The Perceptual Prediction Paradox

Clare Press[1]*, Peter Kok[2], and Daniel Yon[1,3]

[1] Department of Psychological Sciences, Birkbeck, University of London

[2] Wellcome Centre for Human Neuroimaging, UCL

[3] Department of Psychology, Goldsmiths, University of London

* Correspondence: c.press@bbk.ac.uk (C. Press)

Abstract

From the noisy information bombarding our senses our brains must construct percepts that are *veridical* – reflecting the true state of the world – and *informative* – conveying what we did not already know. Influential theories suggest that both challenges are met through mechanisms that use expectations about the likely state of the world to shape perception. However, current models explaining how expectations render perception either veridical or informative are mutually incompatible. While the former propose that perceptual experiences are dominated by events we expect, the latter propose that perception of expected events is suppressed. To solve this paradox we propose a two-process model in which probabilistic knowledge initially biases perception towards what is likely, and subsequently upweights events that are particularly surprising.

The Computational Challenge of Perception

Our sensory receptors are constantly bombarded with enormous quantities of information that change rapidly across space and time [1]. This information is noisy, partly due to imperfect signal transmission in the brain but also due to indeterminacy in the signals that reach our receptors. For example, parts of an object we need to grasp may be visually occluded or the sound of the metro may mask our conversational partner. The brain must make sense of this noisy information rapidly if it is to be useful – it is not worthwhile to generate perceptual representations of a cup slipping from our fingers if it is already broken on the floor, or to register that an acquaintance has smiled when passing us if they are halfway down the street.

Our sensory systems face two principal challenges. First, they must construct percepts that are *veridical* – reflecting what the world is actually like. A largely accurate model of the extracranial world – including the external environment and our own and others' bodies – is often important for effective behaviour. Second, the perceptual representations must be *informative* – telling the organism what it needs to know for updating its models and beliefs, i.e., what it did not already know. Theories from psychology and neuroscience propose that we achieve these aims by using expectations about what we are likely to encounter to shape our perceptual experiences. However, current models explaining how expectations render perception either veridical or informative are mutually incompatible. While the former propose that our perceptual experiences are dominated by events we expect, the latter propose that perception of expected events is suppressed. In the present article we first outline this conflict, before considering how ideas from research on learning and inference may suggest a two-process model that could resolve the paradox – enabling us to use our expectations to optimise both veridical and informative perception. More broadly, we argue it is essential that models from learning and perception begin to inform each other systematically and outline ways in which they could be coordinated.

Current Theoretical Accounts of Perceptual Optimisation

Bayesian Theories

In recent years researchers have been particularly interested in how we use our expectations (see Box 1) to generate veridical perceptual experiences rapidly in an inherently ambiguous sensory world. According to Bayesian theories, we increase the accuracy of our perceptual representations by biasing them in line with our prior expectations [2,3]. For example, if we are currently at work, the figure waving at us in a shady corridor is more likely to be a colleague than an old classmate from school. Biasing perception in line with these statistical likelihoods will render it more veridical on average, even if such a mechanism may lead to occasional misperceptions when these regularities are disrupted. In other words, while it may seem counterintuitive that biases make perception more accurate, this will in fact often be true given high sensory noise and indeterminacy. If we simply perceived the input without bringing to bear our prior expectations, we may not perceive the correct words through a noisy radio transmission or road signs at dusk, and perception would often be too slow to be useful. This optimisation may be realised through mechanisms that alter the weights on sensory channels – effectively turning up the volume (or 'gain') on expected relative to unexpected inputs ([2,4,5]; Figure 1, Key Figure).

These Bayesian theories are supported by psychophysical findings that we are biased to report the presence of predicted over unpredicted events [6,7], are more sensitive to their presence [8–10], and perceive them to have higher clarity [11]. For example, concave faces are frequently misperceived to have the more typical convex structure ([12,13]; see Figure 1) and we report ambiguously-coloured fruit in line with its typical colour [14]. Aberrations in these illusions are observed across certain clinical conditions – e.g., schizophrenia and autism – and interpreted as evidence for atypical sensory prediction [12,15–17]. At the neural level, concurring evidence demonstrates higher fidelity representations of expected sensory events

[5,18] and sensory representations biased towards expected inputs [19] – with neural biases predicting the perceptual biases [13].

Cancellation Theories

These Bayesian theories provide a plausible account of how we may generate veridical representations of our world rapidly, but in fact they conflict with 'Cancellation' models (also referred to as 'Dampening' theories) that have been popular across cognitive sciences for the last few decades [20–22]. These models outline how – given the limited capacity of our sensory systems – we focus on the most informative perceptual information, prioritising unexpected sensory inputs that signal the need for belief updating and new courses of action. This function is achieved by suppressing (not facilitating) processing of expected inputs. For example, when reaching out to grasp a cup, reduced processing of the predicted sensation of the cup touching our fingers enables us to focus on unexpected events – like the cup slipping. This prioritisation enables rapid updating of models – and new courses of action where appropriate – when the unexpected occurs. These theories are especially prominent in the action control literature (although are also found in the wider sensory cognition literature [22–24]), focusing on the benefit of cancelling out predictable self-generated inputs, and thereby optimising detection of potentially crucial externally-generated (unexpected) signals [25]. Such a cancellation mechanism is thought to explain why we cannot tickle ourselves ([20]; Figure 1).

The theories have drawn wide support from studies reporting that predictable tactile, auditory and visual inputs evoke lower sensory cortical activation [20,22,23,26–28] and are perceived less intensely [21,29–31, see also 32] than unexpected inputs. The theories are particularly popular in computational neuropsychiatry where aberrant cancellation mechanisms are thought to generate atypicalities in the sense of agency in delusional populations [28,33] – by making the (predictable) outcomes of actions appear unusually intense.

5

A Perceptual Prediction Paradox?

While Bayesian and Cancellation theories in isolation appear adaptive ways to optimise veridical and informative perception, respectively, they cannot both be true. When it comes to the contents of perception, monolithic Bayesian theories that suggest perception is dominated by what we expect conflict with monolithic Cancellation theories making the opposite suggestion. This conflict extends to adaptive arguments that motivate both kinds of theories – what we expect to be there is *more* likely, but also *less* informative. As such, any mechanism that uniformly biases what we perceive towards or away from our expectations forfeits one adaptive advantage to secure another. Most work reporting support for Bayesian or Cancellation theories simply ignores the paradox, only outlining reasoning why predicted information may be perceptually facilitated or attenuated, respectively [e.g., 2,3,25]. However, given the conflict, these models can only be partially complete and it is necessary to consider how the paradox can be resolved.

The first possibility to consider is that there is in fact no paradox – Bayesian and Cancellation mechanisms both operate, but in different domains. For example, it may be adaptive for Bayesian mechanisms to optimise veridical perception of the wider environment using sensory context and for Cancellation mechanisms to optimise informative perception during action using sensorimotor predictions. Indeed, while Cancellation models are found in the broad sensory cognition literature [22–24] they predominate in action disciplines. However, we believe the adaptive arguments and recent empirical results do not support domain-specific interpretations.

First, the Bayesian adaptive arguments – rapid generation of veridical perception from noisy inputs – would appear just as useful during action. For example, if we are drinking a cup of coffee in a dark kitchen before sunrise, we will generate more veridical estimates of our ongoing actions if we are biased to perceive the sight of a moving hand approaching a cup. Similarly, a Cancellation mechanism that highlights the unexpected in a limited capacity

system would prove useful regardless of whether we predict the contents of perception on the basis of action or not. If we open our front door to a tiger lying on our sofa, we need to update our beliefs (and act).

Second, the theoretical distinction between sensorimotor and other predictions may not actually be supported by empirical discrepancies. The paradigms and analyses typical in the two fields differ. Specifically, Bayesian accounts frequently consider evidence of event detection and quality of neural representation [e.g., 5,7], whereas Cancellation accounts are typically supported by reports of perceived intensity and quantity of neural activation [e.g., 20,21]. Use of paradigms and analyses popular in the normative sensory cognition literature can yield similar support for Bayesian models in action [18,34–36]. There are also findings from the wider sensory cognition literature that may be incompatible with a Bayesian account, e.g., cancelled neural responses for predicted visual sequences in lateral occipital cortex [22,23].

We believe that some of the apparent differences between fields have arisen due to distinctions in other processes that do operate differently during action [37]. Most notably, there are mechanisms that attenuate perception during action but do not differentially influence expected and unexpected events. They are therefore dissociable from prediction mechanisms (Box 1). For example, when we move, we suppress *all* tactile input to a moving effector – regardless of whether or not it is a predicted outcome of action [29,38–40] – perhaps due to spinal gating mechanisms ([41]; see also 'active inference' theories [42–44]). Studies cited in support of Cancellation theories frequently compare either the processing of predictable events presented during action against events when passive, or processing where the action and sensory events overlap less due to temporal misalignment. Generalised suppression mechanisms may therefore contribute to effects found in studies using these comparison conditions. We thus contend that to isolate the functional influence of predictive mechanisms on perception of action outcomes, empirical efforts should directly compare predicted and unpredicted events always in the presence (or absence) of action [34,45].

To sum, it seems an unlikely solution for Cancellation logic to prevail in action disciplines and Bayesian logic elsewhere. Well-controlled experiments must further interrogate any differences in the action domain, but we assume here that *both* veridical *and* informative perception will be required in any perceptual domain – even if there turn out to be nuances in relative weighting of the two requirements in the action domain [see also 46].

A Theoretical Resolution: The Opposing Process Theory

Given that we believe current Bayesian and Cancellation logic is mutually incompatible, here we consider a possible resolution that allows the adaptive functions of both models to co-exist. We propose that insights can be gained by considering reasoning and findings from the learning and inference literature. Learning theorists are concerned with understanding how we update our models of the world rather than the contents of perception *per se*. This field has made great progress in establishing how we establish causal relationships between events (model uncertainty, e.g. Does being in the arctic predict the sight of polar bears? [47,48]). This type of inference is different from perceptual inferences, which relate instead to how we establish what is present in our environment right now (sensory uncertainty, e.g. Am I looking at a polar bear? see [48]). However, it is undisputed that learning and perception are related. As already outlined, our current learnt models about environmental contingencies inform perceptual processing, and we will only update our learnt models to the extent that we perceive information that contradicts them [49].

Learning models frequently focus on the concept of 'surprise', developing formal descriptions of the extent to which an event is unexpected. Computational models often operationalise surprise as the 'Kullback Leibler Divergence' (KLD) – a quantity that captures the change between beliefs about environmental states before (priors) and after (posteriors) the sensory evidence in question has been processed. Importantly, surprise operationalised in this way reflects both the deviation between the modal value of the prior and posterior, and the

'precision' (inversely related to variance) of these representations. In other words, it conceptualises the extent of overlap between prior and posterior distributions (see Figure 1). When surprise is high (or overlap is low), the organism should learn [50]. KLD thus captures the intuition that you cannot be very surprised when you did not have strong expectations in the first place, nor can you be surprised by an event unless you are sure it actually occurred.

The learning and inference literature reports processes that operate upon the receipt of surprising information that would influence one's propensity to learn about it. These processes would also influence perception. For example, a number of learning studies demonstrate phasic catecholamine release shortly after the presentation of surprising events, which is thought to mediate learning by relatively increasing the gain of sensory inputs [47,51–54]. These findings are relevant to understanding perception because higher gain on sensory inputs would be hypothesised to generate higher intensity percepts [42,55]. It has also been demonstrated [50,54] that we saccade towards events generating surprise (high KLD) – perhaps facilitated by phasic catecholamine release [52,54]. Again, this work is of relevance to the perception community because foveating surprising events will increase visual processing of them. Due to crossmodal spatial representations, such saccades may also increase perceptual processing in other sensory modalities [56]. These surprise-driven mechanisms may thus alter perception by highlighting unexpected, informative inputs, thereby aiding learning and updating of models. When surprising input is detected they allow 'double checking' of sensory evidence or increase vigilance to further change to support adaptive model updating.

We therefore propose that a two-process model could resolve the paradox concerning how to optimise perception so that it is both veridical and informative. As outlined in Bayesian accounts, perception is first biased towards prior knowledge. This aids the rapid generation of largely accurate perceptual experiences. This process may be operational from the point at which predictions can be made, that is, often before an event is presented, through pre-activation of expected units [1,57,58]. However, when an event generates high surprise – as

would be the case only for highly 'unexpected' events (see below) – later processes are triggered (after event presentation) which highlight these events. Engaging these processes only for events that are sufficiently surprising will highlight unexpected inputs but only insofar as they are likely informative for model updating. As such, the rationale motivating Cancellation models – that focusing on unexpected inputs makes perception more informative – applies only when the unexpected inputs are likely informative rather than the result of noise. In short, we broadly perceive what we expect unless we have good reason not to, in which case we then divert extra perceptual processing resources to the unexpected to determine whether it is appropriate to update our models. This account therefore allows the adaptive functions of both Bayesian and Cancellation models to co-exist, but shifts the 'Cancellation' mechanism to a later process engaged only by some unexpected stimuli.

Empirical Support

There is some evidence supporting the hypothesis that opposing perceptual processes operate on different timescales. One psychophysical study demonstrates that while expected events are perceived with greater intensity 50 ms after presentation, this bias reverses by 200 ms such that unexpected events are perceived with greater intensity [45], and an electroencephalography (EEG) study suggests neural switches may operate at similar timescales [59]. The classically reported expectation suppression effects in the EEG and magnetocephalography (MEG) signal – perhaps more in line with Cancellation accounts – are also most commonly found ~150 ms after event presentation or later [30,60–63, see also 64]. Furthermore, an interesting EEG study in 12-month-old infants [65] found early enhanced perceptual processing of expected events, followed by a later switch to favouring unexpected events. Note that the timescale of this switch was longer than in the adult studies, in line with frequent reports of temporal shifts in developmental work.

Empirical support for this account can also be gleaned from differences in the types of paradigms in which evidence for Bayesian and Cancellation theories is typically reported.

Behavioural paradigms demonstrating facilitated perception of the expected often present low intensity stimulation in signal detection paradigms, where we hypothesise this later opposing process would not operate [e.g., 7]. That is, 'unexpected' inputs in these studies have low sensory precision and any surprise is correspondingly low (Figure 2). Conversely, paradigms reporting cancellation often present unexpected events that will generate high surprise and therefore where we predict the later process would operate. For example, in the action literature researchers often measure perceived intensity of suprathreshold stimulation rather than detection of noisy events [e.g., 21,29]. Under situations of strong (or intermediate) surprise and time insensitive methods (e.g., fMRI) it is less clear which process should dominate and therefore more difficult to form hypotheses about the influence of expectation on perceptual processing, so in principle it could also explain null findings under these conditions [66].

Future Directions

We propose that future work should focus on temporally sensitive experiments to address this theory directly. It would be interesting to apply the model to a range of related disciplines currently employing conflicting Bayesian and Cancellation models (Box 2) – e.g., those examining language processing, pain perception and social cognition, and ask whether it can account for the complex perceptual symptomatology present in a range of psychiatric disorders.

Additionally, it will be essential to determine the level of overlap and interaction between expectation-based and attention-based processes when examining these mechanisms. Predictive mechanisms – operating on the basis of probabilities – may appear conceptually similar to attention mechanisms given that many classic attentional manipulations are probabilistic, e.g., the Posner cueing task [67] and oddball paradigms [68]. However, recent work aiming to differentiate these concepts considers attentional mechanisms to be only those that highlight task-relevant input [4,5,7,18,66,69–70]. Dissociating mechanisms in this way

may be key, given that many apparent effects of probability could be the result of confounded task-relevance – events that are likely are also often important for task performance [7]. Nevertheless, there may sometimes be good reason to define attention in a different way (e.g., labelling effects of task-irrelevant probabilities on neural gain as 'attentional' [71]), which may often render attention and expectation mechanisms conceptually and empirically indistinguishable. Regardless of how one draws the distinction, under our theory mechanisms generating upweighting of the expected and unexpected rely on probabilistic information, not information about task–relevance, but empirical efforts must examine whether the operation of these mechanisms is altered when task relevance is manipulated.

We have focussed on how specific insights from learning research could resolve the perceptual paradox generated by Bayesian and Cancellation reasoning. However, the learning literature is much broader than that outlined above and there are a multitude of bidirectional insights to be gained between communities. Currently researchers who study perception frequently assume that expectations reflect conditional probabilities present in the environment, but a more accurate characterisation must account for how we learn and represent these probabilities. For example, debate has raged around which cues we learn about – those that precede predictable versus uncertain information [72–74] – and neurochemical processes have been characterised that may operate differently according to whether uncertainty itself was expected [47]. The boundary conditions are difficult to establish concerning the expectedness of uncertainty [47], but if a distinction proves crucial for these learning mechanisms it is likely to prove crucial for perception as well. For example, we have focussed on opposing processes that operate on the timescale of hundreds of milliseconds within a single event, but sensory processing may also change on a longer timescale according to whether we believe we are in an environment where the cue-outcome mappings are stable or volatile [47,75]. Conversely, our Opposing Process theory makes specific predictions about processes that will bias perception in specific ways. Learning models will benefit from considering these kinds of influence on perception, because it will be our

perception of events – rather than their objective properties – that determines whether and how we update models.

If the Opposing Process theory turns out to be broadly correct, it will also be important to test a precise model of exactly how these processes operate. Predictive coding schemes [76–78] of cortical functioning have become popular in recent years and our Opposing Process theory could certainly be implemented using such a scheme. Under predictive coding assumptions, the brain contains distinct units representing the 'best guess' about the outside world (hypothesis units) and the discrepancy between these guesses and incoming sensory evidence (error units). Perception unfolds as error units pass information up the cortical hierarchy and hypotheses are refined to reflect the external world with greater fidelity. It is frequently suggested that the contents of perception reflect activity across the hypothesis population [10]. If this is the case, our theory would predict that hypothesis units are initially more strongly weighted towards what we expect – i.e., 'representational sharpening' [77]. However, when agents encounter events that elicit surprise, increased gain on surprising inputs leads to high fidelity representations of unexpected events across hypothesis units – perhaps with even greater fidelity than when events unfold as we thought they would. Therefore, we suggest that future work establishes precisely how to fit the Opposing Process theory into predictive coding frameworks.

Finally, insights into the role of expectations in shaping our perceptual beliefs will likely be gained by separating out the relative influences of decisional, primary perceptual, and memory-based processes on perceptual decisions [79–82]. Importantly, most psychophysical experiments reported in this article are unable to dissociate these processes, and when we discuss 'perception' we cannot know whether it is primary perceptual processes that are influenced, or whether it is subsequent memory- or response-based representations of that information. For instance, does our expectation that a face has a convex structure lead us to perceive convexity at the specific time when the ambiguous stimulus is presented, remember it as being convex, or simply report it as being so – because we believe it should have been?

The paradox associated with the need to generate beliefs that are both veridical and informative holds regardless of the specific stage in processing. Certainly decisional processes likely contribute to many of these observed effects [83] and it will be necessary to dissociate the influences empirically, as well as determine the functional role played by the different processes.

Concluding Remarks

Bayesian and Cancellation models in the cognitive sciences suggest that observers use probabilistic knowledge to optimise perception in a rapidly changing and ambiguous world. Though appealing, monolithic Bayesian and Cancellation models cannot both be true, yet they are both apparently supported by large bodies of literature. Here we have outlined a theory that could solve this paradox, with separate mechanisms using the same probabilistic knowledge to bias perception initially towards what is likely, and subsequently to upweight events that elicit high 'surprise'. Future experiments should crucially examine the timecourse of processing, as well as consider whether 'unexpected' events in fact elicit high surprise (see Outstanding Questions). This account therefore allows the adaptive functions of both Bayesian and Cancellation models to co-exist and articulates necessary experiments to determine how top-down knowledge shapes perception of the world around us in health and disease.

Box 1. The Nature of a Prediction

This article considers a sensory prediction to be a sensory state expected by the organism on the basis of environmental regularities. Concurring with predictive coding frameworks, we assume that predictions have a mean and variance. When moving our index finger towards a piano key we might have a highly 'precise' tactile prediction, such that we expect stimulation only on the index finger. However, depending upon our ability, our auditory prediction may be less precise, expecting probably a G, but possibly an F or A. These predictions will typically be learnt through domain-general processes based upon contingencies and contiguities between sensory states and any predictive information [49] – e.g., actions performed [84] and other sensory states [2,3]. We assume a cognitive definition of prediction – with effects mediated by representations of what is probable, even if these probabilities are represented by subpersonal mechanisms (i.e., not necessarily consciously accessible).

We also assume that it is theoretically possible to have no predictions at all, such that the prior distribution is flat. Adults will rarely exhibit flat priors – even if listening to the avant-garde composer John Cage for the first time we will hold expectations on the basis of other composers and likely establish the new set of regularities swiftly. Nevertheless, when comparing perception of expected and unexpected events it is frequently queried where to place a hypothetical 'neutral' (flat) baseline. For example, when poorer processing of expected relative to unexpected events is found, researchers often consider whether the processing of the expected has been enhanced or the unexpected suppressed [22,85]. We concur that this is an interesting discussion point but is likely to represent a quantitative rather than qualitative distinction. If one holds an imprecise flat prior then all sensory events are equally unexpected, but will be less surprising than events that are wrongly predicted. Whether events yield sufficient surprise to trigger the opposing process will not require simply knowing the shape of the prior distribution but also the shape of the posterior distribution, and therefore there may not be a qualitative distinction between processing in wrongly predicted and unpredicted

scenarios. Considering the level of surprise (KLD) evoked by sensory events may prove a fruitful line of enquiry for drawing hypotheses about these differences.

Box 2. Wider Applications

This article has focussed on the literature from low-level visual cognition and action cognition, but the Opposing Process model has the potential also to resolve debates and inconsistencies in related disciplines. For example, the conflict between Bayesian and Cancellation theories is also debated by those examining speech perception, with neural and behavioural findings in line with both accounts ([11,86,87]; Figure I). Similar arguments are raised by those aiming to understand how we perceive and interpret the actions of others – when we have information about their likely underlying intentions [88, see also 89] – and how expectations influence pain [90] and threat [91] perception. The reasoning presented in the present article would apply to any domain aiming to understand how expectations influence perception. It may therefore inform some of these current controversies, as well as related debates in the memory literature concerning whether expected or unexpected events should be remembered with greater accuracy [92,93].

It would be particularly interesting to determine whether the Opposing Process theory could also provide a more accurate description of pathological processing in psychiatric and neurological disorders. For example, recent predictive coding models propose that unusual experiences seen in psychosis could arise because some kinds of predictions (e.g., about the self) are weakened and others (e.g., about the external world) strengthened [15]. It may prove crucial to consider how altering the balance between different kinds of predictions interacts with our suggestions concerning when expected events will be up- or down-weighted. It may also help to explain detail-focussed perception [94], atypical learning [75,95] and atypical action [96] in autism, all of which have been proposed to relate to atypical sensory prediction, and associations between neurochemical and cognitive change in Parkinson's disease [97].

Furthermore, it would be interesting to establish whether current debates concerning the influence of event repetition on perceptual processing could be informed with such opposing process assumptions [cf. 63,98]. It has been disputed whether repetition of events should facilitate or attenuate the neural response, as well as whether it should lead to priming or habituation in perception [99,100]. An Opposing Process model may shed light on this dispute. Specifically, if a repeated event typically reflects an expected scenario, there may be initial biases towards perceiving repetitions followed by later upweighting of surprising non-repeat events. Testing this theory would require using time resolved methods in repetition paradigms, as well as varying the expectedness of repetition.

References

1   Press, C. and Yon, D. (2019) Perceptual prediction: Rapidly making sense of a noisy world. *Curr. Biol.* 29, R751–R753

2   de Lange, F.P. *et al.* (2018) How do expectations shape perception? *Trends Cogn. Sci.* 22, 764–779

3   Kaiser, D. *et al.* (2019) Object vision in a structured world. *Trends Cogn. Sci.* 23, 672–685

4   Summerfield, C. and de Lange, F.P. (2014) Expectation in perceptual decision making: Neural and computational mechanisms. *Nat. Rev. Neurosci.* 15, 745–756

5   Kok, P. *et al.* (2012) Less is more: Expectation sharpens representations in the primary visual cortex. *Neuron* 75, 265–270

6   Chalk, M. *et al.* (2010) Rapidly learned stimulus expectations alter perception of motion. *J. Vis.* 10, 2–2

7   Wyart, V. *et al.* (2012) Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. *Proc. Natl. Acad. Sci. U. S. A.* 109, 3593–3598

8   Pinto, Y. *et al.* (2015) Expectations accelerate entry of visual stimuli into awareness. *J. Vis.* 15, 13

9   Stein, T. and Peelen, M.V. (2015) Content-specific expectations enhance stimulus detectability by increasing perceptual sensitivity. *J. Exp. Psychol. Gen.* 144, 1089–1104

10   Teufel, C. *et al.* (2018) Prior object-knowledge sharpens properties of early visual feature-detectors. *Sci. Rep.* 8, 10853

11   Sohoglu, E. *et al.* (2014) Top-down influences of written text on perceived clarity of degraded speech. *J. Exp. Psychol. Hum. Percept. Perform.* 40, 186–199

12   Palmer, C.J. *et al.* (2017) Bayesian approaches to autism: Towards volatility, action, and behavior. *Psychol. Bull.* 143, 521–542

13   Kuhn, G. (2019) *Experiencing the impossible: The science of magic*. Cambridge, MA: MIT Press.

14   Vandenbroucke, A.R.E. *et al.* (2016) Prior knowledge about objects determines neural color representation in human visual cortex. *Cereb. Cortex* 26, 1401–1408

15   Corlett, P.R. *et al.* (2019) Hallucinations and strong priors. *Trends Cogn. Sci.* 23, 114–127

16   Keane, B.P. *et al.* (2016) Seeing more clearly through psychosis: Depth inversion illusions are normal in bipolar disorder but reduced in schizophrenia. *Schizophr. Res.* 176, 485–492

17   Powers, A.R. *et al.* (2017) Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors. *Science* 357, 596–600

18   Yon, D. *et al.* (2018) Action sharpens sensory representations of expected outcomes. *Nat. Commun.* 9, 4288

19   Kok, P. *et al.* (2013) Prior expectations bias sensory representations in visual cortex. *J. Neurosci.* 33, 16275–16284

20   Blakemore, S.J. *et al.* (1998) Central cancellation of self-produced tickle sensation. *Nat. Neurosci.* 1, 635–640

21   Kilteni, K. and Ehrsson, H.H. (2017) Body ownership determines the attenuation of self-generated tactile sensations. *Proc. Natl. Acad. Sci.* 114, 8426–8431

22   Richter, D. *et al.* (2018) Suppressed sensory response to predictable object stimuli throughout the ventral visual stream. *J. Neurosci.* 38, 7452–7461

23   Meyer, T. and Olson, C.R. (2011) Statistical learning of visual transitions in monkey inferotemporal cortex. *Proc. Natl. Acad. Sci.* 108, 19401–19406

24   Kumar, S. *et al.* (2017) Encoding of predictable and unpredictable stimuli by inferior temporal cortical neurons. *J. Cogn. Neurosci.* 29, 1445–1454

25   Enikolopov, A.G. *et al.* (2018) Internally generated predictions enhance neural and behavioral detection of sensory stimuli in an electric fish. *Neuron* 99, 135-146.e3

26   Kikuchi, T. *et al.* (2019) Neural responses to action contingency error in different cortical areas are attributable to forward prediction or sensory processing. *Sci. Rep.* 9, 9847

27   Limanowski, J. *et al.* (2018) Different responses of the right superior temporal sulcus to visual movement feedback during self-generated vs. externally generated hand movements. *Eur. J. Neurosci.* 47, 314–320

28   Shergill, S.S. *et al.* (2014) Functional magnetic resonance imaging of impaired sensory prediction in schizophrenia. *JAMA Psychiatry* 71, 28–35

29   Bays, P.M. *et al.* (2006) Attenuation of self-generated tactile sensations is predictive, not postdictive. *PLoS Biol.* 4, e28

30   Hughes, G. *et al.* (2013) Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychol. Bull.* 139, 133–151

31   Wolpe, N. *et al.* (2016) Ageing increases reliance on sensorimotor prediction through structural and functional differences in frontostriatal circuits. *Nat. Commun.* 7, 13034

32   Müsseler, J. and Hommel, B. (1997) Blindness to response-compatible stimuli. *J. Exp. Psychol. Hum. Percept. Perform.* 23, 861–872

33   Sterzer, P. *et al.* (2018) The predictive coding account of psychosis. *Biol. Psychiatry* 84, 634–643

34   Yon, D. *et al.* (2017) Time on your hands: Perceived duration of sensory events is biased toward concurrent actions. *J. Exp. Psychol. Gen.* 146, 182–193

35   Yon, D. and Press, C. (2018) Sensory predictions during action support perception of imitative reactions across suprasecond delays. *Cognition* 173, 21–27

36   Dogge, M. *et al.* (2019) Perception of action-outcomes is shaped by life-long and contextual expectations. *Sci. Rep.* 9, 1–9

37   Press, C. and Cook, R. (2015) Beyond action-specific simulation: Domain-general motor contributions to perception. *Trends Cogn. Sci.* 19, 176–178

38   Angel, R.W. and Malenka, R.C. (1982) Velocity-dependent suppression of cutaneous sensitivity during movement. *Exp. Neurol.* 77, 266–274

39   Juravle, G. *et al.* (2017) Tactile suppression in goal-directed movement. *Psychon. Bull. Rev.* 24, 1060–1076

40   Limanowski, J. *et al.* (2019) Action-dependent processing of touch in the human parietal operculum and posterior insula. *Cereb. Cortex* DOI: 10.1093/cercor/bhz111

41   Seki, K. and Fetz, E.E. (2012) Gating of sensory input at spinal and cortical levels during preparation and execution of voluntary movement. *J. Neurosci. Off. J. Soc. Neurosci.* 32, 890–902

42   Brown, H. *et al.* (2013) Active inference, sensory attenuation and illusions. *Cogn. Process.* 14, 411–427

43   Yon, D. *et al.* (2019) The predictive brain as a stubborn scientist. *Trends Cogn. Sci.* 23, 6–8

44   Friston, K.J. and Frith, C.D. (2015) Active inference, communication and hermeneutics. *Cortex* 68, 129–143

45   Yon, D. and Press, C. (2017) Predicted action consequences are perceptually facilitated before cancellation. *J. Exp. Psychol. Hum. Percept. Perform.* 43, 1073–1083

46   Dogge, M. *et al.* (2019) Moving forward: On the limits of motor-based forward models. *Trends Cogn. Sci.* DOI: 10.1016/j.tics.2019.06.008

47   Yu, A.J. and Dayan, P. (2005) Uncertainty, neuromodulation, and attention. *Neuron* 46, 681–692

48   Dayan, P. and Yu, A. (2003) Uncertainty and learning. *IETE J. Res.* 49, 171–181

49   Rescorla, R.A. and Wagner, A.R. (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical conditioning: Current research and theory*

50   Itti, L. and Baldi, P. (2009) Bayesian surprise attracts human attention. *Vision Res.* 49, 1295–1306

51   Dayan, P. and Yu, A.J. (2006) Phasic norepinephrine: A neural interrupt signal for unexpected events. *Netw. Comput. Neural Syst.* 17, 335–350

52   Devauges, V. and Sara, S.J. (1990) Activation of the noradrenergic system facilitates an attentional shift in the rat. *Behav. Brain Res.* 39, 19–28

53   Marshall, L. *et al.* (2016) Pharmacological fingerprints of contextual uncertainty. *PLOS Biol.* 14, e1002575

54   Redgrave, P. and Gurney, K. (2006) The short-latency dopamine signal: A role in discovering novel actions? *Nat. Rev. Neurosci.* 7, 967–975

55   Cutrone, E.K. *et al.* (2014) Attention enhances contrast appearance via increased input baseline of neural responses. *J. Vis.* 14, 16–16

56   Driver, J. and Spence, C. (1998) Attention and the crossmodal construction of space. *Trends Cogn. Sci.* 2, 254–262

57   Gandolfo, M. and Downing, P.E. (2019) Causal evidence for expression of perceptual expectations in category-selective extrastriate regions. *Curr. Biol.* 29, 2496-2500.e3

58   Kok, P. *et al.* (2017) Prior expectations induce prestimulus sensory templates. *Proc. Natl. Acad. Sci.* 114, 10473–10478

59   Press, C. *et al.* (2010) Action preparation helps and hinders perception of action. *J. Cogn. Neurosci.* 22, 2198–2211

60   Baess, P. *et al.* (2011) Selective suppression of self-initiated sounds in an auditory stream: An ERP study. *Psychophysiology* 48, 1276–1283

61   Klaffehn, A.L. *et al.* (2019) Sensory attenuation prevails when controlling for temporal predictability of self- and externally generated tones. *Neuropsychologia* 132, 107145

62   Le Bars, S. *et al.* (2019) Event-related brain potentials to self-triggered tones: Impact of action type and impulsivity traits. *Neuropsychologia* 125, 14–22

63   Todorovic, A. and Lange, F.P. de (2012) Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. *J. Neurosci.* 32, 13389–13395

64   Alilović, J. *et al.* (2019) No evidence that predictions and attention modulate the first feedforward sweep of cortical information processing. *Cereb. Cortex* 29, 2261–2278

65   Kouider, S. *et al.* (2015) Neural dynamics of prediction and surprise in infants. *Nat. Commun.* 6, 8537

66   Richter, D. and de Lange, F.P. (2019) Statistical learning attenuates visual activity only for attended stimuli. *eLife* 8, e47869

67   Posner, M.I. (1980) Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25

68   Courchesne, E. *et al.* (1975) Stimulus novelty, task relevance and the visual evoked potential in man. *Electroencephalogr. Clin. Neurophysiol.* 39, 131–143

69   Kok, P. *et al.* (2012) Attention reverses the effect of prediction in silencing sensory signals. *Cereb. Cortex* 22, 2197–2206

70   Gordon, N. *et al.* (2019) Expectation and attention increase the integration of top-down and bottom-up signals in perception through different pathways. *PLOS Biol.* 17, e3000233

71   Auksztulewicz, R. *et al.* (2017) The cumulative effects of predictability on synaptic gain in the auditory processing stream. *J. Neurosci.* 37, 6751–6760

72   Pearce, J.M. and Hall, G. (1980) A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* 87, 532–552

73   Mackintosh, N.J. (1975) A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychol. Rev.* 82, 276–298

74   Esber, G.R. and Haselgrove, M. (2011) Reconciling the influence of predictiveness and uncertainty on stimulus salience: A model of attention in associative learning. *Proc. R. Soc. B Biol. Sci.* 278, 2553–2561

75   Lawson, R.P. *et al.* (2017) Adults with autism overestimate the volatility of the sensory environment. *Nat. Neurosci.* 20, 1293–1299

76   Clark, A. (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 36, 181–204

77   Friston, K. (2018) Does predictive coding have a future? *Nat. Neurosci.* 21, 1019

78   Friston, K. (2010) The free-energy principle: A unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138

79   Rungratsameetaweemana, N. and Serences, J.T. (2019) Dissociating the impact of attention and expectation on early sensory processing. *Curr. Opin. Psychol.* 29, 181–186

80   Firestone, C. and Scholl, B.J. (2016) Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *Behav. Brain Sci.* 39, e229

81   Vetter, P. and Newen, A. (2014) Varieties of cognitive penetration in visual perception. *Conscious. Cogn.* 27, 62–75

82   Fritsche, M. *et al.* (2017) Opposite effects of recent history on perception and decision. *Curr. Biol.* 27, 590–595

83   Bang, J.W. and Rahnev, D. (2017) Stimulus expectation alters decision criterion but not sensory signal in perceptual decision making. *Sci. Rep.* 7, 17072

84   Cook, R. *et al.* (2014) Mirror neurons: From origin to function. *Behav. Brain Sci.* 37, 177–192

85   Hsu, Y.-F. *et al.* (2015) Distinctive representation of mispredicted and unpredicted prediction errors in human electroencephalography. *J. Neurosci.* 35, 14653–14660

86   Blank, H. *et al.* (2018) Neural prediction errors distinguish perception and misperception of speech. *J. Neurosci.* 38, 6076–6089

87   Blank, H. and Davis, M.H. (2016) Prediction errors but not sharpened signals simulate multivoxel fMRI patterns during speech perception. *PLoS Biol.* 14, e1002577

88   Hudson, M. *et al.* (2018) Perceptual teleology: Expectations of action efficiency bias social perception. *Proc. R. Soc. B Biol. Sci.* 285, 20180638

89   Tamir, D.I. and Thornton, M.A. (2018) Modeling the predictive social mind. *Trends Cogn. Sci.* 22, 201–212

90   Geuter, S. *et al.* (2017) Functional dissociation of stimulus intensity encoding and predictive coding of pain in the insula. *eLife* 6, e24770

91   Sussman, T.J. *et al.* (2017) Here comes trouble: Prestimulus brain activity predicts enhanced perception of threat. *Cereb. Cortex* 27, 2695–2707

92   Greve, A. *et al.* (2019) Knowledge is power: Prior knowledge aids memory for both congruent and incongruent events, but in different ways. *J. Exp. Psychol. Gen.* 148, 325–341

93   Kim, G. *et al.* (2014) Pruning of memories by context-based prediction error. *Proc. Natl. Acad. Sci.* 111, 8997–9002

94   Pellicano, E. and Burr, D. (2012) When the world becomes 'too real': A Bayesian explanation of autistic perception. *Trends Cogn. Sci.* 16, 504–510

95   Lawson, R.P. *et al.* (2014) An aberrant precision account of autism. *Front. Hum. Neurosci.* 8,

96   Bolis, D. and Schilbach, L. (2018) Observing and participating in social interactions: Action perception and action control across the autistic spectrum. *Dev. Cogn. Neurosci.* 29, 168–175

97   Galea, J.M. *et al.* (2012) Action reprogramming in Parkinson's Disease: Response to prediction error is modulated by levels of dopamine. *J. Neurosci.* 32, 542–550

98   Grotheer, M. and Kovács, G. (2016) Can predictive coding explain repetition suppression? *Cortex* 80, 113–124

99   Henson, R. *et al.* (2000) Neuroimaging evidence for dissociable forms of repetition priming. *Science* 287, 1269–1272

100 Segaert, K. *et al.* (2013) The suppression of repetition enhancement: A review of fMRI studies. *Neuropsychologia* 51, 59–66
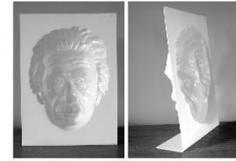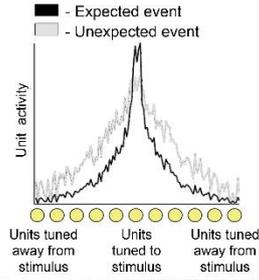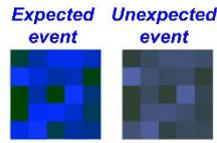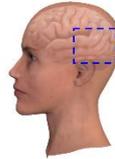
Figure 1, Key Figure. Existing conflicting theories and a proposed resolution. (a) Bayesian theories propose that perception is biased towards what we expect (e.g., a polar bear in the arctic) – and therefore, on average, most likely to be true – by increasing the gain on expected relative to unexpected units (optimising *veridicality*). Matrices with higher contrast reflect stronger sensory representations (high signal to noise ratio). These models can explain illusions when typical regularities are disrupted, e.g., perceiving concave faces to be convex (reproduced with permission from [13]). (b) Cancellation theories propose that the limited resources of our perceptual systems are devoted to unexpected signals (e.g., an elephant in the arctic) that may require us to update our beliefs and perform corrective actions, by suppressing – not increasing – expected sensory activity (optimising *informativeness*). These theories have been used to explain why we cannot tickle ourselves. These theories are in conflict with Bayesian theories. (c) Our Opposing Process theory proposes a resolution. It posits that perception is initially biased towards the expected to aid rapid generation of largely veridical experiences. When the input is sufficiently in line with expectations such that any deviation is likely attributable to sensory noise, no other process operates. (d) However, if the input is different enough from the prediction to generate surprise (high KLD), catecholamine release boosts perception of these surprising inputs, thus aiding learning.

Figure 2. Dependency of the opposing process on signal strength. (a) When an unexpected signal is weak – i.e., low precision – then surprise (KLD) is low and we will typically simply continue to perceive what we expect. (b) When an unexpected signal is strong – i.e., high precision – then this highly surprising input will generate processes characterised in the learning literature that increase sensory gain, upweighting processing of these inputs to aid model updating. In this way, unexpected events in our environment that may signal the need for model updating are perceptually highlighted, but 'unexpected' events that are inline with sensory noise (and therefore less informative for model updating) are not. Matrices with higher contrast reflect stronger sensory representations (high signal to noise ratio), but relative comparisons should only be made within panel (a) or within panel (b) - i.e., all signal to noise ratios will be higher in (b).

Figure I, Box 2. Domain-generality. We assume that *both* veridical *and* informative perception will be required in any perceptual domain, and therefore that the same opposing processes operate regardless of whether we generate expectations on the basis of (a) sensory context, (b) the actions we perform, or (c) the words and syllables that we hear.
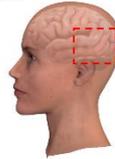
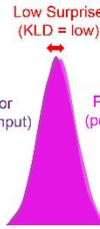**a)** **Bayesian theory**

(e.g.) Expected event

Expected event   *Unexpected event*

■ - Expected event
▨ - Unexpected event

Unit activity

Units tuned away from stimulus | Units tuned to stimulus | Units tuned away from stimulus

b) **Cancellation theory**

(e.g.) Unexpected event

Expected event   *Unexpected event*

■ - Expected event
▨ - Unexpected event

Unit activity

Units tuned away from stimulus | Units tuned to stimulus | Units tuned away from stimulus
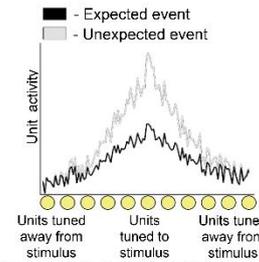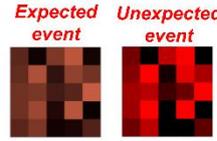
Opposing Process theory

c) Expected event

Expected event

Low Surprise
(KLD = low)
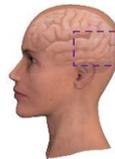
Prior (pre-input)   Posterior (post-input)

Expected event

time (ms)

d) Unexpected event

Unexpected event

High Surprise
(KLD = high)

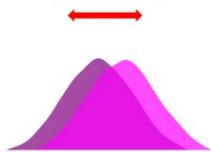Prior (pre-input)   Posterior (post-input)

Unexpected event

time (ms)

24
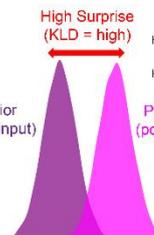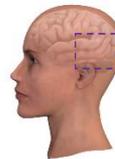
a) *'Weak' signals (e.g.)*
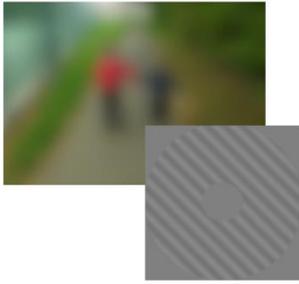
High noise
(low precision)

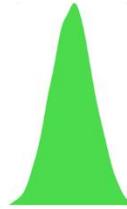*Expected event* *Unexpected event*
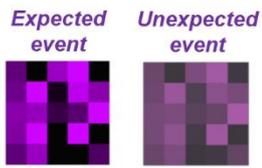
Low Surprise
(KLD = low)

b) *'Strong' signals (e.g.)*

Low noise
(high precision)
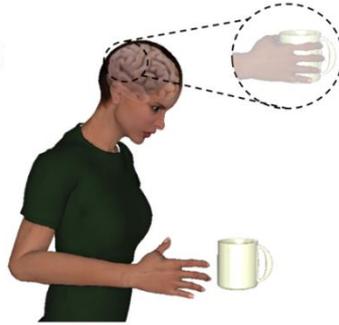
Expected event *Unexpected event* *Expected event* *Unexpected event*

High Surprise
(KLD = high)
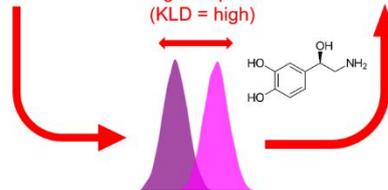
a) Sensory context   b) Action   c) Language