075
116. Mante, V. *et al.* (2013) Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* 503, 78–84
117. Stringer, C. *et al.* (2019) High-dimensional geometry of population responses in visual cortex. *Nature* 571, 361–365