Trends in
Cognitive Sciences

CellPress

Opinion

# Predictive coding: a more cognitive process than we thought?

Kaitlyn M. Gabhart [1], Yihan (Sophy) Xiong [1], and André M. Bastos [1,*]

**In predictive coding (PC), higher-order brain areas generate predictions that are sent to lower-order sensory areas. Top-down predictions are compared with bottom-up sensory data, and mismatches evoke prediction errors. In PC, the prediction errors are encoded in layer 2/3 pyramidal neurons of sensory cortex that feed forward. The PC model has been tested with multiple recording modalities using the global–local oddball paradigm. Consistent with PC, neuroimaging studies reported prediction error responses in sensory and higher-order areas. However, recent studies of neuronal spiking suggest that genuine prediction errors emerge in prefrontal cortex (PFC). This implies that predictive processing is a more cognitive than sensory-based mechanism – an observation that challenges PC and better aligns with a framework we call predictive routing (PR).**

## Highlights

Predictive coding (PC) models propose that predictions are generated in higher-order areas and feed back to lower-order areas, where they are compared with sensory inputs. Mismatches generate prediction errors.

The local–global oddball paradigm is used to study PC. Local oddballs are formed from repetition-based predictions and do not necessarily imply a predictive code. Global oddballs dissociate stimulus repetition from predictability.

Neuroimaging and intracortical spiking data have been used to investigate the local–global paradigm. Signatures of global oddball processing in sensory cortex are found in neuroimaging but not spiking data.

This provides evidence against PC. We provide a conceptual framework to guide future work, which we call predictive routing (PR).

We apply PR to autism spectrum disorder and schizophrenia.

## Predictive coding models of brain function

Brains have evolved to create mental models that explain the regularities in the environment [1,2]. According to **PC** (see Glossary), these mental models issue internal predictions to drive sensation, thought, and action to reduce energetic costs while maintaining an efficient neuronal code [3–5]. This is achieved via the implementation of an approximation of **Bayesian inference**, which involves combining predictions (priors) with sensory inputs (likelihood) to create a posterior probability (posterior). The posterior is an optimal combination of the prior beliefs with current sensory evidence [6,7]. It serves as the brain's best guess of the state of the world and body [8]. When sensory inputs arrive that do not accord with this internal model, a **prediction error** is generated [5]. Prediction errors are modulated according to their **gain** [4,9,10].

'Classical' PC models [1,4,5,11] propose that predictions are generated in higher-order areas of the brain and feed back down the hierarchy to lower-order areas, where they are compared with sensory inputs. Prediction errors travel in the opposite direction. They feed forward up the hierarchy to update internal models to make better predictions. Feedback predictions are thought to be subtractive. Predictable stimuli are uninformative, so they should drive less overall neuronal activity to save energy. By contrast, surprising/unpredictable stimuli enhance neuronal activity [12–14].

This classical PC model has now been tested using both neurophysiological studies in animals and noninvasive studies as subjects were presented with the **local–global oddball paradigm** (Figure 1A), which is designed to evoke prediction errors at two distinct stages of hierarchical processing. Here, we focus mostly on studies in nonhuman primates and humans using this paradigm, given that primates have a deeper and more distinct cortical hierarchy than rodents [15] and that the concept of hierarchy plays a central role in PC.

Studies recording data using electroencephalography (EEG), magnetoencephalography (MEG), fMRI, and intracranial **local field potentials (LFPs)** largely supported classic PC by showing widespread

[1]Department of Psychology, Vanderbilt University, Nashville, TN, USA

*Correspondence: andre.bastos@vanderbilt.edu (A.M. Bastos).

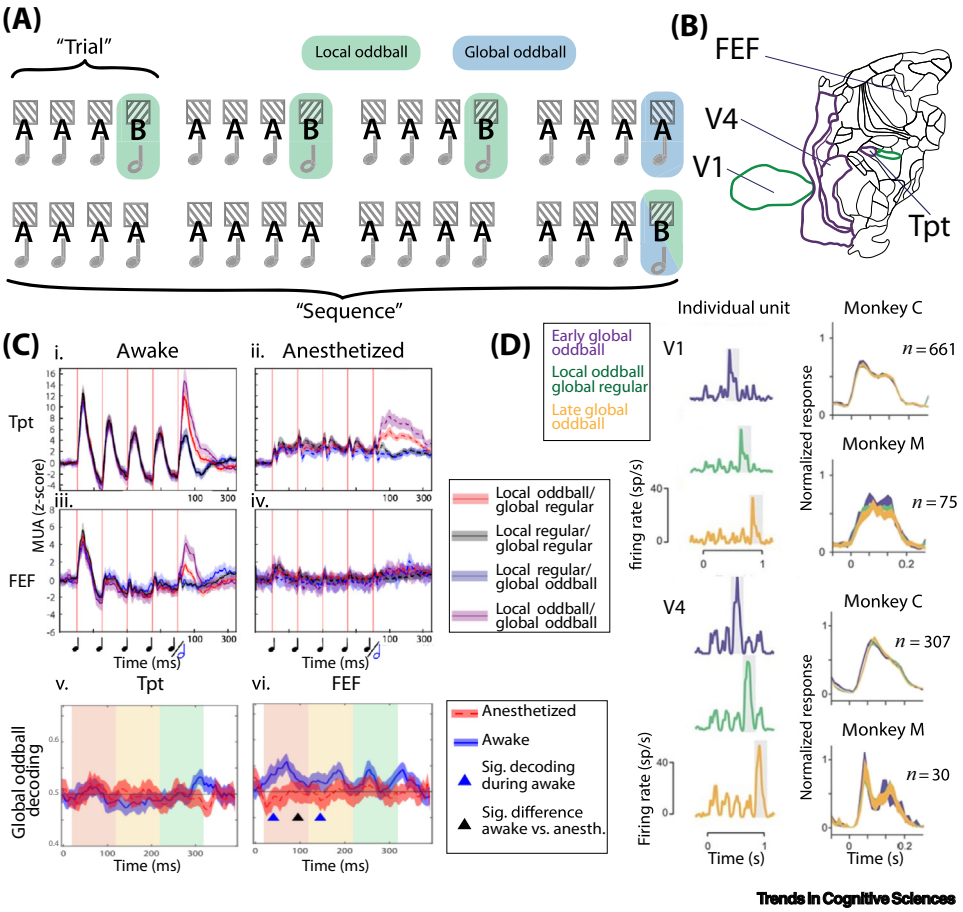


Figure 1. Local and global oddballs and recent evidence of their encoding in neuronal spiking activity. (A) Schematic of the local–global oddball paradigm. Neuronal responses to local oddballs can be caused either by true prediction error signals or by a release from sensory adaptation. Global oddballs are caused by habituation to a stimulus sequence. For example, in the sequence AAAB, AAAB, AAAB, the fourth stimulus is predicted to change. The infrequent stimulus (AAAA) is a global oddball, because the fourth tone is unpredicted despite being a repetition. The timescale of a trial is typically one to a few seconds. The timescale of a sequence is typically tens of seconds. For our review of the literature, we consider studies that have used stimulus repetition-based predictions as local oddballs vs studies that have used more complex designs where prediction and repetition are dissociated as global oddballs. We note that local oddballs that are also global oddballs (AAAB, AAAB, AAAA) also dissociate prediction from adaptation, but only if there is sufficient time between the oddballs to rule out adaptation (see [59]). (B) Area legend on a macaque cortex flat map. (C) A recent study showing the lack of global oddball responses in population spiking [awake Tpt and frontal eye field (FEF) in Ci and Ciii, respectively, and anesthetized Tpt and FEF in Cii and Civ, respectively] [32]. (D) Spiking activity in V1 and V4 [33] is identical to a stimulus when it occurs in the expected position (green traces) as well as when it is unexpectedly early (purple traces) or late (yellow traces) in the sequence.

cortical activation during stimuli that should evoke prediction errors. However, recent evidence from studies measuring **neuronal spiking** in nonhuman primates has shown sparser and higher-order origins for prediction errors, inconsistent with the proposal that feedforward processing represents prediction error. These recent results challenge the classic PC models. Our aim here is to review these studies and suggest an alternative framework for how predictions are implemented: **PR**.

## The local–global oddball paradigm

Initial evidence for prediction error coding came from studies using the **mismatch negativity (MMN)** response [16–18]. These paradigms rely on stimulus repetition to create predictions,

## Glossary

**Autism spectrum disorder (ASD):** a developmental condition characterized by challenges with social communication, limited interests, and repetitive behaviors [92].
**Bayesian inference:** a model that uses known information, termed 'priors', and incoming stimuli to form a prediction about its sensory environment. Priors are used to give probabilities to incoming stimuli, while the incoming stimuli are used to update these probabilities to better represent the current environment.
**Gain:** the brain's ability to determine the importance of prediction errors based on prior knowledge. Prediction errors with low gain will not be signaled to higher-order cortex.
**Local field potential (LFP):** an overall signal of local network activity from a large number of neurons surrounding the recording site, with a spatial resolution of 120–250 mm
**Local–global oddball paradigm:** an auditory oddball paradigm with two levels of regularity used to measure cognitive and attentional capabilities of the brain.
**Mismatch negativity (MMN):** an auditory event-related potential that occurs when a sequence of repetitive auditory stimuli is interrupted by an occasional 'oddball' sound that differs in frequency or duration.
**Neuronal spiking:** electrical impulses generated from single neurons reflecting action potentials that can be recorded using invasive microelectrodes in experimental settings.
**Prediction error:** when environmental sensory inputs arrive that do not accord with the brain's predicted internal model.
**Predictive coding (PC):** a theory of brain function that proposes that the brain is constantly generating and updating a model of its sensory environment. Internal predictions feed back to inform sensory processing, which feeds forward prediction errors.
**Predictive routing (PR):** a theory that hypothesizes a flexible system for predictions that is dependent on the state of learning, context, and conscious state and is implemented via specific cortical layers and neuronal rhythms.
**Schizophrenia (ScZ):** a severe mental illness involving symptoms such as delusions, hallucinations, disorganized speech, troubled thinking, and lack of motivation [92].

and rare deviant stimuli elicit increased neural responses, consistent with a prediction error response. However, the neuronal activity caused by release from adaptation to an unrepeated stimulus versus genuine prediction error is conflated in this experimental design. Neuronal adaptation, a low-level mechanism whereby stimulus repetition causes decreased responses due to neuronal fatigue [19], can explain these responses without invoking hierarchical, Bayesian brain processes [20–22]. To dissociate expectation violation from adaptation, paradigms such as deviance detection [23–25] and the local–global oddball paradigm (Figure 1A; [26]) were developed.

The local–global oddball paradigm uses local oddballs to violate a locally repeated context (e.g., in sequence AAAB, B is the oddball and violates the prediction that A will repeat). Global oddballs violate a pre-habituated pattern [16,27–30]. Global oddballs involve changes in a stimulus within a sequence (in the sequence, AAAB, AAAB, AAAA, the final A is the global oddball). This dissociates a violation of expectation from a release of adaptation.

PC posits that brains construct complex models that are compared with sensory inputs during inference, with mismatches driving prediction error computations. This implies that neurons should actively respond to stimuli, not merely passively adapt. Local oddballs and repetition-based oddballs, which persist under anesthesia [31,32] when top-down connections are functionally inactive (Figure 1C; [32]), provide limited evidence for PC [33]. For PC to explain cortical responses ubiquitously, more complex types of prediction errors (e.g., global oddballs) should evoke responses [1]. Indeed, neuronal model implementations of PC hypothesize that a rare repetition in an environment of frequent alternations (i.e., global or omission oddballs) will elicit prediction error signals at each layer of the network architecture [17,123].

While both local and global oddballs trigger prediction errors, they may do so at distinct hierarchical levels. Basic errors (local oddballs) may be more strongly encoded at earlier levels of the hierarchy, while complex errors (global oddballs) might be more prominent at later levels. However, both types of oddballs should modulate activity in sensory cortical areas for the claim that feedback responses instigate predictive models and that feedforward responses primarily signal prediction error to hold true. If sensory cortex primarily shows passive adaptation (e.g., responds only to local oddballs where prediction error and release from adaptation are conflated), this would weaken the argument for PC as an active, canonical, cortex-wide mechanism.

## fMRI, M/EEG, and LFP studies indicate widespread local and global oddball coding

fMRI [26,34], M/EEG [16–18,35,36], and LFP studies demonstrated widespread representations of local [17,26,37,38] (Figure 2A,C,E, Key figure) and global (Figure 2B,D,F) oddballs, with some studies indicating that even primary and secondary sensory areas (highlighted in green and purple outlines in Figure 2) encode these oddball types [26,31,34,37,39–41]. These studies aligned with the classic PC proposal that sensory cortex compares top-down predictions with bottom-up sensory inputs and issues prediction errors due to their mismatch. They suggested a canonical computation for prediction error [11] and inspired more mechanistic studies of neuronal spiking to investigate the neuronal code associated with prediction error computations.

## Do sensory neurons feed forward prediction errors?

Spiking responses to local oddballs are observed throughout cortex (Figures 1C,D and 2G), from primary sensory areas to higher-order cortex [12,32,33,42–45]. However, studies of neuronal spiking to global oddballs tell a different story (Figure 2H). We recently reported on spike and LFP responses in mid-level auditory cortex area Tpt and from higher-order PFC [the frontal eye fields (FEFs), part of PFC] during the auditory local–global oddball task in macaque monkeys (Figure 1C). Local oddballs were robustly signaled in neuronal spiking both in auditory cortex

## Key figure

Current state of the local and global oddball literature

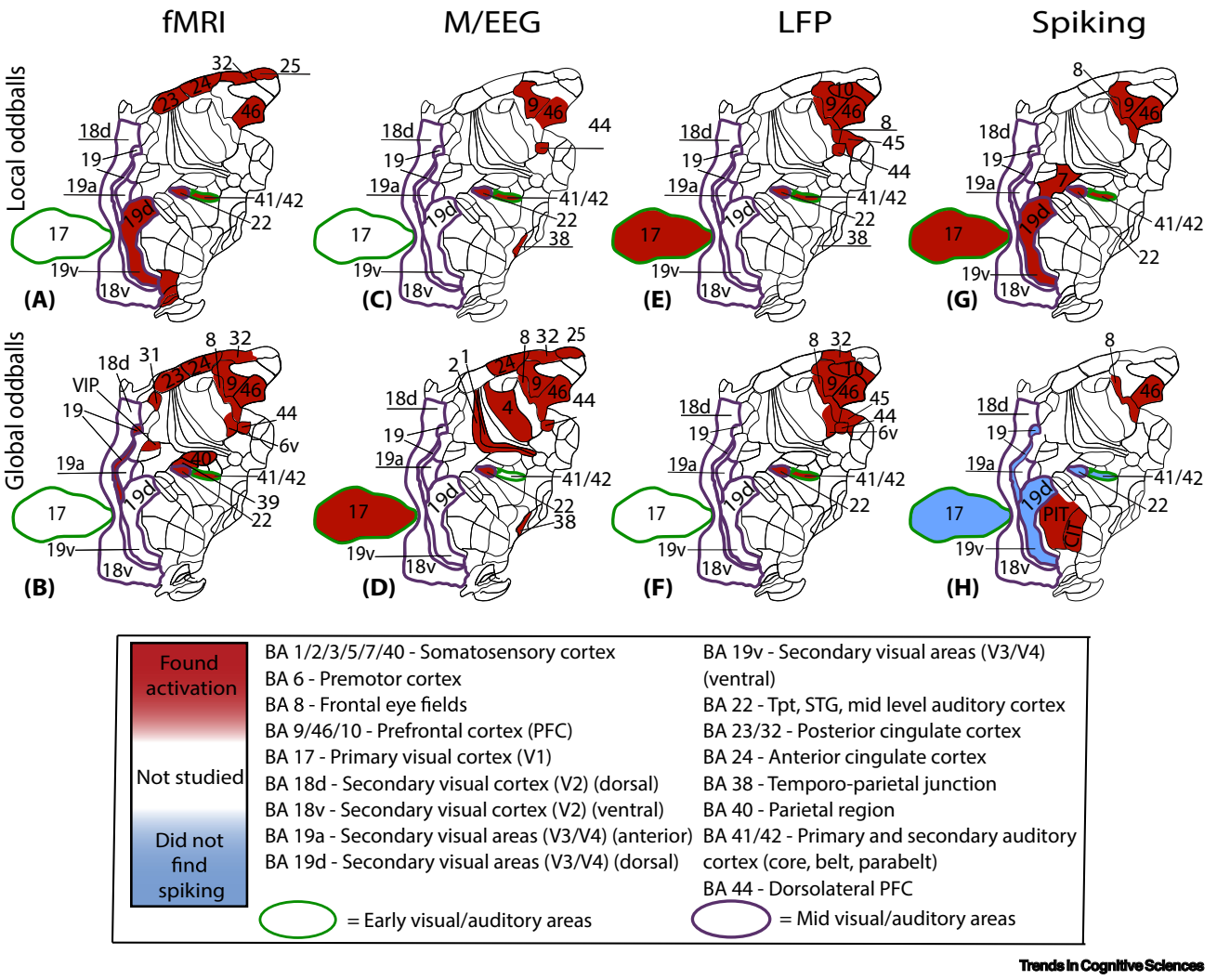


Figure 2. The numbers related in the figure refer to Broadman areas (BAs). The shaded (red vs blue) regions represent activation (or lack of) to local and global. Negative (blue shading) results are only shown for spiking studies. Blue/green borders represented higher- vs lower-order auditory/visual cortices. Flat maps are from macaques [100,101] and human results are depicted in the homologous areas. (A) Local oddball fMRI results [26,34,39]. (B) Global oddball fMRI results [26,34,39,102]. (C) Local oddball magnetoencephalography (MEG)/electroencephalography (EEG) results [26,41,103,104]. (D) Global oddball M/EEG results [26,41,103,104]. (E) Local oddball local field potential (LFP) (note that we consider here also studies employing electrocorticography) results [12,26,31,32,37,39,105]. (F) Global oddball LFP results [26,31,32,37,39]. (G) Local oddball spiking results [12,32,33,42,106–108]. (H) Global oddball spiking results [13,14,32,33,45,48,109]. For further details, see Table S1 in the supplemental information online.

and FEFs, but global oddballs did not register in the population response (Figure 1Ci,iii). Spiking activity to global oddballs was also examined in macaque visual cortex (areas V1 and V4) by varying the list order in which a stimulus would fall relative to a predicted order (Figure 1D). Neurons did not encode prediction errors [33]. Instead, firing rates of neurons in V1 and V4 could be

described by their classic bottom-up properties, including their orientation preference and stimulus repetition, along with enhanced spiking to unrepeated stimuli.

These observations argue against classical PC because they do not show that predictions suppress population activity spiking in sensory areas to save energy during processing of predictable stimuli. Neurons spiked equally to unpredictable and predictable stimuli, after controlling for adaptation. So, if predictions in these studies failed to emerge in spiking studies at the level of V1 and V4 in visual processing (Figure 1D) and in mid-level auditory processing (Figure 1C), is it possible that genuine predictive codes emerge later in the hierarchy than previously thought? A recent study examined this possibility using a visual local–global oddball paradigm and recordings in monkey PFC [45]. Prefrontal neurons spontaneously formed internal models of the task structure, including both local and global oddballs. While global oddballs were decodable from the population response, they did not constitute the primary neuronal representation (i.e., the overall neuronal response in PFC was not enhanced during global oddballs [45]). Only a small fraction of PFC (~2% in [45], 9% in [32]) neurons had a significant response at the individual (spiking) channel level to global oddballs, challenging the idea that prediction errors drive widespread and robust representations.

To determine how widespread population decoding of global oddballs was and whether it was a feature of the conscious state [26,31,40], we examined spike decoding of the stimulus, local oddball, and global oddball status in the auditory cortex (area Tpt) and FEF (a part of the PFC) as we manipulated consciousness with the anesthetic propofol. Notably, global oddball decoding was absent in all states in auditory cortex (Figure 1Cv). In FEF, global oddballs were decodable in the awake state, but decoding fell to chance levels in the unconscious state (Figure 1Cvi; note that spike rate reductions in FEF may have contributed to the lack of decoding). These studies of spiking neurons during local/global predictions (Figures 1C,D and 2H) challenge classic PC models. Genuine predictive codes emerged late in processing, not early as hypothesized. In PFC, neurons flexibly created internal representations of all the sensory and latent task elements, supporting the idea of multidimensional, mixed selectivity in PFC [46,47]. However, these predictions did not result in suppression of overall firing rates, as they were present in only a sparse subspace of neuronal coding. These predictive codes remained undetectable from spiking activity in early to mid-levels of the auditory and visual sensory hierarchies (although genuine predictive codes have been reported in inferotemporal cortex, a late stage of visual sensory processing [13,14,48]).

To summarize, these studies of neuronal spiking diverged in their conclusions from previous fMRI, M/EEG, and LFP data (Figure 2) by demonstrating that feedforward processing in early to mid-levels of sensory cortex does not represent prediction errors, and that predictions did not exert an overall suppressive influence on the population response, as hypothesized in PC. Below we consider, and argue against, the notion that the observed failures in detecting genuine PC in sensory cortex can be attributed to factors such as recording methodology, lack of cell-type specificity, and lack of explicit task engagement. We also consider to what extent the local oddball effects, which are present in spike rate studies in sensory cortex, can be considered a genuine form of PC. We argue that none of these factors can rescue classical PC and that a new perspective, PR (Figure 3), can more parsimoniously explain the current data.

### Recording methodology?

The discrepancy between the fMRI, M/EEG, and LFP versus spikes may be explained by the different nature of the signals recorded across studies. Although fMRI has spatial resolution on the order of a millimeter, it has poor temporal resolution, which means that early (the first ~150 ms of response, reflecting feedforward processing [49]) versus late (after ~150 ms from the onset of a visual stimulus; can reflect both feedforward and feedback influences) response elements are

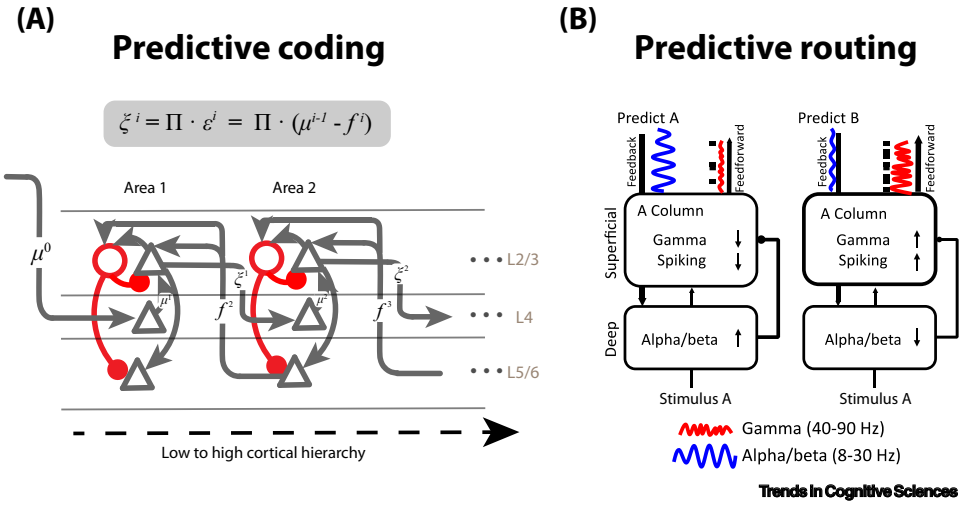


Figure 3. Predictive coding and predictive routing. (A) The classical predictive coding model as realized in a laminar cortical circuit with two levels of hierarchy. Prediction error is computed within a cortical column and then fed forward via superficial-layer (layer 2/3) cells to higher-order cortex. Prediction is fed back via deep-layer (layer 5/6) cells from higher-order area. The proposed computation is given as the shaded equation, where $\xi$ is the precision weighted error, $\varepsilon$ is the error term, and $\Pi$ is the precision ($i$ denotes the level of cortical hierarchy). $\varepsilon$ is calculated with a subtraction between state $\mu$ and top-down prediction $f$, implying a local inhibitory mechanism commonly thought to be realized via superficial-layer inhibitory cells (in red). (B) The predictive routing framework integrates the function of neural oscillations into the predictive processing mechanism in the cortex. Prediction error is carried by gamma, which feeds forward via superficial layers, and prediction is carried by alpha/beta, which feeds back via deep layers. Predictive routing provides the following novel elements: (i) preparatory rhythmic activity, shown as alpha/beta oscillations in the figure (but note that other mechanisms for rhythmic preparation are also possible; e.g., [110]), that (ii) suppresses specific sensory channels (left subpanel: top-down predictions for stimulus A are sent to A selective neurons, to suppress processing of A). This removes the need for explicit prediction error neurons, as prediction 'errors' are the result of sensory inputs to unprepared cortex (right subpanel), and (iii) sparser predictive suppression in sensory cortex (only relevant representations are suppressed). Previous studies are consistent with the idea that prediction error computations emerge from the interaction between cortical rhythms and layers [12,32,37,53,72,111,112].

conflated. As a result, fMRI maps may reflect the processing occurring in each area as well as the top-down inputs to an area. M/EEG and LFP reflect transmembrane currents, which can be the result of the inputs to an area as well as its local computations [50,51]. Only neuronal spiking activity can unambiguously resolve which computations occur where and when in the brain (see Box 1 for more in-depth discussion).

### Box 1. Recording modalities and their relationships: fMRI, EEG/MEG, LFP, and spiking

Extracellular recording of spiking activity represents the action potential output of neurons surrounding the recording contact. By contrast, LFP signal represents the superposition of ionic cellular currents in the brain at the location of the recording [51], with >95% of contributors to the (high-frequency portion) of the LFP from within ~250 μm of the recording electrode [113]. However, fMRI represents the intravascular magnetic susceptibility due to hemodynamic changes and blood oxygenation [50]. Investigations of neurophysiological correlates of the blood oxygenation level-dependent (BOLD) signal have found that BOLD correlates strongly with local field potential signal in the gamma frequency range but very weakly with the spike rate response [50]. This suggests that the BOLD fMRI signal primarily reflects the input of an area instead of the output and local cortical computation, which is better represented in spiking. BOLD signals have also been found to correlate with gamma power particularly in superficial layers of cortex [114]. These are the layers of cortex that receive the primary anatomical feedback connections (via layer 1 [115]) and contain pyramidal neurons that project feedforward outputs (primarily via layers 2 and 3 [101]). This suggests that BOLD signals and gamma (and other frequencies) LFP represent an integration of top-down inputs with local processing, and perhaps an integration of these signals to compute prediction error [116]. In sum, divergence between LFP/fMRI and spiking results can be explained by their respective underlying neural processes, with EEG/LFP/fMRI reflecting composite signals that better represent input to a given area and spiking reflecting the neuronal output and local computation within an area. We note that a similar divergence between EEG/LFP/fMRI results showing extensive top-down modulation in sensory areas but spiking studies showing weak or no top-down modulation has also been encountered in studies of working memory [117] and attention [118].

These considerations suggest that, to understand where and how predictive codes emerge, it is necessary to complement fMRI, M/EEG, and LFP studies with better coverage of neuronal spiking. Traditional single-unit recordings were limited to just one or two areas and a handful of neurons at a time, limiting this method's field of view. This makes it easy to potentially miss a sparse signal such as global oddball encoding which may engage a small proportion of neurons. To mitigate this, we and others have recently worked to expand the scope of brain areas and neurons that can be measured using multi-area, high-density, laminar neurophysiology (MaDeLaNe [52–54]; Box 2). The emerging evidence from studies employing MaDeLaNe is largely consistent with Figures 1 and 2 and shows that genuine predictive codes emerged at much later cortical processing stages than hypothesized by classical PC [54].

### Cell-type specificity?

Although difficult to study in primates with current methods (but see [55–57]), cell-specific circuits during local oddballs have been investigated in mice. These studies often utilized passive tasks and repetition to establish a predictable stimulus [42,58]. For example, deviance detection (a prediction error signal using repetition to establish prediction) is abolished in mouse V1 when top-down inputs from PFC to V1 are suppressed [42]. Two specific cell types may be involved in this type of predictive processing. Pharmacogenetic blockade of somatostatin$^+$ (SOM$^+$) cells in mouse V1 eliminated deviance detection. Also in mouse V1, a study found that vasointestinal peptide-positive interneurons (VIP$^+$) increased their activity to repeated (predictable) stimuli, suggesting that VIP interneurons play a role in signaling predictions [58]. In addition, chemogenetic blockade of these neurons disrupted deviance detection [58] and VIP$^+$ interneurons signaled unexpected omissions [59]. It is therefore possible that genuine PC does emerge within specific neuronal populations. These signals, contrary to the hypotheses of classical PC, appear insufficient to drive robust prediction error responses that could evoke feedforward processing. They failed to elicit a significant population response to global oddballs in the sensory areas [54,59], largely consistent with the studies in monkeys (Figures 1 and 2).

### Explicit task engagement?

Finally, it is important to consider task and behavioral context. In PC, prediction errors are modulated by their gain ([4,9], but see [5], where prediction errors are assumed to be an automatic feature of neuronal activity). Gain can be parameterized as a form of attention [9]. The primate studies that reported failures in detecting global oddballs in sensory areas [32,33] used paradigms where oddballs were presented passively, and therefore the gain on prediction errors may have been low. However, attentional modulation is unlikely to account for the lack of global oddballs in these studies. First, even without an explicit task, PFC neurons spontaneously formed inner models of the

#### Box 2. The importance of MaDeLaNe recording techniques

Classical PC models have been difficult to test experimentally because the model makes use of multiple stages of cortical processing, with distinct cell types, cortical layers, areas, and directions of feedforward/feedback processing contributing to the hypothesized computation. Rigorous testing of PC therefore requires methodologies that can capture each of these dimensions. A method is needed that delivers simultaneously good spatial (neuronal specific) resolution, temporal resolution (to determine when the computation happens), together with sufficient coverage to track the evolution of the cortical responses across the hierarchy (a comparison of currently used methods is shown in Figure I). We and others have employed multiple high-density electrodes (Figure I), combined with novel analytical tools [53] to gain layer information during predictive processing in multiple cortical areas [44,54,119,120]. Methods such as calcium imaging, optogenetics, and high-density recording can be used in conjunction to distinguish neuronal cell types during predictive processing [54,58,59,63]. Taken together, such methods are necessary to make further progress in understanding PC and PR, to understand the contributions of distinct layers, cortical cell types, and processing stages [54]. We propose to further refine and improve MaDeLaNe methodology to further gain spatial coverage along with cellular specificity, and to utilize this method in diverse species (including at a minimum both rodents and primates; e.g., [54]) and, where deemed clinically appropriate, in the human brain (e.g., [121,122]) to uncover the circuitry underlying distinctly human forms of prediction [112].

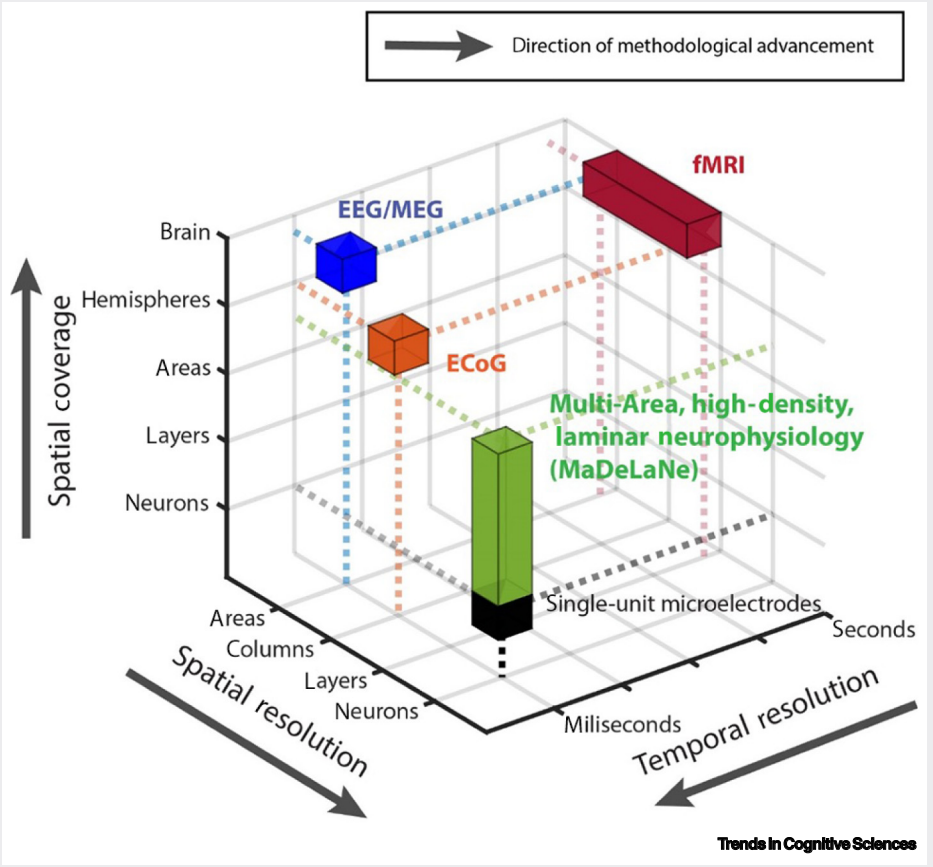


Figure I. Multi-area, high-density, laminar neurophysiology (MaDeLaNe) sampling compared with fMRI, electroencephalography (EEG)/magnetoencephalography (MEG), and local field potential (LFP). Comparison of commonly used neuroimaging and neuronal recording methods in neuroscientific research and their respective spatial coverage, spatial resolution, and temporal resolution. Traditional methods that offer brain-wide sampling (M/EEG and fMRI) suffer from either poor spatial or temporal resolution. Electrocorticography (ECoG) (which we have grouped together with studies measuring LFPs from intracortical electrodes) retains high spatiotemporal resolution with large-scale coverage but does not sufficiently resolve layers and neurons in the spatial resolution axis [64]. Traditional invasive recordings using single-unit microelectrodes have poor spatial coverage (typically, a few neurons at a time in one or two areas). MaDeLaNe methods retain high spatiotemporal resolution (for detecting single neurons) while also sampling densely across layers and in multiple areas [44,54,120]. Note that the goal of future methodological development is greater spatial coverage, and finer spatial and temporal resolution to obtain neuronal and cell-type specificity (illustrated here as towards the reader in the projected 3D diagram).

task structure [32,45], including global context/deviance, which was eliminated during unconsciousness [32]. This implies that some level of awareness/attention was present in the passive local–global oddball tasks. Second, selective attention in V1, V2, and V4 cortex modulates the neuronal response by ~5–23% [60]. It is unlikely that this level of modulation would create a robust global oddball response despite its absence in the reported data. Nevertheless, future studies should carefully control the state of selective attention during oddball processing to explicitly study the neuronal mechanisms of gain modulation of prediction errors [10,61].

Sensory cortical neurons encoded local oddballs, so they may be described as encoding a local prediction that the current features of the sensory environment will persist (and repeat). One might

be tempted to call this a predictive code. Indeed, this locally predictive code based on adaptation at different timescales is sufficient to compute a probability distribution [62]. However, although local oddballs might be modulated by top-down processes, it is worth remembering that these repetition-based local predictions persist in deep anesthesia (Figure 1Ci) during which frontal cortex becomes inactive (Figure 1Civ) [31,32]. Local oddball responses therefore do not require the active feedback machinery envisioned by PC to pass high-level predictions down the hierarchy.

## What functions can top-down predictive inputs serve in sensory cortex if they do not drive spiking?

Based on the current evidence, we believe that genuine prediction errors are computed in higher-order cortex and do feedback to sensory areas but do not drive a population spiking response (e.g., no prediction errors detectable in the mean spike rates) in those areas. What might be the function of this feedback if it does not drive a suprathreshold response? We speculate that this top-down feedback may be sparse and selective, such that a population response is not observed in lower-order cortex but that some neurons do receive preparatory feedback that matters for a small population, perhaps especially involving subpopulations of inhibitory interneurons [59,63] (but see [54]). A second possibility is that top-down feedback exists via extraclassical mechanisms that do not drive a population spike response, but still modulates sensory areas via mechanisms such as subthreshold oscillatory coupling [12,32,64–66], ephaptic coupling [67,68], or dendritic mechanisms [69]. A third possibility is that top-down feedback is used to drive the motor system for behaviorally relevant predictions [63], which many of the current paradigms do not address. Table 1 summarizes the neuronal activity observed during local and global oddballs.

## How to reconcile these observations? Predictive routing

The findings reviewed here on global oddball responses suggest that genuine predictions emerge only at high levels of the hierarchy, inconsistent with classical PC (Figure 3A). We argue that the existing evidence can be well accommodated within an updated model, PR (Figure 3B) [12] (see also [70]). The PR framework builds on several properties of neuronal oscillations. Gamma frequency (40–90 Hz) increases in power with sensory stimuli and is positively correlated with spiking activity [71,72]. Gamma/spiking is anticorrelated with alpha/beta frequency (8–30 Hz) oscillations, which strengthen in PFC (and its coherence with other cortical areas) during top-down tasks, such as selective attention [73,74] and when stimuli are predictable [12,37] (for reviews of the top-down effects of beta during cognition, see [72,75,76]). This suggests that alpha/beta may

Table 1. What activity do local and global oddballs elicit? A chart summarizing activity elicited from local and global oddballs as demonstrated in the current literature

| | Local oddballs | Global oddballs |
|---|---|---|
| Hierarchical emergence | Widespread (Figure 2A,C,E,G) | Higher-order areas in spiking (Figure 2H) |
| Cortical layer | Superficial layers dominant [12,42,58,86] | Agranular layers (activation outside L4) [54] |
| Frequency of neuronal response | Increased gamma; decreased alpha/beta [12,32,37,39,111] | Decreased alpha/beta [39,111] |
| Attentional dependence (gain) | Not dependent [41] | Dependent [26,41] |
| Cell type | All cell types that are released from adaptation [54,58] | Specific inhibitory cell types [29,59,63], but see [54] |
| Consciousness dependent | Yes in higher-order cortex, no for lower-order cortex [31,32,40] | Yes [31,32,40,102] |

act as an executive control mechanism by turning up or down the amount of gamma/spiking needed for a task [72].

PR involves learning-driven formation of dynamic ensembles [77,78] in higher-order areas. This involves mixed selectivity neurons, which have been frequently observed in PFC [45–47]. Neurons with this property can form complex predictions in real time based on any combination of inputs, rather than respond based on static receptive field characteristics. Once a predictive ensemble is formed among these mixed-selectivity neurons, feedback connections can transmit these signals from higher-order to lower-order cortical regions utilizing alpha/beta rhythms [12,37,64,65,79–81] and prepare sensory areas for stimulus processing. Alpha/beta rhythms can have an inhibitory effect on neuronal spiking and gamma [12,82–85] at specific phases [82,84], leading to a state of relative inhibition or 'preparation' in predicted pathways [12]. Unpredicted stimuli arriving in unprepared cortical areas elicit enhanced spiking and gamma-frequency (40–90 Hz) oscillations, engaging enhanced feedforward communication (Figure 3B) [12,31,37,39,64,65,86].

The key difference between PC and PR is that in PR, there are no explicit prediction error neurons [12]. By contrast, in PC there is dedicated canonical circuitry in cortical layers 2 and 3 for prediction error computation (Figure 3A). In PR, higher-order cortex issues selective preparatory signals that suppress sensory processing. A prediction 'error' is a result of inputs arriving at an unprepared cortex (Figure 3B). Higher-order cortex can issue these predictions in a sparse and selective manner, such that predictive suppression targets only relevant representations. This is compatible with the idea of redundant coding in sensory cortex, when neurons in early to mid-levels of the hierarchy fire away to even the most highly predictable stimuli and are not modulated by global oddballs (Figure 1C,D). Higher-order predictions (e.g., global oddballs) depend on temporal integration across longer timescales at the level of the full sequence (Figure 1A; typically, several seconds). Neurons in early sensory cortex have fast timescales of temporal integration [87] and would be ill equipped to receive predictive suppression for sequences with long timescales. PR proposes that complex predictions are formed in PFC and selectively suppress sensory areas with longer time constants, such as high-level visual areas [88]. In PR, predictions are a higher-order, more selective, and sparser signal than in PC, and are implemented via spectrolaminar mechanisms rather than dedicated error circuits.

We recently tested PR using propofol-mediated unconsciousness during the auditory local–global oddball paradigm [32]. Propofol essentially inactivated PFC (Figure 1Civ) while sparing bottom-up sensory drive to auditory cortex (Figure 1Cii). Alpha/beta band power modulation was also eliminated with propofol. Under these conditions of no top-down input from PFC, we presented local oddballs and recorded neuronal activity in sensory cortex. Paradoxically, we found that, during unconsciousness, local oddball-related gamma increased [32] relative to the awake state. We interpret this as evidence for PR: without beta-band activity (which normally increases during processing of predicted stimuli) and without top-down inputs from PFC, sensory cortex became disinhibited and generated more oddball-related gamma (along with temporally exaggerated spiking to oddballs; Figure 1Cii) compared with the conscious state.

## Implications for clinical disorders

The proposed PR model (with its emphasis on selective suppression of specific sensory areas) offers a valuable framework for understanding the neural mechanisms underlying sensory processing and PC, in particular for **schizophrenia (ScZ)** and **autism spectrum disorder (ASD)**. A key neural deficit in ScZ is reductions in MMN signals and reductions in sensory-induced gamma oscillations [89–91], which implies that there may be profound deficits in bottom-up sensory

processing and predictive processing [92]. Reduced bottom-up sensory processing in ScZ may lead to weakened top-down predictive models and disrupted top-down beta [89,93]. Without accurate input from the sensory environment, the brain is left to generate its own internal model of the world, which can lead to false predictions due to the lack of reliable sensory input (i.e., hallucinations and delusions). This feedforward sensory processing via dampened gamma may be the result of decreased functionality of parvalbumin interneurons [94,95], which contribute to gamma oscillations [96] and are decreased in ScZ in specific areas and layers [95,97].

By contrast, ASD symptomology presents as abnormally increased prediction errors [92]. A recent study recording electrocorticography from nonhuman primates with an induced form of ASD showed abnormally high sensory responses to local oddballs in the local–global oddball task [98]. Additionally, individuals with ASD often struggle with shifting attention between different stimuli or tasks [99]. This impairment could affect their ability to process global oddballs, as it requires flexible attentional allocation and the ability to update mental models.

Given the evidence, we propose that sensory cortex primarily employs a redundant code, ensuring robust representation of sensory input. By contrast, higher-order cortical areas may rely more on a predictive code, utilizing top-down predictions to efficiently process information. Imbalances in this process may underlie the distinct clinical presentations of ASD and ScZ. We propose that the local–global oddball task and its variants be systematically applied to studies of ScZ and ASD (and other clinical disorders) as functional assays for sensory and higher-order cognition. The use of these tasks in animal models of these clinical disorders is another exciting avenue, which could generate biomarkers for states of disorder and provide objective markers to drive therapeutics targeted to specific types of processing (e.g., local/global oddballs).

## Concluding remarks

The literature on local–global oddball processing presents a challenge to PC models: M/EEG, fMRI, and LFP data indicated that both local and global oddballs modulate activity in both sensory and higher-order cortex. However, studies of neuronal spiking in primates have failed to find robust global oddballs in lower- and mid-level sensory areas. Predictive codes emerged in PFC but only in a sparse subspace of neuronal encoding. Therefore, predictions may not be as broadly suppressive or as canonical as hypothesized. From the perspective of PR, global oddballs are an emergent feature of higher-order cortex neurons displaying flexible, mixed selectivity. Predictions are sent to the appropriate level of processing. Early to mid-sensory processing may be largely immune from the effects of prediction. More sophisticated task designs and higher-density neuronal recordings will provide further insights into the cortical circuitry for prediction. This may lead to a better understanding of clinical disorders that depend on intact predictions (see Outstanding questions).

## Outstanding questions

Do active tasks (and attention) enhance the gain of prediction error responses to global oddballs in sensory cortex?

Is there a generic canonical microcircuit for prediction error that operates with similar computations for local, global, omission, and other types of violation?

How can new high-density, cell-type-specific methods be used to learn which areas and specific cell types are involved in predictions in the primate brain?

How can the mechanisms of the PR model be causally tested? Does rhythmic alpha/beta activity exert specific suppression via distinct interneuron subtypes?

Do distinct species represent predictions at different stages of cortical processing and with distinct mechanisms?

How can PC/PR-derived experimental paradigms be used to address bottom-up vs top-down theories of clinical disorders such as ScZ? Does feedforward sensory processing via dampened gamma result from decreased functionality of parvalbumin interneurons?

### Acknowledgments

This research was supported by the National Institute of Mental Health (NIMH) R00MH116100, Vanderbilt University startup funds, a Vanderbilt Brain Institute Faculty Fellow Award, the NARSAD Young Investigator Award from the Brain and Behavior Research Foundation, and the National Science Foundation (NSF) Faculty Early Career Development Program (CAREER) grant 2339210. We thank Hamed Nejat, Eli Sennesh, Jacob Westerberg, and Alex Maier for helpful discussions.

### Declaration of interests

No interests are declared.

### Supplemental information

Supplemental information associated with this article can be found online at https://doi.org/10.1016/j.tics.2025.01.012.

## References


1. Friston, K. (2005) A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 815–836
2. von Helmholtz, H. (1948) Concerning the perceptions in general, 1867. In *Readings in the history of psychology. Century psychology series*, pp. 214–230, Appleton-Century-Crofts
3. Srinivasan, M.V. *et al.* (1997) Predictive coding: a fresh view of inhibition in the retina. *Proc. Biol. Sci.* 216, 427–459
4. Friston, K. (2009) The free-energy principle: a rough guide to the brain? *Trends Cogn. Sci.* 13, 293–301
5. Rao, R.P.N. and Ballard, D.H. (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87
6. Hohwy, J. (2012) Attention and conscious perception in the hypothesis testing brain. *Front. Psychol.* 96,
7. Vilares, I. and Kording, K. (2011) Bayesian models: the structure of the world, uncertainty, behavior, and the brain. *Ann. N. Y. Acad. Sci.* 1224, 22–39
8. Seth, A.K. and Friston, K.J. (2016) Active interoceptive inference and the emotional brain. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371, 20160007
9. Feldman, H. and Friston, K.J. (2010) Attention, uncertainty, and free-energy. *Front. Hum. Neurosci.* 4, 215
10. Kok, P. *et al.* (2012) Attention reverses the effect of prediction in silencing sensory signals. *Cereb. Cortex* 22, 2197–2206
11. Bastos, A.M. *et al.* (2012) Canonical microcircuits for predictive coding. *Neuron* 76, 695–711
12. Bastos, A.M. *et al.* (2020) Layer and rhythm specificity for predictive routing. *Proc. Natl. Acad. Sci. U. S. A.* 117, 31459–31469
13. Issa, E.B. *et al.* (2018) Neural dynamics at successive stages of the ventral visual stream are consistent with hierarchical error signals. *eLife* 7, e42870
14. Schwiedrzik, C.M. and Freiwald, W.A. (2017) High-level prediction signals in a low-level area of the macaque face-processing hierarchy. *Neuron* 96, 89–97.e4
15. Glatigny, M. *et al.* (2024) Comparison of the cortical hierarchy between macaque monkeys and mice based on cell-type specific microcircuits. *bioRxiv*, Published online October 31, 2024. https://doi.org/10.1101/2024.10.31.621258
16. Garrido, M.I. *et al.* (2009) The mismatch negativity: a review of underlying mechanisms. *Clin. Neurophysiol.* 120, 453–463
17. Wacongne, C. *et al.* (2012) A neuronal model of predictive coding accounting for the mismatch negativity. *J. Neurosci.* 32, 3665–3678
18. Picton, T.W. (1992) The P300 wave of the human event-related potential. *J. Clin. Neurophysiol.* 9, 456–479
19. Grill-Spector, K. *et al.* (2006) Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci.* 10, 14–23
20. Taaseh, N. *et al.* (2011) Stimulus-specific adaptation and deviance detection in the rat auditory cortex. *PLoS One* 6, e23369
21. Nir, Y. *et al.* (2015) Auditory responses and stimulus-specific adaptation in rat auditory cortex are preserved across NREM and REM sleep. *Cereb. Cortex* 25, 1362–1378
22. Ayala, Y.A. and Malmierca, M.S. (2013) Stimulus-specific adaptation and deviance detection in the inferior colliculus. *Front. Neural Circuits* 6, 89
23. Harms, L. *et al.* (2018) Late deviance detection in rats is reduced, while early deviance detection is augmented by the NMDA receptor antagonist MK-801. *Schizophr. Res.* 191, 43–50
24. Gallimore, C.G. *et al.* (2023) Spatiotemporal dynamics across visual cortical laminae support a predictive coding framework for interpreting mismatch responses. *Cereb. Cortex* 33, 9417–9428
25. Lakatos, P. *et al.* (2020) The thalamocortical circuit of auditory mismatch negativity. *Biol. Psychiatry* 87, 770–780
26. Bekinschtein, T.A. *et al.* (2009) Neural signature of the conscious processing of auditory regularities. *Proc. Natl. Acad. Sci.* 106, 1672–1677
27. Vuust, P. *et al.* (2012) The sound of music: differentiating musicians using a fast, musical multi-feature mismatch negativity paradigm. *Neuropsychologia* 50, 1432–1443
28. Tervaniemi, M. *et al.* (1994) Neural representations of abstract stimulus features in the human brain as reflected by the mismatch negativity. *Neuroreport* 5, 844–846
29. Garrett, M. *et al.* (2020) Experience shapes activity dynamics and stimulus coding of VIP inhibitory cells. *eLife* 9, e50340
30. Awwad, B. *et al.* (2023) Extensive representation of sensory deviance in the responses to auditory gaps in unanesthetized rats. *Curr. Biol.* 33, 3024–3030.e3
31. Nourski, K.V. *et al.* (2018) Auditory predictive coding across awareness states under anesthesia: an intracranial electrophysiology study. *J. Neurosci.* 38, 8441–8452
32. Xiong, Y.S. *et al.* (2024) Propofol-mediated loss of consciousness disrupts predictive routing and local field phase modulation of neural activity. *Proc. Natl. Acad. Sci. U. S. A.* 121, e2315160121
33. Solomon, S.S. *et al.* (2021) Limited evidence for sensory prediction error responses in visual cortex of macaques and humans. *Cereb. Cortex* 31, 3136–3152
34. Uhrig, L. *et al.* (2014) A hierarchy of responses to auditory regularities in the macaque brain. *J. Neurosci.* 34, 1127–1132
35. Charuthamrong, P. *et al.* (2022) Automatic speech discrimination assessment methods based on event-related potentials (ERP). *Sensors* 22, 2702
36. Sauer, A. *et al.* (2023) Spectral and phase-coherence correlates of impaired auditory mismatch negativity (MMN) in schizophrenia: a MEG study. *Schizophr. Res.* 261, 60–71
37. Chao, Z.C. *et al.* (2018) Large-scale cortical networks for hierarchical prediction and prediction error in the primate brain. *Neuron* 100, 1252–1266.e3
38. Strauss, M. *et al.* (2015) Disruption of hierarchical predictive coding during sleep. *Proc. Natl. Acad. Sci. U. S. A.* 112, E1353–E1362
39. Jiang, Y. *et al.* (2022) Constructing the hierarchy of predictive auditory sequences in the marmoset brain. *eLife* 11, e74653
40. Uhrig, L. *et al.* (2016) Cerebral responses to local and global auditory novelty under general anesthesia. *Neuroimage* 141, 326–340
41. Wacongne, C. *et al.* (2011) Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proc. Natl. Acad. Sci. U. S. A.* 108, 20754–20759
42. Hamm, J.P. *et al.* (2021) Cortical ensembles selective for context. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2026179118
43. Keller, G.B. *et al.* (2021) Sensorimotor mismatch signals in primary visual cortex of the behaving mouse. *Neuron* 74, 809–815
44. Siegle, J.H. *et al.* (2021) Survey of spiking in the mouse visual system reveals functional hierarchy. *Nature* 592, 86–92
45. Bellet, M.E. *et al.* (2024) Spontaneously emerging internal models of visual sequences combine abstract and event-specific information in the prefrontal cortex. *Cell Rep.* 43, 113952
46. Rigotti, M. *et al.* (2013) The importance of mixed selectivity in complex cognitive tasks. *Nature* 497, 585–590
47. Mante, V. *et al.* (2013) Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* 503, 78–84
48. Meyer, T. and Olson, C.R. (2011) Statistical learning of visual transitions in monkey inferotemporal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 108, 19401–19406
49. Lamme, V.A. and Roelfsema, P.R. (2000) The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci.* 23, 571–579
50. Logothetis, N.K. *et al.* (2001) Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412, 150–157
51. Buzsáki, G. *et al.* (2012) The origin of extracellular fields and currents – EEG, ECoG, LFP and spikes. *Nat. Rev. Neurosci.* 13, 407–420
52. Jun, J.J. *et al.* (2017) Fully integrated silicon probes for high-density recording of neural activity. *Nature* 551, 232–236
53. Mendoza-Halliday, D. *et al.* (2024) A ubiquitous spectrolaminar motif of local field potential power across the primate cortex. *Nat. Neurosci.* 27, 547–560
54. Westerberg, J.A. *et al.* (2024) Stimulus history, not expectation, drives sensory prediction errors in mammalian cortex. *bioRxiv*, Published online October 3, 2024. https://doi.org/10.1101/2024.10.02.616378

55. Dimidschstein, J. *et al.* (2016) A viral strategy for targeting and manipulating interneurons across vertebrate species. *Nat. Neurosci.* 19, 1743–1749
56. Krienen, F.M. *et al.* (2020) Innovations present in the primate interneuron repertoire. *Nature* 586, 262–269
57. Mehta, P. *et al.* (2019) Functional access to neuron subclasses in rodent and primate forebrain. *Cell Rep.* 26, 2818–2832.e8
58. Bastos, G. *et al.* (2023) Top-down input modulates visual context processing through an interneuron-specific circuit. *Cell Rep.* 42, 113133
59. Jamali, S. *et al.* (2024) Parallel mechanisms signal a hierarchy of sequence structure violations in the auditory cortex. *bioRxiv*, Published online August 22, 2024. https://doi.org/10.1101/2024.08.21.609026
60. Buffalo, E.A. *et al.* (2010) A backward progression of attentional effects in the ventral stream. *Proc. Natl. Acad. Sci. U. S. A.* 107, 361–365
61. Auksztulewicz, R. *et al.* (2017) Task relevance modulates the behavioural and neural effects of sensory predictions. *PLoS Biol.* 15, e2003143
62. Tring, E. *et al.* (2023) A power law describes the magnitude of adaptation in neural populations of primary visual cortex. *Nat. Commun.* 14, 8366
63. Furutachi, S. *et al.* (2024) Cooperative thalamocortical circuit mechanism for sensory prediction errors. *Nature* 633, 398–406
64. Bastos, A.M. *et al.* (2015) Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron* 85, 390–401
65. van Kerkoerle, T. *et al.* (2014) Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proc. Natl. Acad. Sci. U. S. A.* 111, 14332–14341
66. Carlson, B.M. *et al.* (2024) V1 population spiking does not show statistically significant differences between laminar compartments during binocular rivalry flash suppression. In *Proceedings of Neuroscience,* PSTR281.02/E19-V1, Chicago, IL, USA
67. Lee, S.Y. *et al.* (2024) Cell-class-specific electric field entrainment of neural activity. *Neuron* 112, 2614–2630.e5
68. Pinotsis, D.A. and Miller, E.K. (2023) *In vivo* ephaptic coupling allows memory network formation. *Cereb. Cortex* 33, 9877–9895
69. Mikulasch, F.A. *et al.* (2023) Where is the error? Hierarchical predictive coding through dendritic error computation. *Trends Neurosci.* 46, 45–59
70. Arnal, L.H. and Giraud, A.-L. (2012) Cortical oscillations and sensory predictions. *Trends Cogn. Sci.* 16, 390–398
71. Lundqvist, M. *et al.* (2016) Gamma and beta bursts underlie working memory. *Neuron* 90, 152–164
72. Miller, E.K. *et al.* (2018) Working memory 2.0. *Neuron* 100, 463–475
73. Westerberg, J.A. *et al.* (2021) Pop-out search instigates beta-gated feature selectivity enhancement across V4 layers. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2103702118
74. Buschman, T.J. and Miller, E.K. (2007) Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315, 1860–1862
75. Lundqvist, M. *et al.* (2024) Beta: bursts of cognition. *Trends Cogn. Sci.* 28, 662–676
76. Bressler, S.L. and Richter, C.G. (2015) Interareal oscillatory synchronization in top-down neocortical processing. *Curr. Opin. Neurobiol.* 31, 62–66
77. Buschman, T.J. *et al.* (2012) Synchronous oscillatory neural ensembles for rules in the prefrontal cortex. *Neuron* 76, 838–846
78. Antzoulatos, E.G. and Miller, E.K. (2016) Synchronous beta rhythms of frontoparietal networks support only behaviorally relevant representations. *eLife* 5, e17822
79. Turner, W. *et al.* (2023) Visual information is predictively encoded in occipital alpha/low-beta oscillations. *J. Neurosci.* 43, 5537–5545
80. Michalareas, G. *et al.* (2016) Alpha-beta and gamma rhythms subserve feedback and feedforward influences among human visual cortical areas. *Neuron* 89, 384–397
81. Vezoli, J. *et al.* (2021) Brain rhythms define distinct interaction networks with differential dependence on anatomy. *Neuron* 109, 3862–3878.e5
82. Haegens, S. *et al.* (2011) α-Oscillations in the monkey sensorimotor network influence discrimination performance by rhythmical inhibition of neuronal spiking. *Proc. Natl. Acad. Sci. U. S. A.* 108, 19377–19382
83. Lundqvist, M. *et al.* (2020) Preservation and changes in oscillatory dynamics across the cortical hierarchy. *J. Cogn. Neurosci.* 32, 2024–2035
84. Spaak, E. *et al.* (2012) Layer-specific entrainment of gamma-band neural activity by the alpha rhythm in monkey visual cortex. *Curr. Biol.* 22, 2313–2318
85. Bastos, A.M. *et al.* (2018) Laminar recordings in frontal cortex suggest distinct layers for maintenance and control of working memory. *Proc. Natl. Acad. Sci. U. S. A.* 115, 1117–1122
86. English, G. *et al.* (2023) Bayesian surprise shapes neural responses in somatosensory cortical circuits. *Cell Rep.* 42, 112009
87. Murray, J.D. *et al.* (2014) A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci.* 17, 12
88. Mendoza-Halliday, D. *et al.* (2014) Sharp emergence of feature-selective sustained activity along the dorsal visual pathway. *Nat. Neurosci.* 17, 9
89. Uhlhaas, P.J. and Singer, W. (2010) Abnormal neural oscillations and synchrony in schizophrenia. *Nat. Rev. Neurosci.* 11, 100–113
90. Uhlhaas, P.J. and Singer, W. (2012) Neuronal dynamics and neuropsychiatric disorders: toward a translational paradigm for dysfunctional large-scale networks. *Neuron* 75, 963–980
91. Grent-'t-Jong, T. *et al.* (2016) MEG-measured visually induced gamma-band oscillations in chronic schizophrenia: evidence for impaired generation of rhythmic activity in ventral stream regions. *Schizophr. Res.* 176, 177–185
92. Tarasi, L. *et al.* (2022) Predictive waves in the autism-schizophrenia continuum: a novel biobehavioral model. *Neurosci. Biobehav. Rev.* 132, 1–22
93. Liddle, E.B. *et al.* (2016) Abnormal salience signaling in schizophrenia: the role of integrative beta oscillations. *Hum. Brain Mapp.* 37, 1361–1374
94. Dienel, S.J. *et al.* (2023) The nature of prefrontal cortical GABA neuron alterations in schizophrenia: markedly lower somatostatin and parvalbumin gene expression without missing neurons. *Am. J. Psychiatry* 180, 495–507
95. Zhang, Z.J. and Reynolds, G.P. (2002) A selective decrease in the relative density of parvalbumin-immunoreactive neurons in the hippocampus in schizophrenia. *Schizophr. Res.* 55, 1–10
96. Antonoudiou, Y.L. *et al.* (2020) Parvalbumin and somatostatin interneurons contribute to the generation of hippocampal gamma oscillations. *J. Neurosci.* 40, 7668–7687
97. Lewis, D.A. *et al.* (2012) Cortical parvalbumin interneurons and cognitive dysfunction in schizophrenia. *Trends Neurosci.* 35, 57–67
98. Chao, Z.C. *et al.* (2024) Erroneous predictive coding across brain hierarchies in a non-human primate model of autism spectrum disorder. *Commun. Biol.* 7, 851
99. Chen, Y.-Y. *et al.* (2022) Excessive functional coupling with less variability between salience and default mode networks in autism spectrum disorder. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 7, 876–884
100. Felleman, D.J. and Van Essen, D.C. (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47
101. Markov, N.T. *et al.* (2014) A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cereb. Cortex* 24, 17–36
102. Tasserie, J. *et al.* (2022) Deep brain stimulation of the thalamus restores signatures of consciousness in a nonhuman primate model. *Sci. Adv.* 8, eabl5547
103. Chennu, S. *et al.* (2013) Expectation and attention in hierarchical auditory prediction. *J. Neurosci.* 33, 11194–11205
104. Schofield, T.M. *et al.* (2009) Changing meaning causes coupling changes within higher levels of the cortical hierarchy. *Proc. Natl. Acad. Sci. U. S. A.* 106, 11765–11770
105. Cornella, M. *et al.* (2012) Detection of simple and pattern regularity violations occurs at different levels of the auditory hierarchy. *PLoS One* 7, e43604
106. Hayden, D.J. *et al.* (2021) Visual recognition is heralded by shifts in local field potential oscillations and inhibitory networks in primary visual cortex. *J. Neurosci.* 41, 6257–6272

107. Chen, G. *et al.* (2017) Distinct inhibitory circuits orchestrate cortical beta and gamma band oscillations. *Neuron* 96, 1403–1418.e6
108. Gong, Y. *et al.* (2024) Neural correlates of novelty detection in the primary auditory cortex of behaving monkeys. *Cell Rep.* 43, 113864
109. Auksztulewicz, R. *et al.* (2022) Omission responses in field potentials but not spikes in rat auditory cortex. *bioRxiv*, Published online February 11, 2022. https://doi.org/10.1101/2022.02.11.479668
110. Uran, C. *et al.* (2022) Predictive coding of natural images by V1 firing rates and rhythmic synchronization. *Neuron* 110, 1240–1257.e8
111. Chao, Z.C. *et al.* (2022) A quantitative model reveals a frequency ordering of prediction and prediction-error signals in the human brain. *Commun. Biol.* 5, 1076
112. Arnal, L.H. *et al.* (2011) Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nat. Neurosci.* 14, 797–801
113. Katzner, S. *et al.* (2009) Local origin of field potentials in visual cortex. *Neuron* 61, 35–41
114. Scheeringa, R. *et al.* (2016) The relationship between oscillatory EEG activity and the laminar-specific BOLD signal. *Proc. Natl. Acad. Sci. U. S. A.* 113, 6761–6766
115. Rockland, K.S. and Pandya, D.N. (1797) Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey. *Brain Res.* 179, 3–20
116. Theriault, J.E. *et al.* (2023) A functional account of stimulation-based aerobic glycolysis and its role in interpreting BOLD signal intensity increases in neuroimaging experiments. *Neurosci. Biobehav. Rev.* 153, 105373
117. Leavitt, M.L. *et al.* (2017) Sustained activity encoding working memories: not fully distributed. *Trends Neurosci.* 40, 328–346
118. Boynton, G.M. (2011) Spikes, BOLD, attention, and awareness: a comparison of electrophysiological and fMRI signals in V1. *J. Vis.* 11, 12
119. Liu, Y. *et al.* (2024) A high-density 1,024-channel probe for brain-wide recordings in non-human primates. *Nat. Neurosci.* 27, 1620–1631
120. Findling, C. *et al.* (2024) Brain-wide representations of prior information in mouse decision-making. *bioRxiv*, Published online November 18, 2024. https://doi.org/10.1101/2023.07.04.547684
121. Paulk, A.C. *et al.* (2022) Large-scale neural recordings with single neuron resolution using Neuropixels probes in human cortex. *Nat. Neurosci.* 25, 252–263
122. Xie, W. *et al.* (2024) Neuronal sequences in population bursts encode information in human cortex. *Nature* 635, 935–942
123. Kiebel, S.J. *et al.* (2008) A hierarchy of time-scales and the brain. *PLoS Comput. Biol.* 4, e1000209