Review

Expectation (and attention) in visual cognition

Christopher Summerfield[1] and Tobias Egner[2,3]

[1] Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3UD, UK
[2] Department of Psychology and Neuroscience, Duke University, LSRC, Box 90999, Durham, NC 27708, USA
[3] Center for Cognitive Neuroscience, Duke University, LSRC, Box 90999, Durham, NC 27708, USA

Visual cognition is limited by computational capacity, because the brain can process only a fraction of the visual sensorium in detail, and by the inherent ambiguity of the information entering the visual system. Two mechanisms mitigate these burdens: attention prioritizes stimulus processing on the basis of motivational relevance, and expectations constrain visual interpretation on the basis of prior likelihood. Of the two, attention has been extensively investigated while expectation has been relatively neglected. Here, we review recent work that has begun to delineate a neurobiology of visual expectation, and contrast the findings with those of the attention literature, to explore how these two central influences on visual perception overlap, differ and interact.

Introduction

'Everyone', William James famously wrote, 'knows what attention is' [1]. A similar claim might be made for *expectation*: we have a rich and immediate introspective sense of what it means to expect a forthcoming stimulus, or for expectation to be violated ('Surprise!'). By comparison with attention, however, psychologists and neuroscientists have traditionally dedicated relatively little effort to exploring the brain mechanisms by which expectations about the forthcoming sensory world are generated, represented and implemented. On the contrary, the effects of perceptual expectation often tend to be conflated with those of attention in the empirical literature. For instance, many experimental protocols geared at gauging the effects of attention deploy probabilistic cues that manipulate subjects' expectations concerning the likely location or feature of a forthcoming target stimulus [2]. However, the behavioural and neural consequences of such probabilistic cueing are typically attributed exclusively to the effects of attention, even though this procedure also engenders perceptual expectations, as well as their fulfilment or violation. The purpose of this review, therefore, is twofold. First, we aim to highlight and synthesize recent empirical and computational work that is fostering an emergent neurobiology of expectation (with a focus on the visual domain). Second, in parallel to reviewing characteristics of visual expectation, we will emphasize how these might differ from, overlap with, or interact with, the much better-characterized processes underlying visual attention. We begin by considering the purpose and behavioural consequences of visual expectation.

Expectation in visual cognition

Expectations are brain states that reflect prior information about what is possible or probable in the forthcoming sensory environment. For example, on entering a familiar room, we have prior knowledge about the likely configuration of furniture, paintings or plants, gleaned over multiple previous encounters. This type of foreknowledge can be employed to reduce the computational burden of visual perception in two ways. First, expectations can guide the acquisition of visual information; aspects of the environment that are constant (and thus conform to previously formed expectations) do not have to be processed in depth repeatedly. Second, expectation facilitates the interpretation of visual input. Information arriving at the retina is inherently ambiguous: a single pattern of retinal stimulation can be caused by myriad different stimulus configurations, and a single object can elicit many distinct patterns of retinal stimulation, depending on viewpoint, lighting conditions, or the interposition of occluding surfaces. To overcome these ambiguities, the interpretation of a given percept ('Is this a breadbox or a mailbox?') can be guided by contextual probabilities ('Am I in the kitchen or on the street?'). Accordingly, much classical behavioral research (reviewed in [3]) has demonstrated that objects placed within a congruent context (for example, a lamp in a bedroom) are detected faster and more reliably than objects placed in an incongruent context (for example, a lamp on the beach). Similar results have been found in more controlled psychophysical experiments using rudimentary stimuli: for instance, detection of a Gabor patch (sinusoidal gratings) is facilitated by colinearly oriented flanking patches [4]. Despite ongoing controversy [5], these effects do not appear to be attributable to a change in response bias alone [6]. Benefits to visual recognition also accrue from preceding items that have been jointly or conditionally associated with a stimulus or array [7], from the statistics of natural image information [8], or even the mere recent presentation of related information [9]. The perceptual processes underlying visual detection and recognition, thus, have evolved to benefit from probabilistic information about what is likely to occur, via associative learning, priming, or contextual processing of scene information.

Some important similarities and distinctions can be drawn between these characteristics of visual expectation

---

Corresponding authors: Summerfield, C. (christopher.summerfield@psy.ox.ac.uk); Egner, T. (tobias.egner@duke.edu)

1364-6613/$ – see front matter © 2009 Elsevier Ltd. All rights reserved. doi:10.1016/j.tics.2009.06.003 Available online 27 August 2009

Box 1. Neurocomputational mechanisms of facilitation by top-down information.

We consider four ways in which perceptual decisions could be affected by the type of facilitatory top-down processes that might arise from attention or expectation (red arrows, Figure Ia; red ovals, Figure Ib).

1. Top-down facilitation could occur during the anticipatory processing phase, before stimulus onset. Under the framework of the decision model depicted in Figure I, this would be reflected in a higher level of prior evidence for stimulus presence over absence, that is, a higher starting point for the evidence collection process (Figure Ib: height of red oval 1 on the dashed line). This would lead to faster reaction times because evidence collection begins closer to the bound. Neurophysiologically, this might be reflected in enhanced rates of baseline firing in neurons tuned to the anticipated feature or location.
2. Facilitation could occur at an early phase of sensory processing, before combination of the sensory evidence with performance-limiting internal noise. This manipulation would enhance the gain of evidence accumulation ('contrast gain') and increase the signal-to-noise ratio. In the model, this would be reflected by a steeper slope to the accumulation process. Neurophsyiologically, this might be reflected in sharper tuning curves for neurons tuned to attended features.
3. Facilitation following the addition of internal noise ('response gain'). In the decision model, this would be captured by an increased upward drift of the signal+noise trace (Figure Ib: blue line). Although this would hasten reaction times, as evidence would reach the bound faster, it could also lead to more erroneous detections ('false alarms'), because noisy signals would also be enhanced. Neurophysiologically, this might lead to an upward scaling of responses across the tuning curve of a neuron, without increasing sensitivity to its preferred feature.
4. Top-down mechanism could leave perception unaltered, but lower the criterial level of evidence needed for a decision to be made ('response bias'). Reaction times would be shorter, because less evidence would need to be accumulated before the bound were breached. In situations of multiple choices (and bounds), this would lead to an increased tendency to favour speed over accuracy (unlike 1, above).

Neurophysiological studies have described evidence that top-down attention can enhance baseline firing rates [62–64], sharpen neuronal tuning towards task-relevant stimulus features [65,66] and scale up neuronal response curves [67,68]. These findings suggest that attention could impinge on perceptual decision-making via mechanisms 1, 2, and 3. Two recent articles provide excellent in-depth review (and integrated quantitative accounts) of this issue [69,70]. Although there is evidence that prior learning might modulate gain control in neurons in the prefrontal cortex [71], little is known about how perceptual decisions might be biased by expectations.

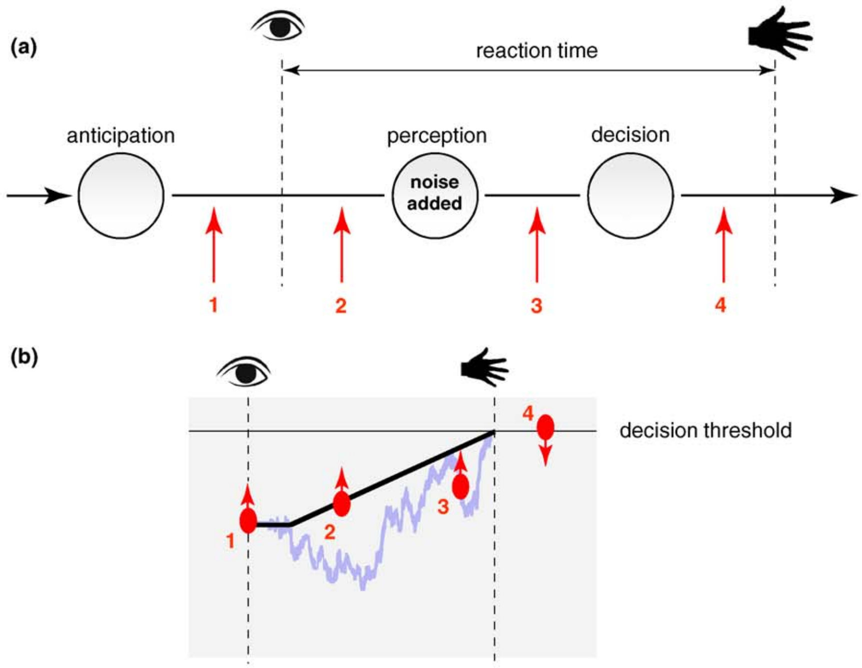


Figure I. Both attended and expected stimuli are responded to more rapidly than their unattended or unexpected counterparts. To help understand possible mechanisms underlying these effects, we consider the evolution of a perceptual decision across a representative trial in a detection task (**a**). The epoch considered encompasses anticipatory processing of the stimulus before its onset, stimulus presentation (dashed line by eye symbol) input to perceptual regions, perceptual processing and transfer to higher regions responsible for decision-making and response selection, and finally execution (dashed line by hand symbol). For a related account, see [72]. Concurrently, we describe a simple model of how perceptual decisions are made (**b**), in which evidence (y-axis) is accumulated serially and stochastically across time (x-axis) towards a decision threshold or 'bound' (thin black line). This model is related to classic descriptions of the decision process, reviewed in [73]. The slope of the accumulation process (thick black line) reflects the rate at which evidence can be accumulated; the blue line reflects an example trial, in which both signal and noise contribute to perceptual decision-making.

and those of visual attention. Expectation facilitates visual perception by constraining interpretation on the basis of prior likelihood, whereas attention alleviates computational burden by prioritizing processing of that subset of visual information deemed to be of the highest relevance to the organism's goals [10]. Thus, attention is directed endogenously ('top-down') to favour processing of features, objects, or spatial locations that are associated with a task goal, reward, or potential danger [2,11], and attention is attracted exogenously ('bottom-up') towards stimuli of high physical salience, such as sudden motion [12]. Critically, the motivational considerations that guide attention are in principle orthogonal to the perceptual regularities that guide visual expectations - although in practice, expectations and attention often coincide and interact. For example, learned perceptual context can guide attention towards a likely location of a behaviourally relevant item [13] (even unconsciously [14]), and the violation of a perceptual expectation might be employed as a signal for attracting attention towards a potentially significant event [15]. Conversely, attention facilitates the formation of stimulus–stimulus associations [16], and a modicum of attentional resources might in fact be a prerequisite for the acquisition of contextual knowledge [17,18]. Regardless of these conceptual distinctions, it is particularly noteworthy that the behavioural effects of expectation and attention are superficially identical; both attended and expected stimuli are detected and recognized more readily than unattended or unexpected. The exact mechanisms underlying this facilitation, however, remain subject to ongoing debate (Box 1). Given the similarity of their facilitative effects on object recognition, the question arises as to whether expectation and attention might share neural mechanisms, discussed next.

Neural sources of expectation

A natural way of understanding the role of expectation in visual cognition is as a 'top-down' process, in which prior information, perhaps stored in higher brain regions, feeds back to modulate basic perceptual processing in sensory cortices. This conceptualization has close contact with that of top-down attentional biasing. However, while the major sources of top-down attentional biasing have long been located principally (although not exclusively) in the dorsolateral prefrontal and posterior parietal cortices [19], only recently have attempts been made to identify the sources of expectation during object recognition. Intriguingly, the emerging evidence suggests that the neural origins of contextual biasing signals lie more medially in the posterior cortices as well as more ventrally in the frontal lobe than those of attentional biasing. For example, parahippocampal and retrosplenial zones seem to encode information about visual scenes that might prompt object recognition [20] and these regions are activated by objects that have strong relative to weak contextual associations [21]. Pioneering work by Bar and colleagues has emphasized the role that the orbitofrontal cortex (OFC) might play in 'top-down' object recognition [22,23], arguing that during scene perception, a rapid extraction of low-level, low spatial frequency information might be fed to the OFC, where associative information allows the formation of a predictive 'initial guess' about what objects or elements are likely in the scene [24]. This information is then fed back to the extrastriate visual cortices to bias recognition towards contextually congruent objects. This account draws support from a variety of literatures, including the reports of very early, scene- and object-related potentials over the frontopolar cortices [25], models of the OFC as a heteromodal nexus for stimulus–stimulus associations [26], and the long-standing implication of frontopolar regions in retrieval of information from long-term memory [27]. Empirical support has also been provided by brain imaging studies using functional connectivity techniques, revealing a unidirectional flow of information from the medial OFC to the visual cortices during categorization of degraded face images [28] (Figure 1) and perception of low-frequency objects [29].

Neural effects of expectation in visual cortex

A large electrophysiological literature has documented that top-down attention acts by enhancing the responsivity of sensory neurons that code for a motivationally relevant feature, object or location [11], leading to an increase in the neural activity elicited by an attended relative to an unattended stimulus, although the precise nature of this modulation remains controversial (Box 1). Given that attention and expectation have similar facilitatory effect on visual object recognition, one might anticipate that expected (relative to unexpected) stimuli would also be associated with enhanced sensory responses. However - strikingly - the opposite is in fact typically the case: expected stimuli tend to elicit reduced visual responses, relative to their

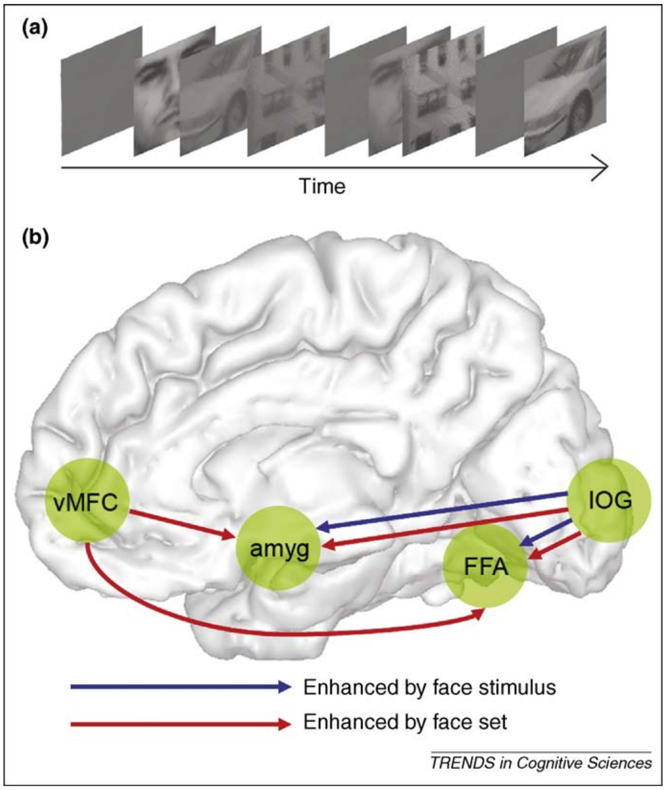
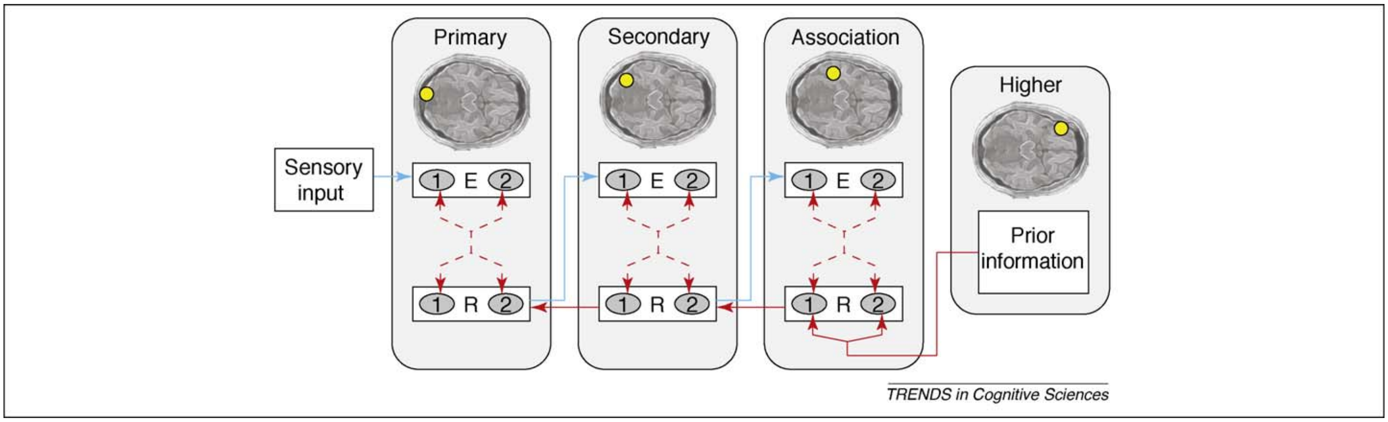


Figure 1. Frontal contributions to perceptual inference. **(a)** In this fMRI study [28], subjects were presented with randomly intermixed degraded and masked (masks not shown) images of faces, houses, and cars, tailored to each individual's perceptual threshold. Via two response buttons, subjects were required to indicate whether the current stimulus was a face or not a face (in 'face set' blocks), or whether the current stimulus was a house or not a house (in 'house set' blocks). Thus, while the perceptual input was kept identical across blocks, the task encouraged subjects to employ distinct top-down perceptual sets (or templates) in the two block types, one for detecting face stimuli, and one for detecting house stimuli. Contrasting face with non-face stimuli (irrespective of set type) revealed stimulus-driven activation in core regions of the face processing network, including the inferior occipital gyrus (IOG), fusiform face area (FFA), and amygdala (amyg), whereas face set-related activity (irrespective of stimulus type) was observed prominently in the ventral medial frontal cortex (vMFC) (data not shown). **(b)** Dynamic causal modeling [74] revealed that during face set blocks (as compared with house set blocks), the vMFC displayed enhanced top-down effective connectivity with the amygdala and FFA, while both face stimuli and face set affected bottom-up connectivity from the IOG to the FFA and amygdala. These data support the notion that ventral frontal regions might provide top-down perceptual 'priors' to posterior regions, where these predictions are matched against incoming sensory information. Adapted, with permission, from Ref. [28].



TRENDS in Cognitive Sciences

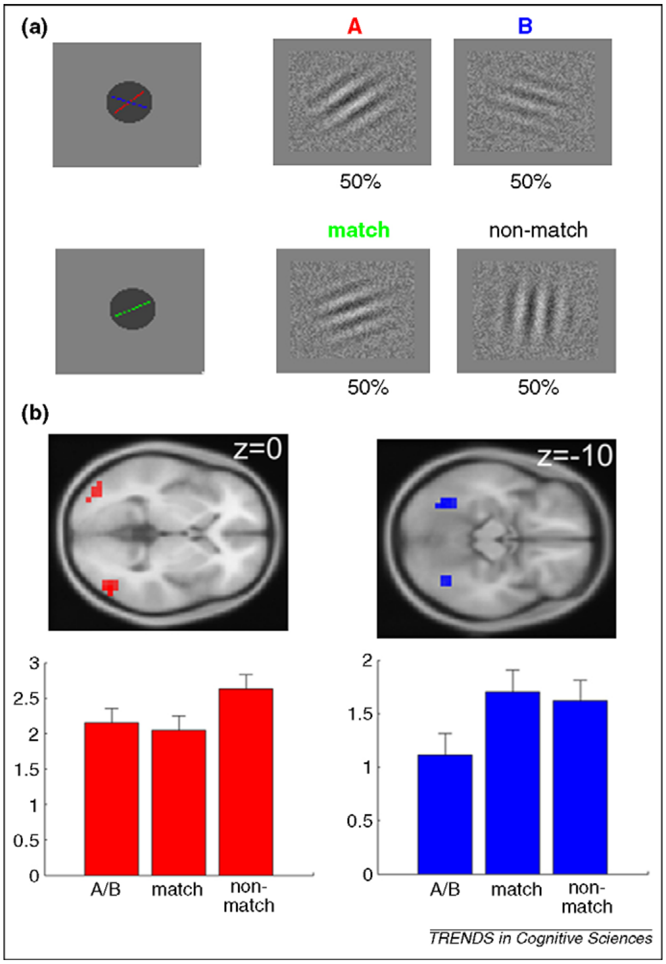
Figure 2. Predictive coding in visual perception. Predictive coding [40,41,43] argues that perception unfolds as expected and observed information are gradually reconciled, via iterative interactions within a hierarchically organised sensory system. Here, this hierarchy is shown schematically, with grey panels representing the passage of sensory information between distinct sites (yellow dots on an axial slice of the brain) stretching from primary to higher association cortex. The key elements of the scheme as described in [43] are (i) two functionally distinct classes of neurons. Representation units (R) (lower white boxes) encode expectations about what is possible or likely in the coming sensory world. Error units (E) (upper white boxes) are driven by sensory input from the preceding layer of the hierarchy (blue arrows). (ii) Predictions are 'subtracted' from E unit responses; surprising events elicit larger neuronal responses. (iii) R unit predictions are updated on the basis of the mismatch between expected and observed information. Both (ii) and (iii) depend on local circuit interactions (dashed red arrows) occurring via many-to-many interactions between individual nodes (grey ellipses; for simplicity, only two are shown). (iv) Resultant predictions in turn are used as priors for the preceding layer (red arrows). A related account has been described elsewhere [41,42].

unexpected counterparts [30,31], and an extensive literature has documented the corresponding phenomenon in the auditory domain in the form of the 'mismatch negativity' event-related potential [32]. In line with these findings, recent neuroimaging studies employing formal measures derived from information theory [33], or reinforcement learning [34], have documented increased visual cortical responses to visual surprise. Unfortunately, an unambiguous characterization of expectation effects at the level of single neurons is currently lacking in the literature (but see the section 'Is there evidence for error and representation neurons?', below, for a discussion of suggestive evidence). One possible interpretation of the event-related potential and neuroimaging results is that unexpected stimuli evoke stronger responses than expected ones because they attract attention. However, attentional effects alone are unlikely to account for these data. First, many experiments have demonstrated sensory responses to expectancy violations in a task-irrelevant stimulus stream, while attention was occupied elsewhere [35–37]. Second, the differential response between expected and unexpected stimuli is not solely a function of enhanced activity evoked by the unexpected stimulus, but is substantially driven by accumulative decreases in the response to stimuli that conform to expectation [38,39]. If attention is not responsible for the enhanced neural responses to unexpected relative to expected stimuli, what could be the mediating mechanism?

Predictive coding models of visual expectation

Models of visual attention are concerned with explaining the enhancement of neural responses to attended relative to unattended stimuli (Box 1). As described in the previous section, the effects of expectation on visual responses, however, exhibit the exact opposite pattern. A class of models aimed at explaining the effects of prior knowledge on visual perception offers an intriguing explanation for these neuroscientific findings. The formal quantitative [40] and computational [41–43] framework known as 'predictive coding' proposes that information flowing forwards through a hierarchy of sensory regions is met at each stage by a top-down 'prediction' projected back from the layer immediately above (Figure 2). Neural implementations of this framework assume that there are two interacting classes of sensory neurons at each level of the visual processing stream. *Representational neurons* (R units) encode conditional (context-dependent) expectations about the identity of a stimulus and project their predictions to the next lower level of the processing hierarchy. These units are complemented by *error neurons* (E units) that encode 'prediction error' when the expected and observed sensory information diverge, and pass this error message to the R units, where predictions are updated accordingly (these proposed error units represent a sensory equivalent to the well-studied dopaminergic reward prediction error neurons of the midbrain [44]). Predictive coding argues that perceptual inference proceeds as expected and observed information are iteratively reconciled across multiple levels of the visual processing hierarchy, resulting in a progressive reduction in prediction error as the visual system settles on a single perceptual interpretation of the sensory input. Importantly, the proposal of co-existing representational and error neuron populations provides a natural account for the observation that cortical responses scale with sensory surprise: when stimulation conforms to expectations, error neuron activity will be minimal, but when sensory evidence diverges from predictions, error units will respond vigorously.

Indirect support for predictive coding includes the realization that the spatial range over which contextual effects in visual detection (such as facilitation by colinear flankers) have been observed precludes a mechanism that relies exclusively on lateral interactions among neurons in a given cortical stage (for example, via horizontal connections within V1). Rather, these effects - and 'extra-classical' receptive field (RF) effects, where the response properties of a neuron are modulated by information falling beyond its classical RF - can only be explained by feedback from subsequent cortical stages [41,45]. Furthermore, inactivation of higher extrastriate visual regions (such as V5) alters the response properties of neurons in lower regions (such as V1), and impairs perceptual inference, for example during figure/ground segmentation [46]. Predictive coding also makes some counter-intuitive predictions that have been borne out by recent empirical work. One intriguing implication of this scheme is that the omission of an expected stimulus should elicit a (prediction error) response in sensory cortex; this hypothesis has been confirmed in a elegant brain imaging experiment, in which subjects learned task-irrelevant, probabilistic audio–visual associations [34]. fMRI responses in V1 to unexpected visual stimuli were greater than to expected visual stimuli, but were largest in trials where a visual stimulus was predicted by an auditory cue, but did not occur. Similar results have been obtained in the auditory modality [47]. Finally, predictive coding holds that if higher sensory regions are able to 'explain' a perceptual input, activity elicited by this input in lower regions should be reduced (or 'explained away'). This prediction has also been confirmed [48]. For instance, activity in V1 is reduced while activity in LOC is enhanced when viewing line elements arranged in coherent shapes compared with when the same elements are arranged randomly [48]. Another recent fMRI study providing support for predictive coding [49] has described dissociable visual regions that respond to the presence of 'prior' information (when a visual expectation is engendered) and to 'prediction error' (when that expectation is violated) (Figure 3).



TRENDS in Cognitive Sciences

Figure 3. Expectations and prediction error responses in the ventral visual stream. (**a**) In a recent study [49], subjects matched the orientation of contrast-modulated Gabor patches to target line orientations. In one condition (A/B task), matching was performed as a two-alternatives forced-choice discrimination, where subjects decided which of two target line orientations (red and blue lines) was matched by the current grating (top panel). This condition encourages an unbiased evidence accumulation process [75]. In a second condition (A/~A task), gratings were judged in reference to a single target orientation line (green line), resulting in matches and non-matches (bottom panel). This condition invokes a biased evidence accumulation process, where subjects employ the target line as a perceptual template (a 'prior' or expectation) against which to match incoming evidence [40]. (**b**) Regions in bilateral inferior temporal gyrus (in red, bottom left) displayed visual prediction error responses, reflected in higher fMRI activity in the A/~A non-match condition, where the stimulus violated the prior, than during matches in the A/~A and the A/B task. Regions in the fusiform gyrus (in blue, bottom right) represented prior expectations, displaying higher responses in both conditions of the A/~A task (where a perceptual prior is employed) than in the A/B task. Adapted, with permission, from Ref [49].

How can expectation and attention be integrated?

Predictive coding provides a parsimonious account of the neural signature of visual expectations, but how can the effects of top-down attention be accommodated in this scheme? This question provides a major challenge to visual neuroscience, but only few theoretical proposals have been put forward thus far and, to our knowledge, not a single empirical study has manipulated attention and expectation orthogonally and assessed the resulting effects on visual responses. One proposal has attempted to reconcile top-down mechanisms of expectation and attention, by merging predictive coding [41] with standard accounts of attentional biasing under a common computational framework [43]. Under this scheme, visual neurons receive top-down input that facilitates their responding, but rather than competing for processing resources via lateral inhibition as previously proposed [50], these neurons (R units) compete for inputs from a distinct population of neurons [51] that in turn receive feedforward input from the immediately preceding cortical or thalamic level - akin to the error neurons (E units) in the predictive coding scheme (Figure 2). Because top-down information upregulates responding in R units, the model accounts for a range of data describing single neuron responses to attended and unattended features and locations [52]. Critically, however, the iterative interactions between R and E units preserve the main tenet of predictive coding, that predictions are 'subtracted' away from observed sensory input via feedback from representation to error neurons [40]. Theoretically, thus, visual 'surprise' responses could be accounted for as the responses of these E units to unexpected stimuli. However, it has yet to be reported whether the model can account for data from experiments that manipulated contextual expectations, or expectation and attention simultaneously. It is therefore currently not clear whether (and how) this scheme, which in its current form treats expectation and attention as interchangeable top-down influences on visual neurons, could reproduce the seemingly distinct effects of expectation and attention at the level of neuronal populations discussed in the previous section.

Is there evidence for error and representation neurons?

This account [43], and related computational models [42] could yet offer a fruitful path to understanding why visual regions seem to respond to both attended, and yet also to unexpected, visual stimuli. However, one limitation of these models - and of predictive coding in general - is that to date no single-neuron study has systematically pursued the search for sensory prediction error responses. Nevertheless, indirect evidence exists that both representational and error responses might be observed in visual regions. First, some early single-neuron work identified extrastriate visual neurons that responded to orientations that were 'matches' or 'mismatches' to a cued target orientation [53–55]. These 'match' neurons additionally responded to the instruction cue designating the target orientation, even when it was presented in a nonvisual modality [54]. This behaviour seems strikingly like that of the proposed representation (match) and error (mismatch) neurons proposed by predictive coding (see also Figure 3), but contextual expectations were not formally manipulated in these studies. Second, single neurons in the visual cortex have been identified that display robust adaptation effects [56]. Adaptation might reflect neuronal fatigue or sharpening [57], but could also reflect the dampening of surprise responses in error neurons to the second or subsequent presentation of a stimulus [40,58]. Other, more recent studies have emphasized that pyramidal cells and interneurons in V1 might make distinct contributions to attentional biasing (showing enhanced responses to a preferred stimulus and suppressed responses to competitors, respectively [59,60]), but there is as yet no evidence that their responses map onto those of the proposed R and E units. Clearly, additional work is required to understand the local circuit interactions that give rise to visual 'surprise' responses alongside more classically described modulations by attention. Important steps towards realizing this goal will require a careful characterization of visual responses under conditions where expectations and attention are orthogonalized.

Moreover, despite its apparent success in explaining a range of neurophysiological phenomena, predictive coding can be criticized on a number of grounds. First, the sheer speed at which object recognition can occur might preclude a slow, iterative resolution of object identity via local circuit interactions in the visual cortex [22,25]. Second, researchers have noted that linear models of recognition based on subtractive, match–mismatch interactions might become catastrophically unstable [61], and a divisive variant of predictive coding, in which observed sensory evidence is normalized by expectations, might be a more plausible alternative [52].

Conclusions

We have reviewed recent advances in understanding a traditionally neglected aspect of top-down biases on visual processing, namely, the use of perceptual expectation to contextually bias the accumulation and interpretation of visual information. This literature suggests that fulfilled expectations are associated with facilitated object recognition but attenuated neural responses; predictive coding models provide an elegant theoretical account for these effects. A central goal of the current review was to emphasize similarities and distinctions between expectation and top-down attention, in particular because these processes might often be conflated in studies of attention. The above survey of their paradoxical behavioural and neural effects reinforces the importance of carefully distinguishing between expectation and attention in experimental design and interpretation: for instance, when attention and expectation toward a target stimulus coincide, both processes will contribute to behavioural benefits in stimulus recognition, but they will exert opposing effects on neural activity in visual regions representing the stimulus. However, the detailed nature of these interactions between expectation and attention is not currently understood, as modelling work on this topic in its infancy, and experimental data are scarce (for some particularly pressing questions, see Box 2). We hope that this review serves to stimulate additional interest in pursuing an improved understanding of this important and exciting issue.

Box 2. Outstanding questions

- How do behavioral and neural effects of expectation and attention relate to each other when these factors are manipulated independently in a single experiment?
- What is the precise mechanism (or mechanisms) by which expectations affect perceptual decision-making (Box 1)?
- Do dedicated 'prediction error' neurons exist in sensory cortices?
- Where might expectation-related codes reside in the brain?
- To what degree are neural and behavioral effects that have traditionally been attributed to attention (for example, in probabilistic cueing paradigms) confounded by perceptual expectations?

References

1 James, W. (1890) *The Principles of Psychology*. Henry Holt
2 Posner, M.I. *et al.* (1980) Attention and the detection of signals. *J Exp Psychol* 109, 160–174
3 Bar, M. (2004) Visual objects in context. *Nat Rev Neurosci* 5, 617–629
4 Polat, U. and Sagi, D. (1994) Spatial interactions in human vision: from near to far via experience-dependent cascades of connections. *Proc Natl Acad Sci U S A* 91, 1206–1209
5 Henderson, J.M. and Hollingworth, A. (1999) High-level scene perception. *Annu Rev Psychol* 50, 243–271
6 Auckland, M.E. *et al.* (2007) Nontarget objects can influence perceptual processes during object recognition. *Psychon Bull Rev* 14, 332–337
7 Fiser, J. and Aslin, R.N. (2001) Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychol Sci* 12, 499–504
8 Geisler, W.S. (2008) Visual perception and the statistical properties of natural scenes. *Annu Rev Psychol* 59, 167–192
9 Maljkovic, V. and Nakayama, K. (1994) Priming of pop-out: I. Role of features. *Mem Cognit* 22, 657–672
10 Tsotsos, J.K. (1997) Limited capacity of any realizable perceptual system is a sufficient reason for attentive behavior. *Conscious Cogn* 6, 429–436
11 Maunsell, J.H. and Treue, S. (2006) Feature-based attention in visual cortex. *Trends Neurosci* 29, 317–322
12 Yantis, S. and Jonides, J. (1990) Abrupt visual onsets and selective attention: voluntary versus automatic allocation. *J Exp Psychol Hum Percept Perform* 16, 121–134
13 Summerfield, J.J. *et al.* (2006) Orienting attention based on long-term memory experience. *Neuron* 49, 905–916
14 Chun, M.M. and Jiang, Y. (1998) Contextual cueing: implicit learning and memory of visual context guides spatial attention. *Cognit Psychol* 36, 28–71
15 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
16 Chun, M.M. and Turk-Browne, N.B. (2007) Interactions between attention and memory. *Curr Opin Neurobiol* 17, 177–184

17 Jiang, Y. and Chun, M.M. (2001) Selective attention modulates implicit learning. *Q J Exp Psychol A* 54, 1105–1124

18 Troyer, A.K. and Craik, F.I. (2000) The effect of divided attention on memory for items and their context. *Can J Exp Psychol* 54, 161–171

19 Corbetta, M. and Shulman, G.L. (2002) Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* 3, 201–215

20 Epstein, R. and Kanwisher, N. (1998) A cortical representation of the local visual environment. *Nature* 392, 598–601

21 Bar, M. and Aminoff, E. (2003) Cortical analysis of visual context. *Neuron* 38, 347–358

22 Bar, M. *et al.* (2006) Top-down facilitation of visual recognition. *Proc Natl Acad Sci U S A* 103, 449–454

23 Bar, M. *et al.* (2001) Cortical mechanisms specific to explicit visual object recognition. *Neuron* 29, 529–535

24 Fenske, M.J. *et al.* (2006) Top-down facilitation of visual object recognition: object-based and context-based contributions. *Prog. Brain Res* 155, 3–21

25 Thorpe, S. *et al.* (1996) Speed of processing in the human visual system. *Nature* 381, 520–522

26 Rushworth, M.F. *et al.* (2007) Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. *Trends Cogn Sci* 11, 168–176

27 Nyberg, L. *et al.* (1995) Functional brain maps of retrieval mode and recovery of episodic information. *Neuroreport* 7, 249–252

28 Summerfield, C. *et al.* (2006) Predictive codes for forthcoming perception in the frontal cortex. *Science* 314, 1311–1314

29 Kveraga, K. *et al.* (2007) Magnocellular projections as the trigger of top-down facilitation in recognition. *J Neurosci* 27, 13232–13240

30 Marois, R. *et al.* (2000) A stimulus-driven approach to object identity and location processing in the human brain. *Neuron* 25, 717–728

31 Yoshiura, T. *et al.* (1999) Functional MRI study of auditory and visual oddball tasks. *Neuroreport* 10, 1683–1688

32 Naatanen, R. *et al.* (2007) The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin Neurophysiol* 118, 2544–2590

33 Strange, B.A. *et al.* (2005) Information theory, novelty and hippocampal responses: unpredicted or unpredictable? *Neural Netw* 18, 225–230

34 den Ouden, H.E. *et al.* (2009) A Dual Role for Prediction Error in Associative Learning. *Cereb Cortex* 19, 1175–1185

35 Czigler, I. *et al.* (2002) Memory-based detection of task-irrelevant visual changes. *Psychophysiology* 39, 869–873

36 Opitz, B. *et al.* (1999) Combining electrophysiological and hemodynamic measures of the auditory oddball. *Psychophysiology* 36, 142–147

37 Rinne, T. *et al.* (2005) Superior temporal and inferior frontal cortices are activated by infrequent sound duration decrements: an fMRI study. *Neuroimage* 26, 66–72

38 Baldeweg, T. (2007) ERP repetition effects and mismatch negativity generation. *J Psychophysiol* 21, 204–213

39 Haenschel, C. *et al.* (2005) Event-related brain potential correlates of human auditory sensory memory-trace formation. *J Neurosci* 25, 10494–10501

40 Friston, K. (2005) A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci* 360, 815–836

41 Rao, R.P. 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

42 Rao, R.P. and Ballard, D.H. (2004) Probabilistic models of attention based on iconic representations and predictive coding. In *Neurobiology of Attention* (Itti, L. *et al.*, eds), pp. 553–561, Elsevier Academic Press

43 Spratling, M.W. (2008) Reconciling predictive coding and biased competition models of cortical function. *Front Comput Neurosci* 2, 4

44 Schultz, W. and Dickinson, A. (2000) Neuronal coding of prediction errors. *Annu Rev Neurosci* 23, 473–500

45 Angelucci, A. and Bullier, J. (2003) Reaching beyond the classical receptive field of V1 neurons: horizontal or feedback axons? *J Physiol Paris* 97, 141–154

46 Hupe, J.M. *et al.* (1998) Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature* 394, 784–787

47 Hughes, H.C. *et al.* (2001) Responses of human auditory association cortex to the omission of an expected acoustic event. *Neuroimage* 13, 1073–1089

48 Murray, S.O. *et al.* (2002) Shape perception reduces activity in human primary visual cortex. *Proc Natl Acad Sci U S A* 99, 15164–15169

49 Summerfield, C. and Koechlin, E. (2008) A neural representation of prior information during perceptual inference. *Neuron* 59, 336–347

50 Deco, G. and Rolls, E.T. (2005) Neurodynamics of biased competition and cooperation for attention: a model with spiking neurons. *J Neurophysiol* 94, 295–313

51 Harpur, G.F. and Prager, R.W. (1994) A fast method for activating competitive self-organising neural networks. In *Proceedings of the International Symposium on Artificial Neural Networks*, 412–418

52 Spratling, M.W. (2008) Predictive coding as a model of biased competition in visual attention. *Vision Res* 48, 1391–1408

53 Haenny, P.E. *et al.* (1988) State dependent activity in monkey visual cortex. II. Retinal and extraretinal factors in V4. *Exp. Brain Res* 69, 245–259

54 Motter, B.C. (1994) Neural correlates of feature selective memory and pop-out in extrastriate area V4. *J Neurosci* 14, 2190–2199

55 Miller, E.K. and Desimone, R. (1994) Parallel neuronal mechanisms for short-term memory. *Science* 263, 520–522

56 Muller, J.R. *et al.* (1999) Rapid adaptation in visual cortex to the structure of images. *Science* 285, 1405–1408

57 Grill-Spector, K. *et al.* (2006) Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn Sci* 10, 14–23

58 Summerfield, C. *et al.* (2008) Neural repetition suppression reflects fulfilled perceptual expectations. *Nat Neurosci*

59 Tiesinga, P.H. (2005) Stimulus competition by inhibitory interference. *Neural Comput* 17, 2421–2453

60 Mitchell, J.F. *et al.* (2007) Differential attention-dependent response modulation across cell classes in macaque visual area V4. *Neuron* 55, 131–141

61 Grossberg, S. (1988) Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural Networks* 1, 17–61

62 Chelazzi, L. *et al.* (1998) Responses of neurons in inferior temporal cortex during memory-guided visual search. *J Neurophysiol* 80, 2918–2940

63 Chelazzi, L. *et al.* (1993) A neural basis for visual search in inferior temporal cortex. *Nature* 363, 345–347

64 Kastner, S. *et al.* (1999) Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron* 22, 751–761

65 Spitzer, H. *et al.* (1988) Increased attention enhances both behavioral and neuronal performance. *Science* 240, 338–340

66 Yeshurun, Y. and Carrasco, M. (1998) Attention improves or impairs visual performance by enhancing spatial resolution. *Nature* 396, 72–75

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

68 Williford, T. and Maunsell, J.H. (2006) Effects of spatial attention on contrast response functions in macaque area V4. *J Neurophysiol* 96, 40–54

69 Reynolds, J.H. and Heeger, D.J. (2009) The normalization model of attention. *Neuron* 61, 168–185

70 Boynton, G.M. (2009) A framework for describing the effects of attention on visual responses. *Vision Res* 49, 1129–1143

71 Bichot, N.P. *et al.* (1996) Visual feature selectivity in frontal eye fields induced by experience in mature macaques. *Nature* 381, 697–699

72 Liston, D.B. and Stone, L.S. (2008) Effects of prior information and reward on oculomotor and perceptual choices. *J Neurosci* 28, 13866–13875

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

74 Friston, K.J. *et al.* (2003) Dynamic causal modelling. *Neuroimage* 19, 1273–1302

75 Ratcliff, R. (1978) A theory of memory retrieval. *Psychol Rev* 85, 59–108