Please cite this article in press as: Kay et al., Tasks and their role in visual neuroscience, Neuron (2023), https://doi.org/10.1016/j.neuron.2023.03.022

Neuron

CellPress

Review

Tasks and their role in visual neuroscience

Kendrick Kay,[1,*] Kathryn Bonnen,[2] Rachel N. Denison,[3] Mike J. Arcaro,[4] and David L. Barack[5]
[1]Center for Magnetic Resonance Research, Department of Radiology, University of Minnesota, Minneapolis, MN 55455, USA
[2]School of Optometry, Indiana University, Bloomington, IN 47405, USA
[3]Department of Psychological and Brain Sciences, Boston University, Boston, MA 02215, USA
[4]Department of Psychology, University of Pennsylvania, Philadelphia, PA 19146, USA
[5]Departments of Neuroscience and Philosophy, University of Pennsylvania, Philadelphia, PA 19146, USA
*Correspondence: kay@umn.edu
https://doi.org/10.1016/j.neuron.2023.03.022

SUMMARY

Vision is widely used as a model system to gain insights into how sensory inputs are processed and interpreted by the brain. Historically, careful quantification and control of visual stimuli have served as the backbone of visual neuroscience. There has been less emphasis, however, on how an observer's task influences the processing of sensory inputs. Motivated by diverse observations of task-dependent activity in the visual system, we propose a framework for thinking about tasks, their role in sensory processing, and how we might formally incorporate tasks into our models of vision.

INTRODUCTION

Visual neuroscientists seek to understand how patterns of light are received, processed, and interpreted by the brain. Historically, visual neuroscience has focused intensely on *stimuli*, exerting considerable effort to calibrate stimulus presentation, carefully design and quantify visual stimuli, and to pinpoint exactly how stimulus manipulations affect neural activity. However, in nearly all experiments involving an awake observer, there is an additional component, namely, a *task* that is explicitly instructed, learned via training, or automatically performed by the observer (Figure 1). Often, there is an implicit assumption that the task is ancillary to the main endeavor of characterizing neural stimulus selectivity: all that matters is that the carefully controlled stimulus is accurately delivered to the observer's retina. Indeed, task instructions in physiological studies often amount to "fixate the small dot at the center of the screen."

This stimulus-focused mindset can be readily found in classic treatments of visual neurophysiology, where emphasis is placed on characterizing stimulus properties, the association of specific stimulus properties with specific visual areas, and the anatomical and functional organization of the visual pathways.[1–4] From these treatments, one might believe that the primary aim of visual neuroscience is to determine the ways in which neurons encode, or represent, features of the visual environment. Interestingly, these characterizations of the visual system typically do not even mention the task performed by the observer while physiological measurements are made (but see an interesting exception[5]). That these characterizations emerged is sensible, given that many early neurophysiological experiments were conducted under anesthesia.

The portrait painted by classic visual neurophysiology is that the visual system is a sophisticated feature extractor, akin to a camera that is coupled with advanced image processing. While this stimulus-focused mindset has led to important insights, there are two fundamental limitations. First, it is well established that neural activity in the visual system is not invariant to the task. That is, stimulus-evoked responses—and neural activity more generally—depend on the task performed by the observer. The stimulus is therefore not the only relevant variable. Given that an important goal in visual neuroscience is to predict and explain neural activity,[6,7] the task must be taken into account to construct complete and accurate models. Second, measuring behavioral outcomes while engaging observers in meaningful tasks on visual inputs is of central importance in visual psychophysics. Although psychophysics and neurophysiology have been fruitfully connected for a few types of tasks (e.g., attentional tasks[8]), in other domains there is a potential disconnect between psychophysics and physiological studies that do not employ the same tasks. These complications motivate our main question: how can we incorporate tasks into our working models of the visual system?

In this article, we propose a framework for thinking about tasks and their relationships to sensory processing—a topic that has been discussed[9–12] but has not yet been the focus of a review. We adopt a broad perspective, integrating insights across diverse domains of inquiry including psychophysics, neurophysiology, neuroimaging, cognitive psychology, cognitive neuroscience, computational neuroscience, and philosophy. Our proposal can be viewed therefore as a synthesis of diverse views across different fields. Initially, our exposition is general—the concept of tasks applies to a broad range of sensory, cognitive, and motor paradigms—but we then narrow our scope to vision-related tasks. We expect that our framework is applicable equally to studies of humans, non-human primates, and other animals, with the caveats that running experiments expressly focused on task manipulation might be less practical in animals and that certain complex tasks may be less feasible to study in animals. Although there is growing recognition that tasks (and behavior more generally) should play a larger role in visual neuroscience,[10,13–15] a focused

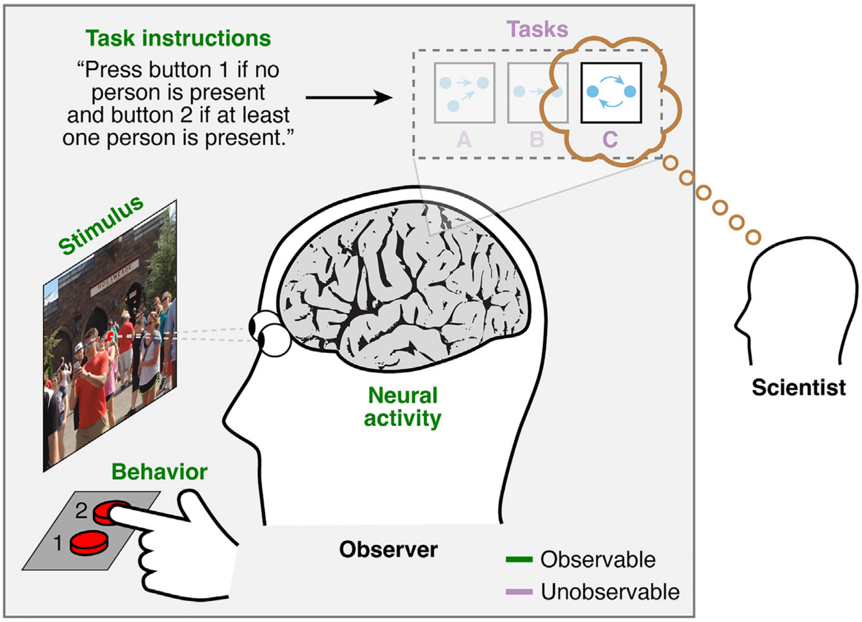

Figure 1. Tasks are an integral component of experiments
In a typical experiment (gray box), the *observer* receives a *stimulus* and a set of *task instructions*. Task instructions are interpreted by the observer and lead to a *task*, defined as the goal of the observer and information-processing operations deployed to achieve that goal (blue dots and arrows). This task is instantiated in *neural activity* of the observer. As a result of the stimulus and task, *behavior* is usually elicited. Critically, while the stimulus, task instructions, neural activity, and behavior are observable, the task is not directly observable and must be inferred by the *scientist* (gold thought cloud). The scientist may favor one hypothesized task (C) over others (A and B), and this could be supported by evidence. Note that some experiments may not involve task instructions, but tasks are nonetheless present, either due to training or due to intrinsic motivation of the observer.

treatment of tasks (what they are, how they relate to neural measurements, and so on) is critical for establishing a foundation for how we might formally incorporate tasks into our models of vision. The overarching goal of this article is to sharpen our thoughts about tasks, to clarify what is at stake when we deploy tasks, and to make the case that the study of tasks is essential for moving the field of visual neuroscience forward.

WHAT ARE TASKS?

Precision of terminology is critical for avoiding confusions that might arise when the same word is used to refer to different things (e.g., "attention,"[16,17] "hierarchy,"[18] and "causality"[19]). The term "task" is used widely in brain and behavioral research with many different meanings, depending on the context. Thus, we start by providing working definitions of a set of basic terms (Box 1). Under our definitions, "goal" refers to the implicit or explicit aim of an observer in a given context, "strategy" refers to the information-processing operations being used to achieve that aim, and task refers to both the goal and strategy of the observer.

It is important to distinguish our definition of task from other possible definitions of the term:

(1) Tasks are not *experimental paradigms*, even though it is common, especially in cognitive neuroscience, to encounter phrasing like "we ran the movie-watching task." An experimental paradigm may help induce a certain task in the observer, but it is not identical to the task. Using our terminology, we might say, "We ran the experiment where the task instructions given to the participant are to watch the movie." The specific task performed by the participant during the experiment (e.g., the participant might seek to look at colorful objects) is not necessarily known to the experimenter.

(2) Tasks are distinct from *states*, defined as sets of external or internal properties that persist over time and partly determine the context in which tasks are performed. For example, the ability of a security agent to screen luggage X-rays at the airport might depend on their internal state—before lunch, they are alert, albeit hungry, whereas after lunch, they are tired, albeit satiated. These state changes may impact neural activity and task performance, but they do not change the task.

(3) Tasks should not be confused with the *functions* of neurons or brain regions. For example, it is sometimes said that the visual system has certain tasks or that perception involves certain tasks,[1] but these usages refer to various functions (or processes) that may be occurring in the brain.

(4) The effects of tasks are related to, but are not the same as, the effects of *feedback* in the visual system. While feedback pathways are likely to be deeply involved in the implementation of tasks in the brain,[10] feedback is not necessarily a signature of task-related computations. For example, consider extra-classical receptive field effects that are thought to be mediated by lateral and feedback connections.[22] Such effects can be fruitfully studied in an experiment where the stimuli are irrelevant to the observer's task (or even during anesthesia).

It is crucial to distinguish the task intended by the experimenter from the task actually performed by the observer (see Figure 1). The former, which we call the *experimenter-defined task*, is conveyed through the design of the experimental paradigm and reflects the point of view of the experimenter. The latter, which we call the *observer's task*, refers to the goal and strategies that the observer actually adopts and reflects the point of view of the observer. For instance, even if the optimal policy in a given experiment is to choose randomly, an observer might form superstitious interpretations of the experiment and carry out complex strategies.[23] The distinction between experimenter-defined task and

Box 1. Working definitions of basic terms

Stimulus: an aspect of the physical environment (usually outside but possibly inside the body) that impinges on the sensors of an organism.
Observer: a biological or artificial entity that processes stimuli from the environment. The observer might exhibit behavior that depends on properties of the stimulus.
Motor act: an output of an observer's motor system that impinges on the environment.
Behavior: an event for which there is a meaningful interpretation with respect to the organism's implicit or explicit goals. The event can have external consequences, such as a motor act, or only internal consequences, such as a change in intention or belief. In different situations, the same motor act (e.g., raising your hand) might be associated with very different behaviors (e.g., you are waving hello or you are waving for help[20]).
Information-processing operation: any transformation of information performed by the observer. This includes sensory transformations, cognitive transformations, and motor transformations.
Context: states of the environment, motor acts that the observer can perform, and causal relationships between motor acts and environmental states.
Goal: the implicit or explicit aim of an observer in a given context. Often, the goal is to achieve a certain behavioral outcome. Observers may neither be aware of their goals nor understand the strategies they are using to achieve those goals. Note that by "goal," we do not mean a normative long-term goal that one might interpret an organism as possessing (e.g., evolutionary goals). Rather, we mean a goal, typically short term, that the organism has in the current context.
Strategy: the information-processing operations that an observer uses to achieve a given goal.
Task: the goal and strategies used by an observer in a given context. Tasks can be hierarchical and consist of multiple subtasks (e.g., the task of grabbing an apple consists of identifying the apple, orienting the hand to grab the apple, initiating the movement, and so on). Tasks need not involve desire (internal motivation), volition (they may be relatively automatic), or an externally observable motor act. Exactly how to individuate tasks, as well as determining when one task has switched to another, is a challenging theoretical question.[21]
Task instructions: the information provided by an experimenter regarding the task the experimenter wants the observer to perform. This information is often communicated in spoken or written form to human observers and in non-verbal form (e.g., task cues) to animal observers. Sometimes, task instructions specify not only the desired goal but also the specific strategy (information-processing operations) to be used.

observer's task is reminiscent of the broader point that "what is meaningful from the point of view of the organism need not be from the point of view of the scientist studying it, and vice versa."[24] Misalignment of tasks can happen if observers engage in other tasks besides the experimenter's task (e.g., thinking about the purpose of the experiment, daydreaming, planning the rest of one's day, etc.). Alternatively, task instructions might be ambiguous or interpreted differently by different individuals. In fact, under our definitions, two observers who execute identical motor acts but use different information-processing operations (strategies) to accomplish those acts are performing different tasks. Because of these various considerations, it is important to remember that the descriptions scientists use for the tasks they incorporate into their experiments are *fallible:* they embody hypotheses regarding the operations executed by an observer, and these hypotheses may neither be correct nor complete and therefore *must be tested*.

HOW DO TASKS AFFECT ACTIVITY IN THE VISUAL SYSTEM?

Role of tasks in visual neuroscience

Suppose that neural activity in the visual system was solely determined by the stimulus (i.e., the patterns of light impinging on the retina). Then, the visual neuroscientist would have the clearly defined job of measuring and characterizing the stimulus-response mappings implemented by neurons in the various visual areas. These mappings are often described in terms of tuning curves along one or more stimulus dimensions, where a tuning curve might consist of a stimulus value that maximally drives the neural response, a bandwidth over which responses maintain their strength, a baseline response level, and a gain parameter characterizing the maximum response level. But if the task of the observer affects neural activity, two qualitatively distinct factors—namely, the stimulus properties present (stimulus) and the observer's goal and strategies for achieving that goal (task)—must be jointly considered when interpreting neural activity. This complicates the situation because the neuroscientist must then grapple with a number of questions. Do observed tuning curves depend on the task? Might some tasks actually change tuning curves, through an increase in gain or bandwidth, or perhaps even a tuning shift? Is it possible that a certain tuning curve arises *only* when the observer performs a particular task?

A common strategy for dealing with the possibility of task influences is to engage the observer in a highly demanding fixation task in which the stimuli chosen for the experiment are irrelevant (e.g., a task involving rapid serial judgments on a small fixation cue). The motivation for this approach is to attempt to achieve neural responses that are free of task influences and hence interpretable as reflecting pure sensory processing. We term such tasks "control tasks" (see Box 2), as they are intended to control the behavioral state of the observer and reduce the possibility of its confounding influence. One way to view these tasks is that

Box 2. Types of experimenter-defined tasks

Note that the categories proposed below are not mutually exclusive (e.g., naturalistic tasks are often continuous tasks). The goal of this survey is to help the reader navigate tasks used in the literature, not necessarily to develop a formal taxonomy of tasks.

Perception-focused tasks: observers are instructed or trained to report on aspects of their perception of a viewed stimulus. This can pertain to either low- (e.g., orientation and color), mid- (e.g., contour, texture, and slant), or high-level (e.g., category and animacy) stimulus properties. The theory is that the task generates behavioral data that enable the scientist to make inferences about the brain's sensory representations.

Cognition-focused tasks: studies in psychology and neuroscience often use tasks that operationalize and target a specific cognitive phenomenon such as attention, memory, or decision-making.[27,28] Experiments often involve simple controlled stimuli whose properties are manipulated to produce variation in behavior. Examples include varying the duration of a presented scene and measuring how well an observer can assess properties of the scene[29] and varying a spatial attention cue and measuring contrast sensitivity at different visual field locations.[30]

Classic psychophysical tasks: early work in visual psychophysics led to the development of detection and discrimination tasks.[31] In the widely used two-alternative forced-choice task paradigm, the observer is forced to choose between two options, even in very difficult (or impossible) situations. Typically, the goal is to discover properties of internal stimulus representations, guided by signal detection theory.

Control tasks: studies in visual neuroscience often attempt to control behavioral state by encouraging the observer to process visual stimuli using a seemingly generic task, such as an oddball detection task[32] or one-back task.[33] Additionally, studies sometimes use tasks that de-emphasize the visual stimuli chosen by the experimenter. These range from highly demanding tasks such as performing rapid judgments on a central visual cue[34] to less demanding tasks such as maintaining central fixation.[35]

"Resting" tasks: resting-state experiments are widely conducted in the field of neuroimaging. These can be construed as involving a task even though it is common to distinguish between resting-state fMRI and task-based fMRI. This is because explicit goals are conveyed to the observer (e.g., "stay still, stay awake, stare at the cross, but otherwise rest"). Furthermore, despite task instructions, the observer may nonetheless engage in a number of uninstructed tasks,[36,37] which can lead to complications if the experimenter assumes that nothing is occurring during rest.

Continuous tasks: many experiments use rigid trial structures where brief, discrete events are conducted with breaks between successive events. Recently, some researchers have explored continuous psychophysics tasks in which stimuli are continuously presented and behavioral measures are continuously obtained, providing efficiency gains and mirroring what occurs in naturalistic contexts.[38,39] Examples include tracking a moving luminance-defined target[38] and tracking the focus of expansion in a field of dots.[40]

Naturalistic tasks: tasks that biological organisms have in the real world,[41] as opposed to ones that are typically studied in the laboratory. Examples include navigating, foraging, fighting, and reproducing.[24,42,43] Naturalistic tasks involve rich dynamic stimuli and are complex (the visual system is just one of many systems involved) and continuous (behavior is continually relevant over time).

Natural behavior tasks: with the advent of technology enabling wireless electrophysiological recordings, recent studies have conducted experiments in which organisms behave freely while brain and/or behavioral data are acquired.[44–46] The key difference with respect to naturalistic tasks is the lack of task instructions or task training. Nonetheless, organisms are still engaged in implicit tasks driven by intrinsic motivations.

they divert cognitive processing as much as possible, as if to render neurons in more central parts of the visual system (e.g., extrastriate visual cortex) to behave like neurons in the peripheral visual system (i.e., retina) where task influences are likely absent or weak (though arousal may influence retinal activity[25]). Insofar as this diversion is successful, neural activity is viewed as building sensory representations, that is, performing information-processing operations that transform the stimulus into more useful formats.[26]

While control tasks are extremely valuable for studying sensory representation, they have clear limitations as the sole approach to understanding vision. Presumably, a major goal of vision science is to understand how an observer performs tasks involving visual inputs. Hence, measurements of neural activity while an observer is engaged in such tasks are of obvious importance. Viewing tasks as part and parcel of the characterization of the function of the visual system is perhaps more natural and commonplace in research programs that are more cognitive in nature. For example, consider the use of vision for spatial navigation.[47] Imagine that a researcher wants to understand how observers perform spatial navigation but measures neural activity only during control tasks and not while an observer actually engages in a spatial navigation task! While some types of visual information-processing operations might be relatively insensitive to the observer's task (e.g., construction of orientation representations in primary visual cortex), below, we draw attention to situations where this is not the case.

Effect of tasks on stimulus-evoked responses

At some level, the influence of tasks on activity in the visual system has long been known. Foundational studies in visual electrophysiology have demonstrated effects of spatial attention[48,49] and feature-based attention[50,51] on the firing rates of single neurons in visual cortex. (Here, we define attention as the selection

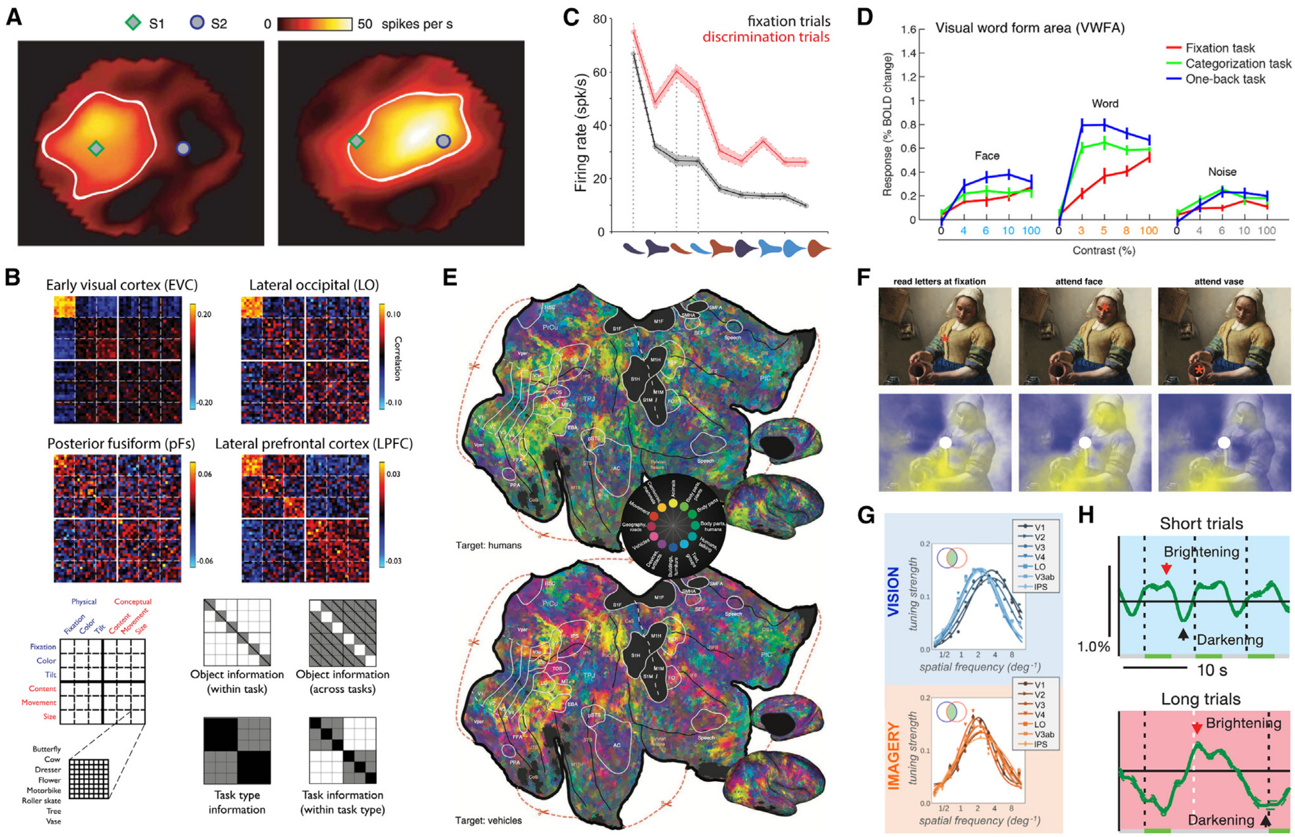

Figure 2. Examples of the effects of tasks on activity in the visual system

(A) Responses of a single neuron in macaque MT (middle temporal area) to stimulus probes presented at different visual field locations. The task was to maintain central fixation while detecting a brief change in a target positioned either at S1 (left) or S2 (right). Reproduced from Womelsdorf et al.[52] with permission from Springer Nature.

(B) Split-half correlation of fMRI activity patterns measured while observers performed different tasks on a common set of objects. While above-chance object decoding (both within and across tasks) was found in EVC, LO, and pFs, task context affected results in all regions. Reproduced from Harel et al.[58] under PNAS license to publish.

(C) Responses of a single neuron in macaque V4 to colored shapes during a fixation task (gaze at a central dot) and a shape discrimination task (judge whether stimulus shape is same or different compared with a previously presented reference stimulus). Rank-order selectivity differed across tasks. Reproduced from Popovkina and Pasupathy[35] under CC-BY license.

(D) Human fMRI responses in visual word form area to faces, words, and noise at different contrast levels. Observers performed either a fixation task (judge color of a rapidly changing central dot), categorization task (judge category of presented images), or one-back task (judge whether current image is same as previous image). Reproduced from Kay and Yeatman[59] under CC-BY license.

(E) Estimates of voxel-wise semantic tuning obtained while the observer covertly searched for humans (top) or vehicles (bottom) in natural movies during central fixation. Reproduced from Çukur et al.[60] with permission from Springer Nature.

(F) Visualization of fMRI activity in V1 and V2 while the observer either reads rapidly presented letters at fixation (left) or maintains central fixation while attending to the face (middle) or attending to the vase (right). Colored images (bottom) reflect summation of voxel-wise receptive field estimates weighted by measured activity (yellow, high; blue, low). Reproduced from Zipser[61] under CC-BY license.

(G) Spatial frequency tuning curves estimated from fMRI responses to viewing of actual stimuli (top) and fMRI responses to internally generated visual imagery (bottom). Imagery-induced responses exhibit stronger tuning to lower spatial frequencies. Reproduced from Breedlove et al.[62] with permission from Elsevier.

(H) Optical imaging of blood volume in macaque V1 during a task in which the observer alternates between looking at a tiny fixation point (green periods) and rest (gray periods). Signal darkening, indicating increase in blood volume, anticipates fixation onsets (black dotted lines) and generalizes across different trial timings. Reproduced from Sirotin and Das[63] with permission from Springer Nature.

or prioritization of certain aspects of sensory inputs, such as visual field location [spatial attention] or other non-spatial attributes [feature-based attention].) For example, using a task that directs spatial attention can produce shifts in the responses of a macaque MT neuron to different visual field locations[52] (Figure 2A). Human neuroimaging has, in turn, demonstrated the impact of attention on blood-oxygen-level-dependent (BOLD) activity in human visual cortex.[53–56] While there is variability in the specific ways attention has been reported to impact tuning curves (ranging from gain effects, additive effects, shifts of contrast-response functions, etc.), when attention is allocated to a portion of visual space or a visual feature that matches the tuning of a given neuron, the activity of that neuron is typically enhanced. Attention, however, is just one of many possible information-processing operations involved in the execution of a task, and other operations might also impact activity in the visual system. For example, a previous study used images of faces to map population receptive fields in face-selective regions of ventral temporal cortex while observers performed three different tasks.[57] Large changes in the position, size, and gain of population receptive fields were observed when the observer performed a face-specific task on the mapping stimulus (one-back task on face identity), compared with a fixation task and also compared with a task in which the observer detected a small dot superimposed on the mapping stimulus. Although the face-specific task and the dot task both involved spatial attention, other operations in the face-specific task (e.g., face recognition, working memory, etc.) were presumably responsible for the distinct outcomes observed under that task.

We highlight several other examples of task-related effects on stimulus-evoked responses beyond attentional effects. We focus specifically on cases where *different* tasks were performed on the *same visual stimulus*, as this helps pinpoint situations where activity was contingent on the goal of the observer and not a consequence of the stimulus itself.[64–66] In one study,[58] fMRI responses were measured to different objects while observers engaged in different perceptual and conceptual tasks. These authors found that task context (i.e., which task was performed by the observer on a given trial) was best decodable from activity in frontal and parietal cortex, but it was also decodable to some degree from visual regions in occipital cortex (Figure 2B). This demonstrates that visual activity is subject to task modulation (see related studies[67–70]). Other studies have provided a closer look at the tuning changes induced by tasks. A recent macaque electrophysiology study[35] demonstrated that the observer's task can alter selectivity in V4 neurons in complex ways that cannot be characterized as a simple scaling or additive effect (Figure 2C). Similarly, a human fMRI study of category-selective visual cortex[59] found task-related modulations of contrast-response functions that do not clearly conform to classic attentional concepts of response gain, contrast gain, or additive shifts (Figure 2D). The observed response modulations were proposed to reflect not attention per se but specific aspects of the perceptual decision-making processes in which the observers were engaged.

Additional examples highlight paradigms that venture away from vision narrowly conceived and move into more cognitive territory. One study[60] monitored whole-brain fMRI activity while human observers maintained central fixation and watched the same naturalistic movie while performing different tasks. Evoked activity while observers covertly searched for humans was strikingly different from evoked activity while observers covertly searched for vehicles, as indexed by large shifts in tuning for various semantic features present in the movie (especially in high-level visual cortex and prefrontal cortex) (Figure 2E). Another study[71] investigated word recognition, that is, the process of converting visually presented letter strings into meaningful concepts. By comparing a lexical decision task (observers judge whether a presented letter string is a real word) to a closely matched control task (observers judge the position of a gap in a string of shapes), these authors provided evidence of highly specific enhancements in word-selective regions of ventral temporal cortex, suggesting that cognitive operations specific to language have privileged top-down control over these word-selective regions. Finally, we draw attention to the point that even seemingly trivial task-related modulations may in fact have significant consequences. Suppose, for example, that a neuron carrying information relevant to the observer's task exhibits an overall gain enhancement (e.g., an increase in tuning curve gain), compared with the situation where the information carried by the neuron is irrelevant to the task. One might conclude that this task-related modulation is not particularly significant, since the relative selectivity exhibited by the neuron is unchanged. However, if one considers the pattern of activity *across* neurons in a given brain region, it becomes clear that information content fundamentally changes. As an intuitive illustration, consider that when spatial attention enhances activity associated with one portion of an image, it changes the overall pattern of activity across the image[61] (Figure 2F).

Effect of tasks on neural activity more generally

Above, we have considered studies investigating the effects of tasks on stimulus-evoked responses. For a more comprehensive characterization of task-related effects, we now broaden our scope to include task influences on neural activity more generally. It is well established that modulations of activity in visual cortex can occur even in the absence of an experimental stimulus. For example, systematic changes in visual activity have been observed for tasks involving generation of visual imagery[62] (Figure 2G), cue-induced memory recall,[72] and allocation of spatial attention in preparation of an upcoming stimulus.[55] In addition, changes in visual activity have been linked to information-processing operations beyond basic stimulus processing such as maintenance of visual working memory,[73–75] prediction,[76,77] temporal attention and expectation,[78–80] and value learning.[81,82] There are even large changes in visual activity caused by locomotion in mouse visual cortex[83,84] (but not so much in primate visual cortex[85,86]). Some studies have revealed processes that appear related to general characteristics of performing a task. Researchers have observed large spatially diffuse activity fluctuations entrained to the timing of a task, reflecting arousal and/or task engagement[63,87,88] (Figure 2H). In addition, the difficulty of a task[89,90] as well as fluctuations in behavioral performance[91] have been shown to have signatures in visual activity. Finally, consider scenarios in which neural responses are altered after an observer is trained to learn specific visual contingencies, such as task-relevant categorization boundaries.[92,93] This is a different, yet interesting, sense in which tasks affect neural activity, one that is less about moment-to-moment changes in neural information processing and more about plasticity of neural response properties.

DIVERSITY OF TASKS USED IN THE STUDY OF VISION

Given that the task of the observer affects activity in the visual system, an immediate question is how to formally approach the study of tasks. We believe a good starting point is to survey the types of tasks that are commonly used in brain and behavioral research (Box 2). It is important to note that the task descriptions provided in the box refer to experimenter-defined tasks: they reflect the point of view of the experimenter and the motivations and mindsets they bring to their experiments. For intuition, we illustrate a variety of example tasks that an experimenter might ask an observer to perform, including cases where different tasks are performed on the same stimulus (Figure 3).

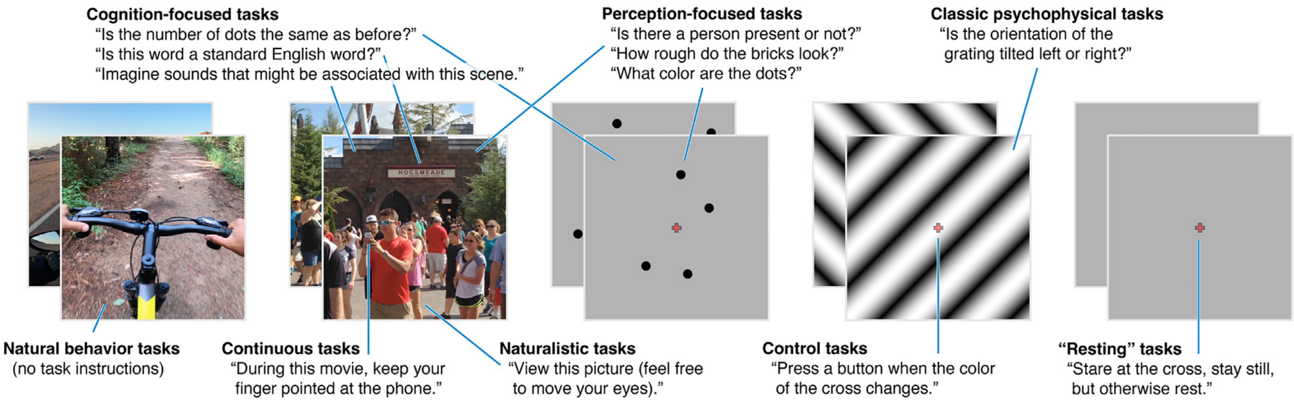

Figure 3. Diversity of tasks that can be performed on visual inputs
Here, we show several examples of visual stimuli, ranging from the complex visual inputs one receives while biking (left) to a simple fixation cross (right). The quoted text indicates different possible task instructions that one might give to an observer. All of these tasks are vision-related in the sense that the goal of the observer depends on properties of the visual input. However, the tasks are notably diverse, with some tasks being more perceptual in nature and other tasks being more cognitive in nature. A central thesis of this article is that understanding how the brain executes a rich diversity of tasks involving visual inputs should be a cornerstone of visual neuroscience.

Experimenter-defined tasks can be understood in terms of variation along two dimensions (Figure 4). One dimension is the level of control that the experimenter exerts over the observer's task. For example, classic psychophysical tasks are designed to tightly control the operations performed by the observer. This contrasts with studies that use natural behavior tasks or "resting" tasks, where little or no control is exerted over the task engaged by the observer.[27] The other dimension is the level of complexity of the task performed by the observer. Some tasks are relatively simple, involving minimal use of brain systems beyond the visual system (e.g., an observer judges the color of a stimulus). Other tasks are relatively complex, involving the coordination of many different sensory, cognitive, and/or motor systems (e.g., an observer watches a movie and remembers the names of the characters).

Above, we surveyed the diverse ways that scientists deploy tasks in different experiments. But what governs the choice of task? We suggest that task choice is often driven by practical considerations. For example, in studies focused on characterizing low-level sensory tuning properties of neurons, the desire to minimize eye movements motivates the use of control tasks that encourage stable central fixation. However, we submit that task choice could instead be driven by the desire to characterize and understand the information-processing operations that observers deploy to execute tasks. This motivation emphasizes the perspective of the observer and what is actually relevant to them during a given experiment.

Using the lens of information processing, we can reconsider how one might decide what tasks to use in a given study. In particular, a basic decision concerns the level of naturalism to employ. On the one hand, rich, complex, and relatively unconstrained naturalistic tasks have ecological validity and are therefore interesting to study in and of themselves.[41] Moreover, we may wish to test how well our understanding of the brain derived from artificial contexts generalizes to naturalistic contexts.[94,95] However, without specially crafted procedures, the complexity of naturalistic paradigms makes it challenging for the experimenter to know what specific information-processing operations might be in play at any given moment in time. In this respect, we express caution concerning approaches in which large-scale annotations of naturalistic experimental paradigms are used to interpret neural activity.[96,97] Even though such annotations may be convenient, they may fail to pinpoint the specific processes engaged by the observer. For example, suppose there is a dog present in a viewed movie and high levels of neural activity are observed. Is this neural activity related to building a coarse sensory representation of the dog independent of attentional allocation, related to attending to the dog and making detailed perceptual decisions about the breed of the dog, or perhaps related to preparation of motor acts such as reaching to pet the dog?

On the other hand, simplified, highly constrained artificial tasks can impose a valuable "behavior clamp"—that is, they can be used to constrain the behavior of an organism to such a degree that the experimenter can have high confidence in identifying the specific cognitive processes underlying that behavior.[28] We acknowledge there is some risk that highly contrived, overconstrained behaviors might lead to misguided lines of research.[98] For example, maintaining central fixation for long periods of time, which has a long tradition in visual psychophysics and visual neuroscience,[99] is highly unnatural and may incur unwanted side effects on the state of the organism (e.g., fatigue[24]). But ultimately, we believe that control is paramount for the goal of understanding information processing, and researchers should therefore consider imposing specific tasks and experimental manipulations on the observer to help control and constrain their behavior.[100]

HOW CAN SCIENTISTS FORMALIZE AND MODEL VISION-RELATED TASKS?

From stimulus-computability to incorporating the observer's task

Visual neuroscientists have made great progress in building models that accurately predict patterns of neural activity in the

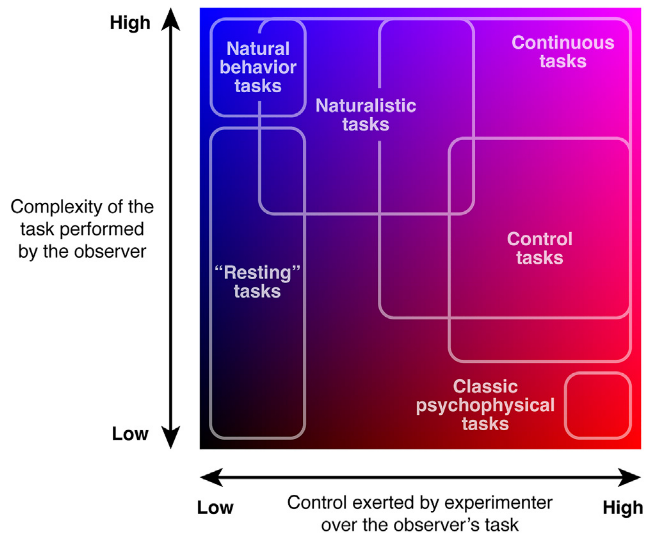

Figure 4. Experimenter-defined tasks vary widely in control and complexity
This schematic illustrates how different kinds of experimenter-defined tasks vary along two key dimensions: the level of control exerted by the experimenter over the actual task performed by the observer (x axis) and the level of complexity of the observer's task (y axis). Some boxes are large, indicating a wide range of possibilities (e.g., some control tasks exert higher levels of control than other control tasks, some observers may engage in more complex tasks during "rest" than other observers, etc.).

visual system evoked by visual stimuli,[6,7] with notable advances made through the use of features found in deep neural networks trained on computer vision objectives.[101–104] These models accept images as input and are often referred to as image-computable, or *stimulus-computable*, models, which emphasizes that they are sufficiently general to generate predictions of neural responses to arbitrary images.[105] In light of the observation that tasks affect activity in the visual system, we face an acute question: how might we augment stimulus-computable models to account for the task of the observer? Can we build upon the remarkable progress in predicting neural responses that has been achieved in stimulus-based modeling and develop a new class of models that can predict outcomes—both brain activity and behavior—for experiments involving arbitrary stimuli *and* arbitrary tasks?

Approach for building task models

Below, we propose a Marrian approach[106] for formalizing and modeling tasks. In brief, the scientist attempts to specify the goal of the observer and the full set of information-processing operations being used to achieve that goal[107,108] and then attempts to assign (map) the information-processing operations to different aspects of neural activity. In our treatment, models of tasks, or *task models*, subsume stimulus-computable models, because they describe not only how stimuli drive visual responses but also how other information-processing operations contribute toward completing a task. The significance of carrying out this endeavor is that it gives serious consideration to all potential information-processing operations deployed in the execution of a task.[109] Note that the approach we describe is not intended to be novel, surprising, or controversial. Indeed, the approach bears general similarity to what has been termed model-based cognitive neuroscience[110,111] and process models in cognitive science.[112] The goal of our contribution here is to distill common ideas into a compact presentation.

David Marr[106] specifies three levels of analysis for the explanation of a cognitive phenomenon, and complete explanations require filling in the details at each level. Here, we take the phenomenon of interest to be task execution by an observer.

(1) At the *computational level*, the scientist develops a formal description of the goal of the observer. This refers to what the observer is trying to compute and why,[113] i.e., the functional relevance of the computation for the observer with respect to the observer's environment. For many tasks studied in the laboratory, the goal is a simple motor act that is contingent on properties of the stimulus. In these cases, the description of the goal may be a relatively straightforward input-output mapping. For complex tasks (e.g., a foraging task), developing a description of the goal may require substantial effort and characterizations of a number of subgoals.
(2) At the *algorithmic level*, the scientist specifies the full set of information-processing operations that the observer uses to achieve the goal, cognizant of the possibility that different observers might employ different operations (e.g., entirely different cognitive strategies) and the possibility that the same observer might even make use of different strategies over time.[114] Ideally, the specification would be formalized into a model using mathematical notation and/or software code such that an arbitrary spatiotemporal sequence of visual inputs could be accepted as input and a time series of motor commands would be output. Cognitive ontologies (i.e., formal descriptions of the types of mental processes and their relationships) could serve as a useful starting point for specifying operations that may be present in different experimental paradigms.[115] For experimental paradigms that involve task switching, it may be critical to consider how the observer represents the task itself,[21,116] i.e., the configuration of information-processing operations constituent of the task, sometimes referred to as task set.
(3) At the *implementational level*, the scientist uses measurements of neural activity to propose possible assignments (mappings) between specific information-processing operations and specific aspects of neural activity (e.g., action potentials, calcium signals, local field potentials, and hemodynamic signals) occurring in various brain structures (e.g., cortical areas, layers, and columns). The underlying presumption is that neural activity is a physical signature of the execution of information-processing operations. The situation, of course, might be complicated: a single operation might be reflected in distributed activity across multiple brain regions (e.g., storage of working memory in sensory, parietal, and prefrontal cortex[117]), multiple operations might be reflected in the activity of a single brain region (e.g., mixed selectivity for task-relevant variables in prefrontal cortex[118]),

and the implementational substrate may be essentially dynamical, with computations achieved by trajectories through a neural state space.[119,120]

Example of task modeling approach

Let's consider an example to help illustrate these general principles. Consider a simple task in which the observer is presented with a series of images and presses a button for each image, indicating whether a person is present ("yes" button) or not ("no" button). To build a task model, the scientist must fill in details at all levels: computational, algorithmic, and implementational.

Details at the computational level concern the overarching goal of the observer in this situation. What are the overall objectives that are being fulfilled? Part of the goal description consists in a specification of which images count as containing a person, and this would address borderline cases such as barely visible persons. Without explicit instruction or training from the experimenter, the observer presumably forms some internal criterion according to which decisions are made. We might try to infer the nature of this criterion from the patterns of button presses made by the observer; this analysis might also reveal periods of time when the observer is not engaged in the task. The goal description might also include a timetable of the speed and frequency at which motor commands are to be issued.

For the algorithmic level, the scientist must form hypotheses regarding the information-processing operations deployed by the observer to achieve the goal. Visual operations are obviously involved, and the functional role of these operations could be to supply a signal that indicates the likelihood that a human body is present in the image. Alternatively, it might be the case that the observer performs the task based on detection of faces. Besides visual operations, a range of other operations may play a role in the task. For example, the observer might deploy a top-down selection process (e.g., feature-based attention) to help find relevant visual information, a decision-making process to manage the visual signal, and an action-selection process to determine which button is ultimately pressed. All of these are distinct computations that require their own formalization in terms of how signals are transformed and how signals from different information-processing operations interact with one another.

Finally, the scientist must specify details at the implementational level. Insofar as high-quality measurements of the observer's brain are available, the scientist must determine which information-processing operations are consistent with the brain data, and discover how the neural circuitry carries out the operations. Many important questions must be resolved. For each brain region, what input-output transformations is that region executing? Are there multiple operations that influence activity in this region? If the observer is disengaged from the task (or is performing a different task), what would the brain data look like? What brain region deploys top-down selection, and how and when does this happen? How do neural circuits manage the speed and resolution of the decision-making process?

Practical challenges of task modeling

The modeling approach we have proposed places deep emphasis on analyzing the structure of tasks. It suggests we should give as much consideration to the analysis of tasks and behavior (computational and algorithmic levels) as we give to the study of neural signals (implementational level), a point that has been compellingly argued by others.[98] A good example of productive work conducted at the computational and algorithmic levels in the context of vision comes from Land and Hayhoe,[12,121] who performed detailed analysis of eye movements and behavior while human observers prepared food (Figure 5). This work delivers rich theory that describes information-processing operations that may be at play during food preparation; developing such theory provides a useful starting point for investigating mechanisms implemented by the brain for completing such a task.

Although our proposed approach for modeling tasks is simple to state, there nonetheless exist deep practical challenges. Building task models is not as straightforward as building stimulus-computable models. Whereas stimuli are directly observable and can serve as a common starting point for different models, tasks must be inferred (e.g., from behavioral data) and thus are already subject to debate. We also highlight the challenge of establishing rigorous connections between hypothesized information-processing operations and neural activity. The field has demonstrated promising initial approaches toward making these connections, for example, in the form of encoding models that construct linear mappings between model features and neural activity[122] and representational similarity analysis (RSA),[123] which matches model features and neural activity at the level of similarity between pairs of experimental conditions. However, we suggest there is further work to be done here: encoding models involve flexible learning of weights, which can lead to distortions of sparsity and dimensionality[124] and interpretation challenges more generally,[125,126] while RSA may be overly coarse and discard too much information.[124,127]

TASKS IN COMPUTER VISION AND ARTIFICIAL INTELLIGENCE

In computer vision and artificial intelligence (AI), the term task is heavily used, and it is common to speak of models as performing tasks, being optimized for certain tasks, exhibiting successful transfer to novel tasks, and so on. A now classic example is the AlexNet artificial neural network (ANN) model trained to perform an object classification task.[128] There has been intense interest in the application of ANN models to neuroscience[129]—in particular, task-optimized deep convolutional ANNs as potential models of the ventral visual pathway.[130,131] Given this interest, it is important to consider similarities and differences between the conceptualization of tasks coming from AI and the one described in this article. In AI, usage of the term task emphasizes the ability to transform a given input (e.g., an image) into some desired output (e.g., a content label). This input-output sense of task is consistent with our description of the goal component of tasks (see Box 1). Thus, we are happy to consider ANNs as performing tasks, insofar as we are willing to consider mathematical objects and software code as having goals. This leads us to a critical question: do ANNs count as full-fledged task models, satisfying all of the requirements proposed in this article?

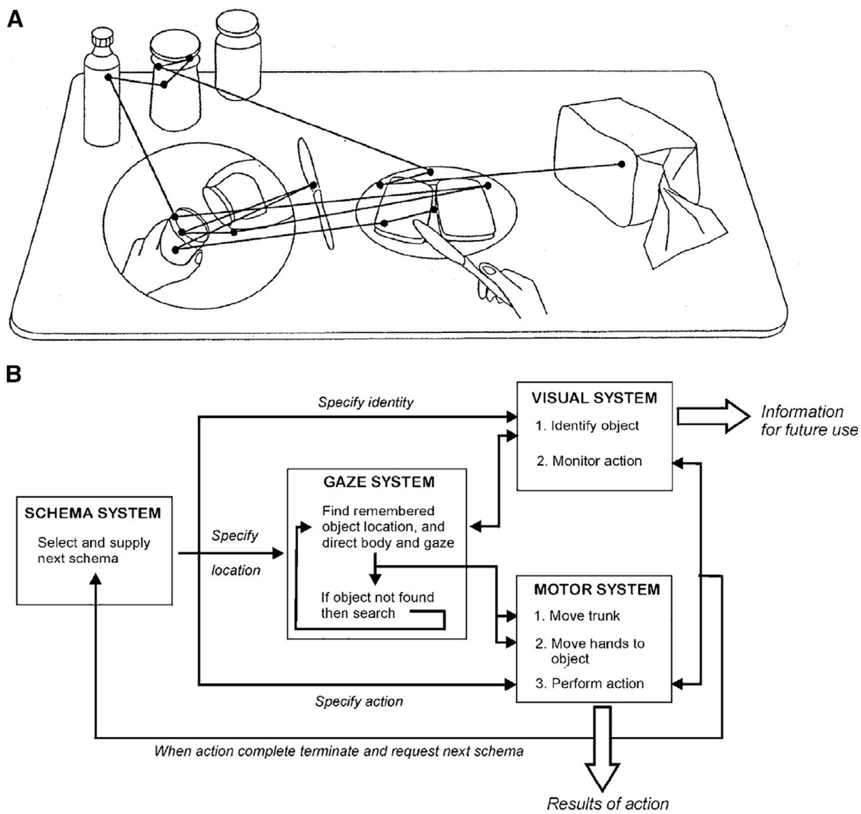

Figure 5. Example of task analysis based on behavioral data

(A) Scan path of a human observer's eyes as the observer makes a sandwich. Dots indicate fixations.

(B) Hypothesized systems involved in executing visually guided action sequences, like those involved in the sandwich-making task. This model was developed based on analyzing and interpreting data like those shown in (A), with particular emphasis on the precise timing of eye movements relative to motor acts. The depicted operations characterize the task at the computational and algorithmic levels of description. Reproduced from Land[12] with permission from Cambridge University Press.

There is a sense in which ANNs are compatible with our proposal for task models. We can view ANNs as a model class that subsumes a large number of model instantiations, each consisting of multiple interacting units that transform information according to certain linear and nonlinear operations. In this sense, it is certainly possible that there exists an ANN that is capable of performing a wide array of vision-related tasks and that mirrors how a biological organism performs these tasks. Compared with the organism, the ANN could be similarly trained and instructed on various experimental paradigms, could manifest similar goals, and could achieve these goals using similar strategies. Hence, our call for building task models does not necessarily break with ongoing development of ANNs in computational and systems neuroscience.[129–131]

However, there are important senses in which ANNs, at least in how they are typically developed and used, are not satisfactory candidates for task models. Although ANNs *perform* tasks, they are not necessarily models *of* tasks. According to our proposed framework, building a task model involves providing descriptions of the information-processing operations that are used to achieve goals; hence, there is an interpretive or explanatory component of the modeling endeavor.[132] Since it is generally difficult to interpret the operations carried out by ANNs,[133] ANNs may fall short of the expectations we have for a task model. Along these lines, it is useful to consider the distinction between a model in the sense of an object that serves as a comparison for a system of interest and a model in the sense of a substantive description of a system of interest. An ANN, like an animal, can be useful as an object to study and to draw parallels with, say, a human. But without further investigation and analysis, the ANN, or animal, provides limited insight into the nature of representations, transformations applied to those representations, and other aspects of information processing. One approach has been to de-emphasize interpretive insights and verbal explanations[105,129] and instead cast understanding in terms of ANN properties such as architecture, objective functions, and learning rules.[129] However, we continue to champion the importance of investigating, identifying, and dissecting information-processing operations when building models and comparing them with the brain.[134]

Another challenge for ANNs as potential task models concerns the scope of the phenomena they characterize. ANNs may provide useful characterizations of computations underlying a given brain region or psychological process of interest, and they may successfully account for experimental observations in certain situations. But we should be cognizant of the full set of potential operations that an organism might bring to bear across different situations. To illustrate, consider the phenomenon of object recognition. From an AI perspective, object recognition might be operationally defined as the successful computation of an appropriate object label. This definition is well matched to thinking of object recognition as a function of a brain region (e.g., inferotemporal cortex), in the sense of transforming sensory inputs into a representation where object category is linearly decodable.[26] But when considering object recognition in the full sense of a task, there are additional phenomena that come into play. Operations that may intervene between sensory inputs and meaningful recognition behavior include not only visual representations but also semantic access, perceptual decision-making, memory retrieval, metacognition, and so on. Certain experimental paradigms can help draw out these operations for deeper study: for example, a human judging the content of a perceptually ambiguous stimulus (e.g., the face-vase illusion) may engage in a variety of cognitive operations (e.g., visual imagery) that may not be strongly engaged in more conventional paradigms.

To be clear, there has been substantial progress in advancing ANNs as possible models of brain computation, and there continue to be many creative ways in which ANNs can be used as investigatory tools. Our enumeration of challenges (concerning interpretation of information-processing operations and scope of modeled phenomena) does not preclude the possibility that ongoing and future research may overcome these challenges. In particular, besides task-optimized deep convolutional ANNs, other flavors of ANNs developed in the cognitive science literature[135,136] might help forge closer links to tasks (in the sense discussed in this article) and cognition.

PROGRESS IN MODELING VISION-RELATED TASKS

We now turn to specific lines of research that have made substantive progress in modeling tasks, focusing on those that are relevant to vision. We briefly discuss these lines of research, highlighting their successes, limitations, and relationship to our proposed approach. A valuable contribution of this discussion is to reveal conceptual links across disparate types of research.

Signal detection theory

Signal detection theory (SDT) is commonly used to model detection and discrimination tasks in psychophysics.[137] SDT posits that the observer transforms a noisy sensory observation into a probabilistic decision variable and uses some criterion to determine a discrete behavioral choice (e.g., ‘‘target present’’ or ‘‘target absent’’). While SDT has been highly successful in vision science, SDT has important limitations as an approach for modeling tasks. First, SDT describes task execution at only the computational and algorithmic levels, although there has been some work pertaining to how SDT might be implemented neurally; notably, the lower envelope hypothesis posits that perceptual thresholds measured in psychophysical experiments are determined by the most sensitive neuron in a population.[138–140] Second, SDT models are highly simplified (by design), with neither a concept of time nor straightforward methods for generalizing from simple one-dimensional representations to the high-dimensional representations that are more representative of naturalistic perception. Most critically, it is not clear how to extend SDT beyond detection and discrimination tasks to model the full diversity of possible tasks (see Figure 3). For example, can the complex perception-action loop involved in, say, riding a bicycle be modeled as a series of detection and discrimination tasks?

Drift diffusion models

Like SDT, drift diffusion models (DDMs) provide a framework for modeling certain types of decision-making tasks. DDM can be viewed as a temporal extension of SDT in which the acquisition of multiple observations over time is explicitly modeled. The core idea underlying DDM is that the observer accumulates evidence (which might be noisy) over time until some bound is reached, at which point the decision is made.[141] There is a long history of success in using DDMs to account for choices and reaction times in perceptual decision-making paradigms,[142] including demonstration of a remarkable correspondence between firing rates in macaque lateral intraparietal area and the evidence-accumulation process hypothesized in DDM.[143] However, like in SDT, there are likely limits on the tasks that can be modeled using DDM. Random dot motion paradigms (often used to study decision-making) are immensely useful but highly simplified and perhaps overly tailored to the DDM framework. Another limitation is that implementations of DDM are often not stimulus-computable (the labels used by the experimenter are assumed to correspond to the actual sensory inputs), and generalization to more complex stimuli therefore remains difficult to assess. A final limitation is that while DDM explicitly addresses time in neural computations, it does not address the complexities of time in behavior. Trial-based paradigms, like those used in many random dot motion experiments, consider a single binary decision made by the observer; however, it is of high interest to understand tasks in which the observer makes high-dimensional decisions continuously over time, mirroring what occurs in naturalistic tasks.[39]

Ideal observer analysis

Ideal observer analysis (IOA) has been used to model visual tasks in which the observer judges some property of a presented image.[144,145] IOA requires a precise specification of the goal (e.g., detect whether a target is present in a natural image patch), the possible visual stimuli that might be presented, and any biological constraints that may be present. The result of IOA is a description of the optimal algorithm that an observer might use to achieve the goal (e.g., which visual features should be extracted and what operations to perform on those features). IOA has delivered important insights into biological vision, especially when coupled with analysis of natural scene statistics.[146,147] Also, many IOA models are stimulus computable. However, the conceptualization of tasks in IOA is narrower than what we have presented in this article. IOA, at least thus far, has primarily focused on sensory properties and how observers estimate these properties. The types of tasks used in IOA experiments are aimed toward isolating perceptual judgments and tend to strip down the behavior of the observer, removing time (but see Straub and Rothkopf[148] where bounded actor models are used to study perception-action loops). Second, IOA is aimed at the computational and algorithmic levels, and there is limited work relating IOA to the implementational level (but see exceptions[149,150]). Finally, much of the power of IOA stems from the assumption that observers are indeed optimal or near-optimal in a given task, but this may not necessarily hold in general.

Partially observable Markov decision processes

In computational reinforcement learning, tasks are often formalized using partially observable Markov decision processes (POMDPs).[151,152] A POMDP is a specification of a set of states, a transition function between states, a reward function, a set of observations from which states are inferred, and a set of possible actions.[153] On the basis of the POMDP and a given loss function, the agent learns a policy that dictates which action to take given the state that the agent is in. For example, we might imagine an experiment where the observer is presented with two colored oriented gratings, color and orientation indicate various probabilities of receiving a reward, and the goal of the observer is to maximize reward through appropriate choice of stimuli. The task in this experiment can be formally modeled using a POMDP specifying how two states (''good'' and ''bad'') are related to the observations (gratings). Although atypical in vision science, POMDPs are appealing as a rigorous and general framework that might accommodate a wide diversity of tasks. However, the POMDP framework adopts a very different perspective on tasks compared with what we have laid out in this article. POMDPs provide a top-down specification that describes the rules by which an experiment unfolds and how an agent *might* behave in the paradigm. This top-down perspective may overlook what the actual task of the observer is (from observer's perspective). Furthermore, POMDPs do not include specification of the information-processing operations used by observers to execute their goals, although it is possible to augment a POMDP with specific proposals (a classic example being the algorithm of temporal-difference learning[154,155]).

Other modeling ideas

As discussed above, SDT, DDM, IOA, and POMDPs are interesting and productive lines of research but have limitations as general approaches to modeling tasks. Hence, we believe the broad endeavor of modeling tasks is still in its infancy, and novel and creative techniques will be critical for forward progress. A number of interesting methods have been demonstrated that may assist in the quest to develop task models. Recent studies have begun tackling the challenge of how a model that performs different tasks even ''knows'' what task it is performing. One approach in the context of an ANN model that is able to perform many tasks is to use an input unit that codes the current task.[156] A different approach is to use natural language embeddings of task instructions,[157,158] an approach that naturally permits generalization to novel tasks.

Other recent studies have used data-driven approaches to advance model development. In one study, visually evoked responses during a task in which attention is diverted away from the stimulus were exploited as a measure of pure sensory evidence in the context of a decision-making model.[59] This study also demonstrated an approach in which neural activity in a specific brain region (intraparietal sulcus) is exploited as a measure of a latent cognitive process (top-down perceptual enhancement) and formally incorporated as an input into a computational model. In another study, a diverse set of experimental paradigms was quantitatively modeled using features obtained by correlating neural activity observed during each paradigm with activity taken from a meta-analytic database of neuroimaging results.[159]

It has been theorized that the visual system infers latent causes of sensory inputs through the use of probabilistic generative models of the environment.[160] Along this line of thinking, one approach is to design stimulus and task manipulations and modeling procedures to probe these generative processes.[161,162] Finally, consider the classic Bubbles technique in which different portions of a stimulus are randomly exposed while the observer attempts to make a certain judgment.[163] This technique has the useful function of isolating the specific parts of the stimulus that are actually used by the observer to perform a task. The identified stimulus parts can then be integrated with brain data in an attempt to pinpoint where, when, and how task-relevant information is processed.[100,164] The long-standing hypothesis that tasks influence perceptual processing by constraining behaviorally relevant stimulus features[9] parallels overall themes in this article.

CONCLUSIONS AND FUTURE OUTLOOK

In this article, we have offered a broad perspective on the concept of tasks, reviewed the diverse ways that tasks impact neural activity in the visual system, and outlined an approach for the formal modeling of tasks. To make our perspective more concrete, we suggest three action items for future research. First, we should keep in mind that observations of neural activity in a given study may depend on the specific task performed by the observer. Moreover, if a neural response property is found to change across conditions or participants, we should consider differences in the task performed as a potential cause, as opposed to differences in stimulus processing per se. Second, to better understand the visual system, we should sample stimulus-evoked responses under a wide range of tasks. Any observation of response variability across tasks should be embraced, not ignored, nor downplayed. Third, we should build models that incorporate both the stimulus and the task of the observer. In particular, we have proposed a modeling approach in which tasks are decomposed into formally specified goals and associated information-processing operations and then mapped to neural activity.

We offer a few remarks to ensure the clarity of our overall message. While we have reviewed studies that demonstrate the effects of tasks on neural activity in the visual system, it is important to put these effects into perspective and to accurately assess the size and nature of the effects. For example, category preferences in high-level visual cortex are maintained across different tasks,[68] and face-related cortical activations occur even when the observer is anesthetized.[165] Moreover, useful progress in understanding response properties in the visual system has been made using control tasks. In particular, substantial mileage has been gained in building quantitative links between neural responses in the ventral visual stream during control tasks (e.g., an observer maintains central fixation while brief images are presented) to certain object-recognition behaviors.[166] Our main contention is that tasks are important, not because they contribute to massive amounts of variance in neural activity (the effects might be relatively small in some particular brain region), but for the deeper reason that they reflect critical information-processing operations engaged by the observer and are worthy of study in and of themselves. While control tasks provide a useful starting point, we should extend the scope of our investigations to a wide range of sensory, cognitive, and motor tasks (e.g., tasks involving eye movements, cognitive judgments, temporally extended goals, or motor goals) to improve our understanding of visual processing. The broader philosophical point here is that there is no pure vision[167]: the visual system is not an isolated brain system that builds representations of the world detached from other concerns. Rather, the visual system is closely and deeply integrated with other brain systems in service of achieving the goals of the observer, especially action-related goals.[11,12]

Looking ahead, we suggest that placing special consideration on time will be crucial for progress in building task models.

Biological organisms operate in dynamic environments and under time constraints. To be complete, models will need to specify, for example, how information is selected or attended over time, that is, the deployment of temporal attention[78,168]; how variations in behavioral performance might manifest over time[41]; and how engagement with the task might itself fluctuate.[87,169] In fact, depending on time constraints, the same set of task instructions may lead to differences in how information-processing operations are deployed,[170] with a well-known example being the attentional blink.[171] We further suggest that comparing models of different tasks could provide major insights. Experimental studies targeting seemingly different phenomena might actually be probing common operations and neural mechanisms (e.g., the overlap between working memory, spatial attention, and visual imagery[172,173]), and building formal models may help scientists generalize across tasks and may unify different experimental traditions. Along these lines, there is evidence that common reusable components (compositionality) underlie the execution of diverse tasks.[156,174,175]

Thinking deeply about tasks has implications for other topics in visual neuroscience. It is commonplace to conduct open-loop experiments in which stimuli are presented briefly, and the observer makes some judgment on the stimulus. However, if we allow eye movements and other motor acts—thus closing the loop[24]—it becomes clear that tasks affect even the earliest stages of sensory processing that occur in the retina. This is because the task of an observer exerts strong influence on where in an environment the observer looks.[11,176,177] Hence, even the study of low-level sensory processing is not immune to the influence of tasks. Consider, for example, theoretical insights into early visual processing that have been gained by analyzing the statistics of natural stimuli.[178] This line of research might benefit from considering variations in stimulus statistics that one might experience depending on the task performed. Efforts to collect databases of naturalistic sensory inputs while observers navigate and act in the real world may be particularly useful in this regard.[43,179–181]

There are also important implications of tasks for visual development. Over the course of development, there are extensive changes in the repertoire of tasks that are performed by an individual. These changes in tasks are partly the result of maturation of basic visual functions such as acuity, stereopsis, motion perception, and contour integration—what an infant can see will determine how they engage and interact with the world.[182,183] Changes in tasks are also influenced by developmental milestones, such as movement-related maturation (e.g., sitting up, crawling, and walking), which dramatically alter the sensory inputs received by an individual.[184–186] Qualitatively new tasks, in the form of novel ways of exploring one's environment and physically interacting with objects and people, enable in turn an influx of new sensory experiences. For example, typical visual experience over the first post-natal year shifts from scenes containing primarily faces and nondescript backgrounds to those with hands and objects.[187] Thus, task dynamics are embedded within a developmental trajectory,[188] with complex bidirectional interactions between tasks, sensory inputs, perceptual abilities, and motor abilities. Measuring and modeling these intertwined phenomena constitute a long-term objective for visual neuroscience.

ACKNOWLEDGMENTS

We thank E. Cooper for extensive conversations and help with the initial version of this manuscript. We also thank I. Charest, M. Cole, C. Curtis, R. Poldrack, A. Schapiro, and M. Waskom for helpful discussions. Finally, we thank J. Breedlove, T. Knapen, T. Naselaris, B. Turner, and A. White for comments on the manuscript. This work was supported by startup funds from Boston University (R.N.D.).

AUTHOR CONTRIBUTIONS

K.K. wrote the paper with content contributions from K.B., R.N.D., M.J.A., and D.L.B. All authors discussed and edited the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. DeYoe, E.A., and Van Essen, D.C. (1988). Concurrent processing streams in monkey visual cortex. Trends Neurosci. *11*, 219–226.

2. Tootell, R.B., Dale, A.M., Sereno, M.I., and Malach, R. (1996). New images from human visual cortex. Trends Neurosci. *19*, 481–489.

3. Van Essen, D.C., and Gallant, J.L. (1994). Neural mechanisms of form and motion processing in the primate visual system. Neuron *13*, 1–10.

4. Zeki, S., and Shipp, S. (1988). The functional logic of cortical connections. Nature *335*, 311–317.

5. Orban, G.A., Dupont, P., Vogels, R., De Bruyn, B., Bormans, G., and Mortelmans, L. (1996). Task dependency of visual processing in the human visual system. Behav. Brain Res. *76*, 215–223.

6. Wu, M.C.-K., David, S.V., and Gallant, J.L. (2006). Complete functional characterization of sensory neurons by system identification. Annu. Rev. Neurosci. *29*, 477–505.

7. Carandini, M., Demb, J.B., Mante, V., Tolhurst, D.J., Dan, Y., Olshausen, B.A., Gallant, J.L., and Rust, N.C. (2005). Do we know what the early visual system does? J. Neurosci. *25*, 10577–10597.

8. Maunsell, J.H.R. (2015). Neuronal mechanisms of visual attention. Annu. Rev. Vis. Sci. *1*, 373–391.

9. Schyns, P.G. (1998). Diagnostic recognition: task constraints, object information, and their interactions. Cognition *67*, 147–179.

10. Gilbert, C.D., and Li, W. (2013). Top-down influences on visual processing. Nat. Rev. Neurosci. *14*, 350–363.

11. Hayhoe, M.M. (2017). Vision and action. Annu. Rev. Vis. Sci. *3*, 389–413.

12. Land, M.F. (2009). Vision, eye movements, and natural behavior. Vis. Neurosci. *26*, 51–62.

13. Conway, B.R. (2018). The organization and operation of inferior temporal cortex. Annu. Rev. Vis. Sci. *4*, 381–402.

14. Pasupathy, A., Popovkina, D.V., and Kim, T. (2020). Visual functions of primate area V4. Annu. Rev. Vis. Sci. *6*, 363–385.

15. Bracci, S., and Op de Beeck, H.P. (2023). Understanding human object vision: a picture is worth a thousand representations. Annu. Rev. Psychol. *74*, 113–135. https://doi.org/10.1146/annurev-psych-032720-041031.

16. Kaldas, A. (2022). Attention: a descriptive taxonomy. Hist. Philos. Life Sci. *44*, 63.

17. Hommel, B., Chapman, C.S., Cisek, P., Neyedli, H.F., Song, J.H., and Welsh, T.N. (2019). No one knows what attention is. Atten. Percept. Psychophys. *81*, 2288–2303.

18. Hilgetag, C.C., and Goulas, A. (2020). "Hierarchy" in the organization of brain networks. Philos. Trans. R. Soc. Lond. B Biol. Sci. *375*, 20190319.

19. Barack, D.L., Miller, E.K., Moore, C.I., Packer, A.M., Pessoa, L., Ross, L.N., and Rust, N.C. (2022). A call for more clarity around causality in neuroscience. Trends Neurosci. *45*, 654–655.

20. Pylyshyn, Z.W. (1986). Computation and Cognition: Toward a Foundation for Cognitive Science (Bradford Books).

21. Hazeltine, E., Dykstra, T., and Schumacher, E. (2022). What is a task and how do you know if you have one or more? In Theory and History in the Human and Social Sciences (Springer International Publishing), pp. 75–95.

22. Angelucci, A., and Bressloff, P.C. (2006). Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. Prog. Brain Res. *154*, 93–120.

23. Seo, H., Cai, X., Donahue, C.H., and Lee, D. (2014). Neural correlates of strategic reasoning during competitive games. Science *346*, 340–343.

24. Gomez-Marin, A., and Ghazanfar, A.A. (2019). The life of behavior. Neuron *104*, 25–36.

25. Schröder, S., Steinmetz, N.A., Krumin, M., Pachitariu, M., Rizzi, M., Lagnado, L., Harris, K.D., and Carandini, M. (2020). Arousal modulates retinal output. Neuron *107*. 487–495.e9.

26. DiCarlo, J.J., and Cox, D.D. (2007). Untangling invariant object recognition. Trends Cogn. Sci. *11*, 333–341.

27. Tibon, R., Geerligs, L., and Campbell, K. (2022). Bridging the big (data) gap: levels of control in small- and large-scale cognitive neuroscience research. Trends Neurosci. *45*, 507–516. https://doi.org/10.1016/j.tins.2022.03.011.

28. Waskom, M.L., Okazawa, G., and Kiani, R. (2019). Designing and interpreting psychophysical investigations of cognition. Neuron *104*, 100–112.

29. Greene, M.R., and Oliva, A. (2009). The briefest of glances: the time course of natural scene understanding. Psychol. Sci. *20*, 464–472.

30. Pestilli, F., and Carrasco, M. (2005). Attention enhances contrast sensitivity at cued and impairs it at uncued locations. Vision Res. *45*, 1867–1875.

31. Fechner, G.T. (1860). Elemente der Psychophysik (Breitkopf u. Härtel).

32. Stigliani, A., Weiner, K.S., and Grill-Spector, K. (2015). Temporal processing capacity in high-level visual cortex is domain specific. J. Neurosci. *35*, 12412–12424.

33. Graumann, M., Ciuffi, C., Dwivedi, K., Roig, G., and Cichy, R.M. (2022). The spatiotemporal neural dynamics of object location representations in the human brain. Nat. Hum. Behav. *6*, 796–811.

34. Hallum, L.E., Landy, M.S., and Heeger, D.J. (2011). Human primary visual cortex (V1) is selective for second-order spatial frequency. J. Neurophysiol. *105*, 2121–2131.

35. Popovkina, D.V., and Pasupathy, A. (2022). Task context modulates feature-selective responses in area V4. J. Neurosci. *42*, 6408–6423.

36. Hurlburt, R.T., Alderson-Day, B., Fernyhough, C., and Kühn, S. (2015). What goes on in the resting-state? A qualitative glimpse into resting-state experience in the scanner. Front. Psychol. *6*, 1535.

37. Delamillieure, P., Doucet, G., Mazoyer, B., Turbelin, M.R., Delcroix, N., Mellet, E., Zago, L., Crivello, F., Petit, L., Tzourio-Mazoyer, N., and Joliot, M. (2010). The resting state questionnaire: an introspective questionnaire for evaluation of inner experience during the conscious resting state. Brain Res. Bull. *81*, 565–573.

38. Bonnen, K., Burge, J., Yates, J., Pillow, J., and Cormack, L.K. (2015). Continuous psychophysics: target-tracking to measure visual sensitivity. J. Vis. *15*. https://doi.org/10.1167/15.3.14.

39. Huk, A., Bonnen, K., and He, B.J. (2018). Beyond trial-based paradigms: continuous behavior, ongoing neural activity, and natural stimuli. J. Neurosci. *38*, 7551–7558.

40. Knöll, J., Pillow, J.W., and Huk, A.C. (2018). Lawful tracking of visual motion in humans, macaques, and marmosets in a naturalistic, continuous, and untrained behavioral context. Proc. Natl. Acad. Sci. USA *115*, E10486–E10494.

41. Miller, C.T., Gire, D., Hoke, K., Huk, A.C., Kelley, D., Leopold, D.A., Smear, M.C., Theunissen, F., Yartsev, M., and Niell, C.M. (2022). Natural behavior is the language of the brain. Curr. Biol. *32*, R482–R493.

42. 't Hart, B.M., and Einhäuser, W. (2012). Mind the step: complementary effects of an implicit task on eye and head movements in real-life gaze allocation. Exp. Brain Res. *223*, 233–249. https://doi.org/10.1007/s00221-012-3254-x.

43. Matthis, J.S., Yates, J.L., and Hayhoe, M.M. (2018). Gaze and the control of foot placement when walking in natural terrain. Curr. Biol. *28*. 1224–1233.e5.

44. Gilja, V., Chestek, C.A., Nuyujukian, P., Foster, J., and Shenoy, K.V. (2010). Autonomous head-mounted electrophysiology systems for freely behaving primates. Curr. Opin. Neurobiol. *20*, 676–686.

45. Voloh, B., Eisenreich, B.R., Maisson, D.J.-N., Becket Ebitz, R., Park, H.S., Hayden, B.Y., and Zimmermann, J. (2021). Hierarchical organization of rhesus macaque behavior. Preprint at bioRxiv. https://doi.org/10.1101/2021.11.15.468721.

46. Yartsev, M.M., and Ulanovsky, N. (2013). Representation of three-dimensional space in the hippocampus of flying bats. Science *340*, 367–372.

47. Dilks, D.D., Kamps, F.S., and Persichetti, A.S. (2022). Three cortical scene systems and their development. Trends Cogn. Sci. *26*, 117–127.

48. Moran, J., and Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. Science *229*, 782–784.

49. Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. Annu. Rev. Neurosci. *18*, 193–222.

50. Maunsell, J.H.R., and Treue, S. (2006). Feature-based attention in visual cortex. Trends Neurosci. *29*, 317–322.

51. Treue, S., and Martínez Trujillo, J.C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. Nature *399*, 575–579.

52. Womelsdorf, T., Anton-Erxleben, K., Pieper, F., and Treue, S. (2006). Dynamic shifts of visual receptive fields in cortical area MT by spatial attention. Nat. Neurosci. *9*, 1156–1160.

53. Silver, M.A., Ress, D., and Heeger, D.J. (2005). Topographic maps of visual spatial attention in human parietal cortex. J. Neurophysiol. *94*, 1358–1371.

54. Dugué, L., Merriam, E.P., Heeger, D.J., and Carrasco, M. (2020). Differential impact of endogenous and exogenous attention on activity in human visual cortex. Sci. Rep. *10*, 21274.

55. Kastner, S., Pinsk, M.A., De Weerd, P., Desimone, R., and Ungerleider, L.G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. Neuron *22*, 751–761.

56. Serences, J.T., and Boynton, G.M. (2007). Feature-based attentional modulations in the absence of direct visual stimulation. Neuron *55*, 301–312.

57. Kay, K.N., Weiner, K.S., and Grill-Spector, K. (2015). Attention reduces spatial uncertainty in human ventral temporal cortex. Curr. Biol. *25*, 595–600.

58. Harel, A., Kravitz, D.J., and Baker, C.I. (2014). Task context impacts visual object processing differentially across the cortex. Proc. Natl. Acad. Sci. USA *111*, E962–E971.

59. Kay, K.N., and Yeatman, J.D. (2017). Bottom-up and top-down computations in word- and face-selective cortex. eLife *6*, e22341.

60. Çukur, T., Nishimoto, S., Huth, A.G., and Gallant, J.L. (2013). Attention during natural vision warps semantic representation across the human brain. Nat. Neurosci. *16*, 763–770.

61. Zipser, K. (2017). Visualizing fMRI BOLD responses to diverse naturalistic scenes using retinotopic projection. J. Vis. *17*, 18.

62. Breedlove, J.L., St-Yves, G., Olman, C.A., and Naselaris, T. (2020). Generative feedback explains distinct brain activity codes for seen and mental images. Curr. Biol. *30*. 2211–2224.e6.

63. Sirotin, Y.B., and Das, A. (2009). Anticipatory haemodynamic signals in sensory cortex not predicted by local neuronal activity. Nature *457*, 475–479.

64. Koida, K., and Komatsu, H. (2007). Effects of task demands on the responses of color-selective neurons in the inferior temporal cortex. Nat. Neurosci. *10*, 108–116.

65. Mante, V., Sussillo, D., Shenoy, K.V., and Newsome, W.T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature *503*, 78–84.

66. Li, W., Piëch, V., and Gilbert, C.D. (2004). Perceptual learning and top-down influences in primary visual cortex. Nat. Neurosci. *7*, 651–657.

67. Bracci, S., Daniels, N., and Op de Beeck, H. (2017). Task context overrules object- and category-related representational content in the human parietal cortex. Cereb. Cortex *27*, 310–321. https://doi.org/10.1093/cercor/bhw419.

68. Bugatus, L., Weiner, K.S., and Grill-Spector, K. (2017). Task alters category representations in prefrontal but not high-level visual cortex. Neuroimage *155*, 437–449.

69. Farzmahdi, A., Fallah, F., Rajimehr, R., and Ebrahimpour, R. (2021). Task-dependent neural representations of visual object categories. Eur. J. Neurosci. *54*, 6445–6462.

70. Hebart, M.N., Bankson, B.B., Harel, A., Baker, C.I., and Cichy, R.M. (2018). The representational dynamics of task and object processing in humans. eLife *7*, e32816. https://doi.org/10.7554/eLife.32816.

71. White, A.L., Kay, K.N., Tang, K.A., and Yeatman, J.D. (2023). Engaging in word recognition elicits highly specific modulations in visual cortex. Curr. Biol. https://doi.org/10.1016/j.cub.2023.02.042.

72. Favila, S.E., Kuhl, B.A., and Winawer, J. (2022). Perception and memory have distinct spatial tuning properties in human visual cortex. Nat. Commun. *13*, 5864.

73. Rademaker, R.L., Chunharas, C., and Serences, J.T. (2019). Coexisting representations of sensory and mnemonic information in human visual cortex. Nat. Neurosci. *22*, 1336–1344.

74. Kwak, Y., and Curtis, C.E. (2022). Unveiling the abstract format of mnemonic representations. Neuron *110*. 1822–1828.e5.

75. Harrison, S.A., and Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. Nature *458*, 632–635.

76. Kok, P., Jehee, J.F.M., and de Lange, F.P. (2012). Less is more: expectation sharpens representations in the primary visual cortex. Neuron *75*, 265–270.

77. Yon, D., Gilbert, S.J., de Lange, F.P., and Press, C. (2018). Action sharpens sensory representations of expected outcomes. Nat. Commun. *9*, 4288.

78. Ghose, G.M., and Maunsell, J.H.R. (2002). Attentional modulation in visual cortex depends on task timing. Nature *419*, 616–620.

79. Lima, B., Singer, W., and Neuenschwander, S. (2011). Gamma responses correlate with temporal expectation in monkey primary visual cortex. J. Neurosci. *31*, 15919–15931.

80. Denison, R.N., Tian, K., Heeger, D.J., and Carrasco, M. (2022). Dynamics of temporal attention revealed by continuous tracking of visual cortical responses. Preprint at bioRxiv. https://doi.org/10.1101/2022.11.18.517084.

81. Henschke, J.U., Dylda, E., Katsanevaki, D., Dupuy, N., Currie, S.P., Amvrosiadis, T., Pakan, J.M.P., and Rochefort, N.L. (2020). Reward association enhances stimulus-specific representations in primary visual cortex. Curr. Biol. *30*. 1866–1880.e5.

82. Itthipuripat, S., Vo, V.A., Sprague, T.C., and Serences, J.T. (2019). Value-driven attentional capture enhances distractor representations in early visual cortex. PLoS Biol. *17*, e3000186.

83. Niell, C.M., and Stryker, M.P. (2010). Modulation of visual responses by behavioral state in mouse visual cortex. Neuron *65*, 472–479.

84. Ayaz, A., Saleem, A.B., Schölvinck, M.L., and Carandini, M. (2013). Locomotion controls spatial integration in mouse visual cortex. Curr. Biol. *23*, 890–894.

85. Talluri, B.C., Kang, I., Lazere, A., Quinn, K.R., Kaliss, N., Yates, J.L., Butts, D.A., and Nienborg, H. (2022). Activity in primate visual cortex is minimally driven by spontaneous movements. Preprint at bioRxiv. https://doi.org/10.1101/2022.09.08.507006.

86. Liska, J.P., Rowley, D.P., Nguyen, T.T.K., Muthmann, J.-O., Butts, D.A., Yates, J.L., and Huk, A.C. (2022). Running modulates primate and rodent visual cortex via common mechanism but quantitatively distinct implementation. Preprint at bioRxiv. https://doi.org/10.1101/2022.06.13.495712.

87. Roth, Z.N., Ryoo, M., and Merriam, E.P. (2020). Task-related activity in human visual cortex. PLoS Biol. *18*, e3000921.

88. Jack, A.I., Shulman, G.L., Snyder, A.Z., McAvoy, M., and Corbetta, M. (2006). Separate modulations of human V1 associated with spatial attention and task structure. Neuron *51*, 135–147.

89. Chen, Y., Martinez-Conde, S., Macknik, S.L., Bereshpolova, Y., Swadlow, H.A., and Alonso, J.M. (2008). Task difficulty modulates the activity of specific neuronal populations in primary visual cortex. Nat. Neurosci. *11*, 974–982.

90. Ress, D., Backus, B.T., and Heeger, D.J. (2000). Activity in primary visual cortex predicts performance in a visual detection task. Nat. Neurosci. *3*, 940–945.

91. Huk, A.C., and Heeger, D.J. (2000). Task-related modulation of visual cortex. J. Neurophysiol. *83*, 3525–3536.

92. Sigala, N., and Logothetis, N.K. (2002). Visual categorization shapes feature selectivity in the primate temporal cortex. Nature *415*, 318–320.

93. Freedman, D.J., Riesenhuber, M., Poggio, T., and Miller, E.K. (2006). Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. Cereb. Cortex *16*, 1631–1644.

94. David, S.V., Vinje, W.E., and Gallant, J.L. (2004). Natural stimulus statistics alter the receptive field structure of V1 neurons. J. Neurosci. *24*, 6991–7006.

95. Jovanovic, V., Fishbein, A.R., de la Mothe, L., Lee, K.F., and Miller, C.T. (2022). Behavioral context affects social signal representations within single primate prefrontal cortex neurons. Neuron *110*, 1318–1326.e4.

96. de la Vega, A., Rocca, R., Blair, R.W., Markiewicz, C.J., Mentch, J., Kent, J.D., Herholz, P., Ghosh, S.S., Poldrack, R.A., and Yarkoni, T. (2022). Neuroscout, a unified platform for generalizable and reproducible fMRI research. eLife *11*. https://doi.org/10.7554/eLife.79277.

97. Sonkusare, S., Breakspear, M., and Guo, C. (2019). Naturalistic stimuli in neuroscience: critically acclaimed. Trends Cogn. Sci. *23*, 699–714.

98. Krakauer, J.W., Ghazanfar, A.A., Gomez-Marin, A., MacIver, M.A., and Poeppel, D. (2017). Neuroscience needs behavior: correcting a reductionist bias. Neuron *93*, 480–490.

99. Leopold, D.A., and Park, S.H. (2020). Studying the visual brain in its natural rhythm. Neuroimage *216*, 116790.

100. Schyns, P.G., Zhan, J., Jack, R.E., and Ince, R.A.A. (2020). Revealing the information contents of memory within the stimulus information representation framework. Philos. Trans. R. Soc. Lond. B Biol. Sci. *375*, 20190705.

101. Yamins, D.L.K., Hong, H., Cadieu, C.F., Solomon, E.A., Seibert, D., and DiCarlo, J.J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. Proc. Natl. Acad. Sci. USA *111*, 8619–8624.

102. Eickenberg, M., Gramfort, A., Varoquaux, G., and Thirion, B. (2017). Seeing it all: convolutional network layers map the function of the human visual system. Neuroimage *152*, 184–194. https://doi.org/10.1016/j.neuroimage.2016.10.001.

103. Güçlü, U., and van Gerven, M.A.J. (2015). Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream. J. Neurosci. *35*, 10005–10014.

104. Khaligh-Razavi, S.M., and Kriegeskorte, N. (2014). Deep supervised, but not unsupervised, models may explain IT cortical representation. PLoS Comput. Biol. *10*, e1003915.

105. Schrimpf, M., Kubilius, J., Lee, M.J., Ratan Murty, N.A., Ajemian, R., and DiCarlo, J.J. (2020). Integrative benchmarking to advance neurally mechanistic models of human intelligence. Neuron *108*, 413–423.

106. Marr, D. (1982). Vision: A Computational Investigation into the Human Representation and Processing of Visual Information (Henry Holt and co. Inc.), pp. 2–46.

107. Anderson, J.R. (1978). Arguments concerning representations for mental imagery. Psychol. Rev. *85*, 249–277.

108. Anderson, J.R. (1990). The Adaptive Character of Thought (Lawrence Erlbaum Associates, Incorp.).

109. Anderson, J.R. (2007). How Can the Human Mind Occur in the Physical Universe? (Oxford University Press).

110. Palmeri, T.J., Love, B.C., and Turner, B.M. (2017). Model-based cognitive neuroscience. J. Math. Psychol. *76*, 59–64.

111. Forstmann, B.U., and Wagenmakers, E.-J. (2015). An Introduction to Model-Based Cognitive Neuroscience (Springer).

112. Jarecki, J.B., Tan, J.H., and Jenny, M.A. (2020). A framework for building cognitive process models. Psychon. Bull. Rev. *27*, 1218–1229.

113. Shagrir, O. (2010). Marr on computational-level theories. Philos. Sci. *77*, 477–500.

114. Ashwood, Z.C., Roy, N.A., Stone, I.R., International Brain Laboratory, Urai, A.E., Churchland, A.K., Pouget, A., and Pillow, J.W. (2022). Mice alternate between discrete strategies during perceptual decision-making. Nat. Neurosci. *25*, 201–212.

115. Walters, J., King, M., Bissett, P.G., Ivry, R.B., Diedrichsen, J., and Poldrack, R.A. (2022). Predicting brain activation maps for arbitrary tasks with cognitive encoding models. Neuroimage *263*, 119610.

116. Miller, E.K., and Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. Annu. Rev. Neurosci. *24*, 167–202.

117. Christophel, T.B., Klink, P.C., Spitzer, B., Roelfsema, P.R., and Haynes, J.D. (2017). The distributed nature of working memory. Trends Cogn. Sci. *21*, 111–124.

118. Rigotti, M., Barak, O., Warden, M.R., Wang, X.J., Daw, N.D., Miller, E.K., and Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. Nature *497*, 585–590.

119. Barack, D.L., and Krakauer, J.W. (2021). Two views on the cognitive brain. Nat. Rev. Neurosci. *22*, 359–371.

120. Vyas, S., Golub, M.D., Sussillo, D., and Shenoy, K.V. (2020). Computation through neural population dynamics. Annu. Rev. Neurosci. *43*, 249–275.

121. Land, M.F., and Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? Vision Res. *41*, 3559–3565.

122. Naselaris, T., Kay, K.N., Nishimoto, S., and Gallant, J.L. (2011). Encoding and decoding in fMRI. Neuroimage *56*, 400–410.

123. Kriegeskorte, N., Mur, M., and Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. Front. Syst. Neurosci. *2*, 4.

124. Saxe, A., Nelli, S., and Summerfield, C. (2021). If deep learning is the answer, what is the question? Nat. Rev. Neurosci. *22*, 55–67.

125. Kriegeskorte, N., and Douglas, P.K. (2019). Interpreting encoding and decoding models. Curr. Opin. Neurobiol. *55*, 167–179.

126. Ivanova, A.A., Schrimpf, M., Anzellotti, S., Zaslavsky, N., Fedorenko, E., and Isik, L. (2022). Beyond linear regression: mapping models in cognitive neuroscience should align with research goals. Preprint at arXiv. https://doi.org/10.48550/arXiv.2208.10668.

127. Dujmović, M., Bowers, J.S., Adolfi, F., and Malhotra, G. (2022). The pitfalls of measuring representational similarity using representational similarity analysis. Preprint at bioRxiv. https://doi.org/10.1101/2022.04.05.487135.

128. Krizhevsky, A., Sutskever, I., and Hinton, G.E. (2012). ImageNet classification with deep convolutional neural networks. NeurIPS Proceedings, 1097–1105.

129. Richards, B.A., Lillicrap, T.P., Beaudoin, P., Bengio, Y., Bogacz, R., Christensen, A., Clopath, C., Costa, R.P., de Berker, A., Ganguli, S., et al. (2019). A deep learning framework for neuroscience. Nat. Neurosci. *22*, 1761–1770.

130. Yamins, D.L.K., and DiCarlo, J.J. (2016). Using goal-driven deep learning models to understand sensory cortex. Nat. Neurosci. *19*, 356–365.

131. Kriegeskorte, N. (2015). Deep neural networks: a new framework for modeling biological vision and brain information processing. Annu. Rev. Vision Sci. *1*, 417–446.

132. Kay, K.N. (2018). Principles for models of neural information processing. Neuroimage *180*, 101–109. https://doi.org/10.1016/j.neuroimage.2017.08.016.

133. Yang, G.R., and Wang, X.J. (2021). Artificial neural networks for neuroscientists: a primer. Neuron *109*, 739.

134. Schyns, P.G., Snoek, L., and Daube, C. (2022). Degrees of algorithmic equivalence between the brain and its DNN models. Trends Cogn. Sci. *26*, 1090–1102.

135. Botvinick, M., Wang, J.X., Dabney, W., Miller, K.J., and Kurth-Nelson, Z. (2020). Deep reinforcement learning and its neuroscientific implications. Neuron *107*, 603–616.

136. Russin, J., O'Reilly, R.C., and Bengio, Y. (2020). Deep learning needs a prefrontal cortex. Bridging AI Cogn. Sci.

137. Green, D.M., and Swets, J.A. (1966). Signal Detection Theory and Psychophysics (John Wiley).

138. De Valois, R.L., Abramov, I., and Mead, W.R. (1967). Single cell analysis of wavelength discrimination at the lateral geniculate nucleus in the macaque. J. Neurophysiol. *30*, 415–433.

139. Barlow, H.B. (1972). Single units and sensation: a neuron doctrine for perceptual psychology? Perception *1*, 371–394.

140. Parker, A.J., and Newsome, W.T. (1998). Sense and the single neuron: probing the physiology of perception. Annu. Rev. Neurosci. *21*, 227–277.

141. Ratcliff, R., and McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. Neural Comput. *20*, 873–922.

142. Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. Annu. Rev. Neurosci. *30*, 535–574.

143. Shadlen, M.N., and Newsome, W.T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. J. Neurophysiol. *86*, 1916–1936.

144. Geisler, W.S. (2011). Contributions of ideal observer theory to vision research. Vision Res. *51*, 771–781.

145. Knill, D.C., and Richards, W. (1996). Perception as Bayesian Inference (Cambridge University Press).

146. Geisler, W.S. (2008). Visual perception and the statistical properties of natural scenes. Annu. Rev. Psychol. *59*, 167–192.

147. Burge, J. (2020). Image-computable ideal observers for tasks with natural stimuli. Annu. Rev. Vis. Sci. *6*, 491–517.

148. Straub, D., and Rothkopf, C.A. (2022). Putting perception into action with inverse optimal control for continuous psychophysics. eLife *11*, e76635. https://doi.org/10.7554/eLife.76635.

149. Burge, J., and Jaini, P. (2017). Accuracy maximization analysis for sensory-perceptual tasks: computational improvements, filter robustness, and coding advantages for scaled additive noise. PLoS Comput. Biol. *13*, e1005281.

150. 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.

151. Kaelbling, L.P., Littman, M.L., and Cassandra, A.R. (1998). Planning and acting in partially observable stochastic domains. Artif. Intell. *101*, 99–134. https://doi.org/10.1016/S0004-3702(98)00023-X.

152. Sutton, R.S., and Barto, A.G. (2018). In Reinforcement Learning: An Introduction, Second Edition (MIT Press).

153. Spaan, M.T.J. (2012). Partially observable markov decision processes. In Reinforcement Learning: State-of-the-Art, M. Wiering and M. van Otterlo, eds. (Springer), pp. 387–414.

154. Sutton, R.S. (1988). Learning to predict by the methods of temporal differences. Mach. Learn. *3*, 9–44.

155. Schultz, W. (1998). Predictive reward signal of dopamine neurons. J. Neurophysiol. *80*, 1–27.

156. Yang, G.R., Joglekar, M.R., Song, H.F., Newsome, W.T., and Wang, X.J. (2019). Task representations in neural networks trained to perform many cognitive tasks. Nat. Neurosci. *22*, 297–306.

157. Riveland, R., and Pouget, A. (2022). A neural model of task compositionality with natural language instructions. Preprint at bioRxiv. https://doi.org/10.1101/2022.02.22.481293.

158. Jang, E., Irpan, A., Khansari, M., Kappler, D., Ebert, F., Lynch, C., Levine, S., and Finn, C. (2022). BC-Z: zero-shot task generalization with robotic imitation learning. https://doi.org/10.48550/arXiv.2202.02005.

159. Nakai, T., and Nishimoto, S. (2020). Quantitative models reveal the organization of diverse cognitive functions in the brain. Nat. Commun. *11*, 1142.

160. Yuille, A., and Kersten, D. (2006). Vision as Bayesian inference: analysis by synthesis? Trends Cogn. Sci. *10*, 301–308.

161. Zhan, J., Ince, R.A.A., van Rijsbergen, N., and Schyns, P.G. (2019). Dynamic construction of reduced representations in the brain for perceptual decision behavior. Curr. Biol. *29*. 319e4–326.e4.

162. Yildirim, I., Belledonne, M., Freiwald, W., and Tenenbaum, J. (2020). Efficient inverse graphics in biological face processing. Sci. Adv. *6*, eaax5979.

163. Gosselin, F., and Schyns, P.G. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. Vision Res. *41*, 2261–2271. https://doi.org/10.1016/s0042-6989(01)00097-9.

164. Jaworska, K., Yan, Y., van Rijsbergen, N.J., Ince, R.A.A., and Schyns, P.G. (2022). Different computations over the same inputs produce selective behavior in algorithmic brain networks. eLife *11*. https://doi.org/10.7554/eLife.73651.

165. Ku, S.P., Tolias, A.S., Logothetis, N.K., and Goense, J. (2011). fMRI of the face-processing network in the ventral temporal lobe of awake and anesthetized macaques. Neuron *70*, 352–362.

166. Majaj, N.J., Hong, H., Solomon, E.A., and DiCarlo, J.J. (2015). Simple learned weighted sums of inferior temporal neuronal firing rates accurately predict human core object recognition performance. J. Neurosci. *35*, 13402–13418.

167. Churchland, P.S., Ramachandran, V.S., and Sejnowski, T.J. (1993). A critique of pure vision. In Large-Scale Neuronal Theories of the Brain, C. Koch and J.L. David, eds. (MIT Press), p. 23.

168. Denison, R.N., Carrasco, M., and Heeger, D.J. (2021). A dynamic normalization model of temporal attention. Nat. Hum. Behav. *5*, 1674–1685.

169. Cowley, B.R., Snyder, A.C., Acar, K., Williamson, R.C., Yu, B.M., and Smith, M.A. (2020). Slow drift of neural activity as a signature of impulsivity in macaque visual and prefrontal cortex. Neuron *108*. 551–567.e8.

170. Barack, D.L., and Gold, J.I. (2016). Temporal trade-offs in psychophysics. Curr. Opin. Neurobiol. *37*, 121–125.

171. Dux, P.E., and Marois, R. (2009). The attentional blink: a review of data and theory. Atten. Percept. Psychophys. *71*, 1683–1700.

172. Albers, A.M., Kok, P., Toni, I., Dijkerman, H.C., and de Lange, F.P. (2013). Shared representations for working memory and mental imagery in early visual cortex. Curr. Biol. *23*, 1427–1431.

173. Jerde, T.A., Merriam, E.P., Riggall, A.C., Hedges, J.H., and Curtis, C.E. (2012). Prioritized maps of space in human frontoparietal cortex. J. Neurosci. *32*, 17382–17390.

174. Cole, M.W., Reynolds, J.R., Power, J.D., Repovs, G., Anticevic, A., and Braver, T.S. (2013). Multi-task connectivity reveals flexible hubs for adaptive task control. Nat. Neurosci. *16*, 1348–1355.

175. Driscoll, L., Shenoy, K., and Sussillo, D. (2022). Flexible multitask computation in recurrent networks utilizes shared dynamical motifs. Preprint at bioRxiv. https://doi.org/10.1101/2022.08.15.503870.

176. Yarbus, A.L. (1967). Eye movements during perception of complex objects. In Eye Movements and Vision, A.L. Yarbus, ed. (Springer), pp. 171–211.

177. Rothkopf, C.A., Ballard, D.H., and Hayhoe, M.M. (2007). Task and context determine where you look. J. Vis. *7*, 16.1–16.20.

178. Simoncelli, E.P., and Olshausen, B.A. (2001). Natural image statistics and neural representation. Annu. Rev. Neurosci. *24*, 1193–1216.

179. Josephs, E.L., Zhao, H., and Konkle, T. (2021). The world within reach: an image database of reach-relevant environments. J. Vis. *21*, 14.

180. Sprague, W.W., Cooper, E.A., Reissier, S., Yellapragada, B., and Banks, M.S. (2016). The natural statistics of blur. J. Vis. *16*, 23.

181. Grauman, K., Westbury, A., Byrne, E., Chavis, Z., Furnari, A., Girdhar, R., Hamburger, J., Jiang, H., Liu, M., Liu, X., et al. (2022). Ego4d: around the world in 3,000 hours of egocentric video. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 18973–18990.

182. Braddick, O., and Atkinson, J. (2011). Development of human visual function. Vision Res. *51*, 1588–1609.

183. Teller, D.Y. (1997). First glances: the vision of infants. the Friedenwald lecture. Invest. Ophthalmol. Vis. Sci. *38*, 2183–2203.

184. Karasik, L.B., Tamis-Lemonda, C.S., and Adolph, K.E. (2014). Crawling and walking infants elicit different verbal responses from mothers. Dev. Sci. *17*, 388–395.

185. Cashon, C.H., Ha, O.R., Allen, C.L., and Barna, A.C. (2013). A U-shaped relation between sitting ability and upright face processing in infants. Child Dev. *84*, 802–809.

186. Libertus, K., and Needham, A. (2011). Reaching experience increases face preference in 3-month-old infants. Dev. Sci. *14*, 1355–1364.

187. Fausey, C.M., Jayaraman, S., and Smith, L.B. (2016). From faces to hands: changing visual input in the first two years. Cognition *152*, 101–107.

188. Smith, L.B., and Thelen, E. (2003). Development as a dynamic system. Trends Cogn. Sci. *7*, 343–348.