

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

Issue: *Competitive Visual Processing Across Space and Time***Prediction, context, and competition in visual recognition**Sabrina Trapp¹ and Moshe Bar²¹Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany. ²Gonda Center for Brain Research, Bar-Ilan University, Ramat Gan, Israel

Address for correspondence: Dr. Sabrina Trapp, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, D-04103 Leipzig, Germany. strapp@cbs.mpg.de

Perception is substantially facilitated by top-down influences, typically seen as predictions. Here, we outline that the process is competitive in nature, in that sensory input initially activates multiple possible interpretations, or perceptual hypotheses, of its causes. This raises the question of how the selection of the correct interpretation from among those multiple hypotheses is achieved. We first review previous findings in support of such a competitive nature of perceptual processing, and then propose which neural regions might provide a platform for rising and using expectations to resolve this competition. Specifically, we propose that it is the rapid extraction and top-down dissemination of a global context signal from the frontal cortices, particularly the orbitofrontal cortex, that affords the quick and reliable resolution of the initial competition among likely alternatives toward a singular percept.

Keywords: perceptual hypothesis; perceptual task; expectations; top-down; object recognition; orbitofrontal cortex; biased competition

Competition as an integrative part of visual recognition

The traditional view of visual recognition has been hierarchical, in that processing advances from the bottom up and analysis gradually increases in complexity. But the input to the visual system is often noisy and ambiguous—objects in natural visual scenes appear with occlusion, suboptimal illumination conditions, motion, and changing viewpoints, hampering an unequivocal mapping from input to object representation. Those obstacles make it seem less likely that an exclusively bottom-up process will be sufficient for efficient and fast visual recognition. Accordingly, a growing body of theoretical work and empirical evidence support the idea that this process is facilitated by top-down expectations.^{1–5} However, where in the brain expectations arise remains a key question to be addressed. We extend this evolving top-down perspective of visual processing here by emphasizing the competitive aspect in visual recognition, suggest how it is resolved by global contextual information, and propose a key neural region that might fulfill the role of mediating such top-down information.

Intuitively, the enormous amount of sensory information around us creates a competition, as we cannot look at everything simultaneously. On the level of single cells, sensory stimuli compete for receptive fields, which become increasingly sparse the higher that visual information is processed in the cortical hierarchy. Researchers have proposed attention as the mechanism that biases such competition toward the most relevant stimuli in the external world.⁶ This was elegantly demonstrated by presenting two visual stimuli within the same receptive field, while a monkey's attention was oriented to only one of the stimuli.⁷ It was observed that, in this condition, the neuronal response mimicked the cell's response when the attended stimulus was presented alone. Similar effects were also found in humans using functional magnetic resonance imaging (fMRI).^{8,9} However, competition does not end with the selection of the most relevant information from the external world. Competition has been extensively investigated in the context of multistable perception and binocular rivalry (for a review, see Ref. 10). Here, an ambiguous visual pattern such as the Necker cube elicits alternating visual percepts (i.e., one perceives either one or the other point of view and interpretation of the cube). It has been

doi: 10.1111/nyas.12680

shown that a large fraction of neurons in the striate and early extrastriate areas maintain a relatively constant level of activity for both interpretations, suggesting ongoing competition between the two different representations.¹¹ But multistable perception is an extreme case, with two interpretations being equally likely.^{12,13} We, however, elaborate a proposal that rests upon the assumption that every sensory input is inherently ambiguous because it is not possible to unequivocally determine a label to the input elicited by a particular object, as it might map onto different neuronal patterns stored in memory.

From a more formal point of view, the drift-diffusion model or related conceptualizations of perceptual decision making have proposed a (competitive) race among multiple candidates until a threshold is reached toward one interpretation.^{14,15} Such models usually incorporate previous knowledge or expectations by postulating some sort of initial bias or, in Bayesian terms, a prior that is assumed to increase the baseline evidence for one candidate or to accelerate its drift rate.^{16,17} However, just like Bayesian frameworks, owing to their formal nature, such models remain silent about where such priors originate. We here build upon the idea that competition is a generic part of object recognition and that it constitutes an important problem that needs to be resolved during perceptual inference. Specifically, we postulate that ambiguous information that is available initially gives rise to multiple possible interpretations of the input or *perceptual hypotheses*. These hypotheses compete, and (conscious) perception is not possible before a selection of one of them is made. In the following sections, we will describe this competition in further detail and propose a neural framework for its efficient resolution with context.

To shed light on postsensory competition and its resolution more broadly, it might be helpful to consider a framework that was suggested in motor control. The *affordance competition hypothesis*¹⁸ suggests that the brain specifies several actions in parallel while further information is collected to bias this competition toward a singular response. This idea was developed on the basis of an intriguing experiment in which neuronal activity in the dorsal premotor cortices of monkeys was recorded during a reach-selection task.¹⁹ First, the monkey received a cue (cue 1) that indicated two opposite potential reaching actions, only one of which would later be

indicated as the correct one by an additional, non-spatial cue (cue 2). Interestingly, they found that, during the period of ambiguity, even after cue 1 vanished, the neural population simultaneously encoded two potential and opposing directions until cue 2 finally specified the correct direction. After cue 2, the representation of the chosen direction was strengthened, while that of the unchosen direction was suppressed. These data suggest that motor control is inherently competitive, and that top-down signals play the part of a sculptor, by pruning overall activity and eliminating irrelevant activity to bias neuronal response toward the organism's current motor goal. These findings provide an intriguing analogy to the idea that we will outline below in that a top-down signal, reminiscent of cue 2 from the experiment described above, also shapes and prunes ongoing visual processes, by inhibiting alternative perceptual hypotheses in favor of the most likely interpretation of the current sensory input. What kind of signal would that be, and where would it be actively maintained to bias ongoing visual information processing? In the next section, we will argue that, from a functional point of view, visual context offers the necessary breadth and degree of abstraction to serve as such a top-down signal to bias a large set of ongoing perceptual processing toward a singular interpretation of the visual input. We further propose that the orbitofrontal cortex (OFC) might be the best candidate to represent and possibly maintain such information.

The dual role of context in generating predictions and in resolving competition

As opposed to highly artificial laboratory settings, objects in our environment never appear in isolation, but rather in typical contexts, together with other objects that usually share the same context, and often with typical spatial and temporal relations. The brain stores these statistical regularities learned with repetitive exposure, possibly in memory structures we have termed *context frames*.²⁰ Hence, when we know where we are, or where we are about to be, we can, by association, gain from information stored in memory to anticipate the range of the most likely objects, events, and occurrences. In the context of a bathroom, for example, one might expect items like a hairdryer, toothpaste, and a towel, and much less a computer, a phone, or a television. There is ample evidence from behavioral studies that contextual information indeed facilitates visual

perception (for a review, see Ref. 21). For example, the efficiency of visual search and recognition of objects greatly benefits from contextual information.^{22–25} Similarly, objects that violate contextual expectations in a scene are processed more slowly and with more errors.²⁶ Context information affords expectations that facilitate and accelerate interpretation of not only what is in front of us, but also about what is in our more global spatial and temporal vicinity.²⁰ Interestingly, similar ideas have also been expressed and tested in relation to lower-level processing (e.g., contour integration) in earlier parts of the visual cortex (for a review, see Ref. 27). One role of context, therefore, is preactivating likely object representations and thereby triggering and generating top-down expectations about identities and relations.

We propose that context also plays a second important role in perception, which is the resolution of competition. A simple example illustrates this idea: an input that initially maps to multiple hypotheses, such as a lamp, a mushroom, and an umbrella, will be biased toward the lamp in the context of a living room. In other words, contextual information might be put in the service of constraining the perceptual hypothesis space about the causes of sensory input and bias processing in favor of the most likely hypothesis, given the current context. Support for this idea would be to demonstrate that context facilitates perception under ambiguity, that is, a condition where multiple perceptual hypotheses about the causes of sensory input are likely to co-exist. Indeed, there is behavioral evidence that context enables an initially ambiguous object to become recognizable if another object that shares the same context is placed in a typical spatial relation to it.²⁰ Similarly, moving multistable stimuli can be stabilized by providing a specific motion of background,²⁸ and Necker cube alternations can be modulated by providing viewpoint information by the background.¹² Similarly, in binocular rivalry, context has been shown to influence perception of rivalrous stimuli.²⁹

A framework for top-down biased competition in visual recognition

To address the aspect of competition and its resolution, we propose extensions of a previously suggested top-down model.⁵ The model focused on processing of singular objects, and proposed that,

based on initial, gist-like information extracted from low spatial frequencies (LSFs), early visual areas send information directly to the OFC. In this region, predictions about the most likely interpretation of this input are activated and are then back-propagated to inferior temporal (IT^a) regions, where slowly arriving high spatial frequencies (HSFs) contribute to the exact representation of a specific exemplar (for illustration see Fig. 1, left side). By providing LSF-driven predictions, the top-down process guides (and significantly constrains) the analysis that needs to be performed by the bottom-up process. Those predictions highlight a subset of the object representations to be considered. That is, rather than all tens of thousands of objects in our memory having equal weight of probability to be considered against the input, only the handful of initial guesses need to be considered as likely interpretations. In that model, we did not assume a mutual competition in a narrow sense among neuronal representations; rather, the framework was more abstract, and top-down processes were assumed to provide information concerning what objects have to be considered from memory (i.e., where to allocate resources in posterior regions of the brain).

Here, we present critical extensions of those ideas, with an emphasis on context, rather than singular isolated objects, which we rarely encounter in the real world. We concede that bottom-up sensory input itself already constrains the options to be considered for further analysis, because it elicits a specific neuronal response. Thus, not all possible interpretations from memory need to be considered, but just a few candidates that resemble the neuronal response and the input it represents. Second, we propose that the OFC extracts a global context signal based on LSFs. Finally, we suggest that the role of top-down

^aHere, we use IT as an abbreviation, but note that the human homology to the monkey IT cortex is likely to be the cortex around the human lateral occipital cortex (for a review, see Ref. 30). We also note that studies in the context of binocular rivalry have demonstrated that most responses in the inferiotemporal cortex were correlated with the conscious percept (i.e., the currently dominant interpretation), while early extrastriate areas such as V4 and MT mostly do not cease firing completely when currently nondominant.¹¹ Therefore, the competitive processes might be more pronounced for those regions.

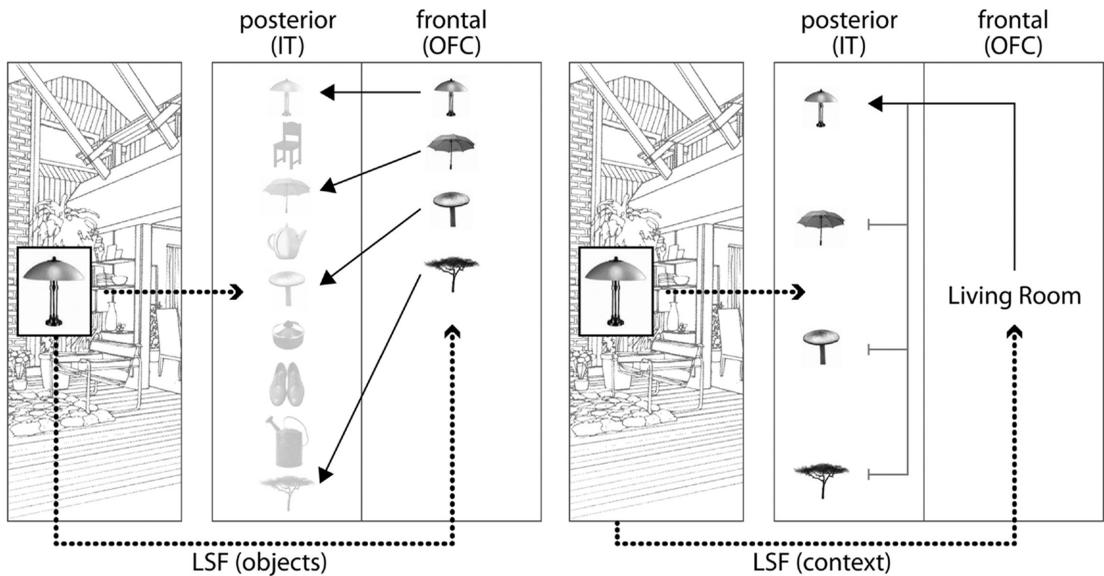


Figure 1. Left: The focus of the original model was on singular objects. We have suggested that early visual areas directly project to the orbitofrontal cortex (OFC), on the basis of gist-like information extracted from low spatial frequencies (LSFs). Here, predictions about the most likely interpretation of this input are activated and are then fed back to inferior temporal (IT) lobe regions. By providing LSF-driven predictions, the analysis that needs to be performed by the bottom-up process is significantly constrained. Here, we did not assume that all possible objects from memory compete for processing; rather, top-down information was considered to allocate resources toward only the most likely candidates or “slots” (therefore, object candidates are illustrated transparently). Right: Extensions of the model with focus on context and competition. Here, bottom-up sensory input already constrains the options to be considered for further analysis, as it elicits a specific neuronal response. Thus, not all possible interpretations or objects in the world need to be considered, but just a few candidates that resemble the neuronal response. As opposed to the extraction of a subset of object candidates, we suggest that the OFC computes a global context signal based on LSFs. The role of this top-down information is to bias the competition by enhancing relevant and suppressing irrelevant interpretations elicited by bottom-up information in posterior visual regions.

information, particularly implemented by the OFC, is to bias the competition elicited by bottom-up information in posterior regions of the brain (for illustration, see Fig. 1, right side). This is the first account where we explicitly consider the competition between alternative hypotheses and propose a mechanism for how it is resolved.

Context as a perceptual task and the OFC

We propose that context is used to constrain the space of perceptual hypotheses. But how is such an intuitively appealing idea implemented in the brain? Our laboratory has conducted some of the first steps toward answering this question. We compared brain activity for stimuli that are strongly linked to unique contextual associations (e.g., a bowling pin) with activity for stimuli that are only weakly related to unique context (e.g., a fly) in that they can be expected in many different settings.^{31,32} In several studies, we found that processing of

strong contextual associations elicits activity within the parahippocampal cortex (PHC), the retrosplenial complex, the medial prefrontal cortex/medial OFC, and the transverse occipital sulcus, which together we accordingly termed the *context network* (for a recent review, see Ref. 33). Here, we propose that rapidly extracted information allows the broad dissemination of global context signals that improve predictions and resolve competition. In our studies so far, however, the contrasts we computed were parametric (strong versus weak), and therefore it is conceivable that global context signals have been subtracted out. Therefore, the questions of how global contextual knowledge is represented in the brain and, most importantly, how it facilitates on-going processing of visual information remain to be addressed. We propose here that initial, coarse information triggers, in parallel, both the activation of multiple object hypotheses and the activation of a global context signal, which helps to quickly resolve

the competition among those multiple hypotheses. Such a global visual context signal is reminiscent of the signal of a *task set*. The psychological construct of a task refers to a configuration of perceptual, attentional, mnemonic, and motor processes, which contains rules of specific associations and cognitive operations for accomplishing a certain task or to achieve a specified goal (for a review, see Ref. 34). Such rule-specific activity was primarily identified in the dorsolateral prefrontal and ventrolateral prefrontal cortices.^{35,36} It was argued that, owing to extensive connectivity with both sensory and motor systems, frontal structures can simultaneously bias processing in different brain systems toward a common theme. We propose that the OFC is the analogy in the domain of perception, which provides such global, constraining and facilitating context signal that biases ongoing perceptual analysis elsewhere in the cortex.

The OFC is a multimodal, extensively connected association structure that receives input from multiple areas of the brain, including visceral, somatic, olfactory, and visual regions.³⁷ The OFC is also highly connected to the object-recognition regions in the IT cortex in monkeys and in humans (for reviews, see Refs. 38 and 39). Structurally, this massive interconnectedness makes the OFC a prime candidate for broadcasting information widely and for biasing ongoing processing in domain-specific areas. Indeed, several studies indicate that the OFC is relevant for processing of contextual top-down information. In a study with words, for instance, presented nouns were preceded by predicting or nonpredicting picture primes.⁴⁰ The authors found that, before the noun was presented, predictive primes triggered enhanced activation in the OFC, supporting its role in providing top-down-based interpretation of the subsequently presented noun in the fusiform word-form area. In another domain, it was demonstrated that the OFC is sensitive to contextual information that facilitates face categorization.⁴¹ In an fMRI study, individuals were asked to categorize faces' race based on only brief exposures. The faces were surrounded by Caucasian, neutral, or Asian scene contexts. Contextual congruent information facilitated categorization of faces by speeding up response times. In addition, the OFC exhibited highly sensitive, graded responses to the compatibility of facial and contextual cues. Specifically, the OFC showed linearly increasing responses as a face became more

Caucasian when in an American context, and linearly increasing responses as a face became more Asian when in a Chinese context.

OFC and LSF information

If the OFC extracts a context signal, to what kind of sensory information would it be sensitive? It has been demonstrated computationally that LSFs are generally sufficient for scene categorization,⁴² and numerous findings with humans have demonstrated that LSFs are relevant for rapid scene categorization or the extraction of the gist of a scene.^{43–47} Accordingly, it follows that a region that extracts and represents gist-like context signals should be preferentially activated by LSF information. Our laboratory has gathered evidence that LSF information is indeed favored by the OFC. In one experiment, we compared fMRI activation patterns elicited by images of objects filtered to contain primarily LSFs versus HSFs.⁴⁸ We controlled for recognition difficulty by matching reaction-time differences to those images, and found that LSF object images activated the OFC significantly more than HSF images. It is important to note that, although the stimuli used in this study were isolated objects devoid of any visual context, it is possible that the OFC is simply more sensitive to LSF information. Such an increased sensitivity might explain the (more) successful recognition. In other words, although the stimuli were isolated objects devoid of any surrounding (which we rarely encounter in the real world), it is possible to explain the findings by the possibility that the OFC is sensitive to LSFs, because it normally uses such information for the computation of the gist of a scene.

In addition, as shown with magnetoencephalography (MEG) signals in the same study, pictures in LSFs contributed to significantly higher synchrony between the OFC and the occipital visual areas and fusiform gyrus, suggesting a dedicated role of such information for top-down-guided processing of visual perception. In another study, we demonstrated indirect evidence for a preference of the OFC for LSFs by showing sensitivity to stimuli that are conveyed by the magnocellular pathway.⁴⁹ There is evidence from physiological findings that indicate that magnocellular pathways convey LSF information early and rapidly.⁴⁶ The cells of this pathway (M cells) have large receptive fields and fast conduction velocities, but are not color selective or able to resolve fine details.^{50,51} In contrast, cells of the

parvocellular pathway (P cells) have small receptive fields that are color sensitive and have low conduction velocities compared with that of the M cells.⁵² We capitalized upon this and created stimuli that were either M-biased (achromatic, low luminance contrast) or P-biased (isoluminant, high chromatic contrast) and found that the M-biased stimuli activated the OFC significantly more, while the P-biased stimuli activated the higher-level object-recognition regions to a greater extent. Finally, we recently demonstrated with fMRI that the OFC is engaged only when such LSF information matches preexisting memory associations (i.e., the stimuli resembled known visual objects as compared to meaningless, oriented gratings).⁵³ This further supports the proposal that the OFC uses LSF information as a guide to facilitate higher-level interpretation of sensory input. Taken together, the results from those studies strongly support the notion that the OFC preferentially processes coarse LSF information. Here, we propose that this information is used for extracting global context signals to bias and refine visual recognition of objects in the IT.

If so, LSF information has to be extracted rapidly, even before object recognition is finished. Individuals can rapidly understand a visual scene with brief exposure durations of around 100–150 ms, even in the near absence of attention.^{54,55} Indeed, there is ample evidence for temporal precedence of LSF processing over HSF in the visual cortex^{56–58} and that LSFs take precedence over HSFs during scene perception.⁵⁹ The authors of the latter study created hybrid images made of two superimposed scenes. One of those two scenes primarily contained LSFs, and the other scene HSFs. It was shown that, for short presentation times (30 ms), a participant's decision was significantly biased toward the scene in LSFs, while the opposite was found for HSF information at longer presentation times (150 ms). If the OFC processes and conveys such global context signals, it must do so rapidly and in parallel, or even before object recognition is completed. Indeed, we have shown with MEG that OFC activity precedes the activity in the IT by 50–80 ms.⁴⁸ Such an early activation of the OFC suggests that it is well positioned to provide top-down information that facilitates and constrains the perceptual analysis required for the recognition of individual objects in posterior regions of the brain.

A context-related functional continuum in the OFC

In our previous studies, two main foci of activation in the OFC were observed. In the recognition of individual objects, it was primarily the inferior part of the OFC that showed selectivity for recognition and for LSF information.^{48,54,60} On the other hand, in studies of contextual recognition and associations, it was a more medial part of the OFC that showed preferred activation, not overlapping with the inferior OFC site.³³ As described above, we originally proposed that the inferior OFC sends information back to the IT about the most likely object interpretations (e.g., drill, gun, hairdryer), either in the form of actual representations or rather “look-up” information about what representations to activate in the IT. The medial OFC activation, on the other hand, was proposed to be more directly related to contextual information proper, as part of the broader contextual network involving the PHC and the retrosplenial cortex. Here, we suggest an alternative whereby perceptual hypotheses about an object identity are activated in the IT as a result of a local process confined to the visual cortex. In this alternative interpretation, both inferior and medial OFC sites are suggested to be the ends of an anatomical and functional continuum that is involved in the generation of a general context signal that is extracted from individual objects, as well as from complete scenes, and is disseminated downstream for resolving competitions.

Supporting evidence for the idea of a continuum in the OFC comes from a study that used contextually associative single objects and reported OFC activation between the medial and the inferior sites described above.⁶¹ However, there is also evidence against a continuum, which comes from synchrony analyses⁴⁸ where the inferior OFC shows significant phase locking with the IT around 130 ms from stimulus onset, whereas the medial OFC shows significant phase locking only around 370 ms, though it is still conceivable that context is extracted and disseminated earlier than synchrony is observed. In either case, given how complicated the OFC is for exact localizations using fMRI (owing to susceptibility artifacts), it is premature to draw strong functional conclusions based on slight anatomical differences, and this particular issue will have to be resolved in the future.

The OFC and reinforcement learning

It should also be noted that parts of the OFC (primarily the medial OFC, also referred to as ventromedial prefrontal cortex) were found to be activated in a plethora of studies of reward learning in rats, monkeys, and humans.^{62–64} It has therefore often been suggested that the activation of the OFC in vision could be related to recognition success.⁶⁵ Indeed, there is evidence that OFC activity is correlated with parameters such as subjective sense of perceptual coherence, confidence, and uncertainty in perceptual decision making.^{66–68} However, we controlled for recognition success by matching reaction times between LSF and HSF pictures.⁴⁸ Additionally, the exact role of the OFC in reinforcement learning is currently under active debate. Several studies point to the interpretation that the representation in the OFC is much more global than previously assumed. Similar to other frontal areas, it has been suggested that the OFC is more related to representing task-relevant and context-sensitive information that guides value-based decision making.⁶⁹ Recently, a unifying theory of the OFC has been proposed that suggests that this region constitutes a cognitive map of *task space* that biases reward-related information processing elsewhere in the cortex.⁷⁰ These ideas bear striking similarity to our proposal of the OFC as processing global context signals to efficiently perform a perceptual task (i.e., object recognition) that is accomplished elsewhere in the cortex (the IT). Certainly, the exact role of the OFC and the nature of its representations remain to be specified, and, just like other frontal structures, may differ depending on the current task.⁷¹ A promising avenue for the future is therefore to elaborate the functional and neural common denominators between reward processing and visual recognition and the orchestrating role of the OFC.

Summary and outlook

While briefly overviewing proposals pertaining to the crucial role of top-down predictions in perception, we focused on the component of competition in recognition. We postulated that, while several perceptual hypotheses about the causes of sensory input are elicited early and simultaneously, the role of top-down signals is to ultimately select relevant and suppress irrelevant hypotheses. We proposed that visual context information provides global top-down signals that can bias ongoing processing to-

ward the most likely perceptual hypotheses. In line with theories of integrative functions of the prefrontal cortex,⁷² we suggested that such a signal is provided by frontal structures, primarily the OFC. Future studies have to address how this biased competition is implemented mechanistically. A potential candidate mechanism is the ability of neuronal populations to increase or decrease the synchrony in their firing. For example, in a MEG study from our own laboratory, we demonstrated enhanced β -band activity for strongly contextual compared with weakly contextual stimuli in the context network.⁷³ Importantly, an elaborated understanding of such context-based perception on neural grounds will also shed light on a question that has not been discussed in other predictive coding frameworks:⁷⁴ namely, where predictions originate, and how top-down processing might be initiated.

Acknowledgments

This project was supported by a Max Planck post-doctoral fellowship (Dr. Med. Anneliese & DSc Dieter Pontius Foundation) (ST) and The Israeli Center of Research Excellence in Cognitive Sciences ICORE Grant No. 51/11 (MB). We thank K. Kveraga for valuable comments on an earlier version of the manuscript.

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Kok, P., M. Failing & F.P. de Lange. 2013. Prior expectations evoke stimulus templates in the primary visual cortex. *J. Cogn. Neurosci.* **26**: 1546–1554.
2. Summerfield, C. & E. Kochlin. 2008. A neural representation of prior information during perceptual inference. *Neuron* **59**: 336–347.
3. Rao, R.P. & D.H. Ballard. 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* **2**: 79–87.
4. Lee, T.S. & D. Mumford. 2003. Hierarchical Bayesian inference in the visual cortex. *J. Opt. Soc. Am. A.* **20**: 1434–1448.
5. Bar, M. 2003. A cortical mechanism for triggering top-down facilitation in visual object recognition. *J. Cogn. Neurosci.* **15**: 600–609.
6. Desimone, R. & J. Duncan. 1995. Neural Mechanisms of selective visual-attention. *Annu. Rev. Neurosci.* **18**: 193–222.
7. Moran, J. & R. Desimone, 1985. Selective attention gates visual processing in the extrastriate cortex. *Science* **229**: 782–784.

8. Kastner, S., P. De Weerd, R. Desimone & L.G. Ungerleider. 1998. Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science* **282**: 108–111.
9. Kastner, S., M.A. Pinsk, P. De Weerd, *et al.* 1999. Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron* **4**: 751–761.
10. Leopold, D. & N.K. Logothetis. 1999. Multistable phenomena: changing views in perception. *Trends Cogn. Sci.* **3**: 254–264.
11. Leopold, D.A. & N.K. Logothetis. 1996. Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature* **379**: 549–553.
12. Sundareswara, R. & P.R. Schrater. 2008. Perceptual multistability predicted by search model for Bayesian decisions. *J. Vis.* **8**(5): 12, 1–19.
13. van Ee, R., W.J. Adams & P. Mamassian. 2003. Bayesian modeling of cue interaction: bistability in stereoscopic slant perception. *J. Opt. Soc. Am. A.* **20**: 1398–1406.
14. Ratcliff, R. & G. McKoon. 2008. The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* **20**: 873–922.
15. Voss, A., M. Nagler & V. Lerch. 2013. Diffusion models in experimental psychology: a practical introduction. *Exp. Psychol.* **60**: 385–402.
16. Mulder, M.J., E.J. Wagenmakers, R. Ratcliff, *et al.* 2012. Bias in the brain: a diffusion model analysis of prior probability and potential payoff. *J. Neurosci.* **32**: 2335–2343.
17. Hanks, T.D., M.E. Mazurek, R. Kiani, *et al.* 2011. Elapsed decision time affects the weighting of prior probability in a perceptual decision task. *J. Neurosci.* **31**: 6339–6352.
18. Cisek, P. 2007. Cortical mechanisms of action selection: the affordance competition hypothesis. *Philos. Trans. R. Soc. Lond. B* **362**: 1585–1599.
19. Cisek, P. & J.F. Kalaska. 2005. Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. *Neuron* **45**: 801–814.
20. Bar, M. & S. Ullman. 1996. Spatial context in recognition. *Perception* **25**: 343–352.
21. Bar, M. 2004. Visual objects in context. *Nat Rev. Neurosci.* **5**: 617–629.
22. Boyce S.J., A. Pollatsek & K. Rayner. 1989. Effect of background information on object identification. *J. Exp. Psychol. Hum. Percept. Perform.* **15**: 556–566.
23. Davenport, J.L. & M.C. Potter. 2004. Scene consistency in object and background perception. *Psychol. Sci.* **15**: 559–564.
24. Chun, M.M. & Y. Jiang. 1998. Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cogn. Psychol.* **36**: 28–71.
25. Palmer, S.E. 1975. The effects of contextual scenes on the identification of objects. *Mem. Cogn.* **3**: 519–526.
26. Biederman, I., A.L. Glass & E.W. Stacy. 1973. Searching for objects in real-world scenes. *J. Exp. Psychol. Gen.* **97**: 22–27.
27. Gilbert, C.D. & W. Li. 2013. Top-down influences on visual processing. *Nat. Rev. Neurosci.* **14**: 350–363.
28. Dawson, M.R. 1987. Moving contexts do affect the perceived direction of apparent motion in motion competition displays. *Vision Res.* **27**: 799–809.
29. Sobel, K.V. & R. Blake. 2002. How context influences predominance during binocular rivalry. *Perception* **31**: 813–824.
30. Orban, G.A., D. Van Essen & W. Vanduffel. 2004. Comparative mapping of higher visual areas in monkeys and humans. *Trends Cogn. Sci.* **8**: 315–324.
31. Bar, M. & E. Aminoff. 2003. Cortical analysis of visual context. *Neuron* **38**: 347–358.
32. Aminoff, E.M., N. Gronau & M. Bar. 2007. The parahippocampal cortex mediates spatial and nonspatial associations. *Cereb. Cortex* **17**: 1493–1503.
33. Aminoff, E.M., K. Kveraga & M. Bar. 2013. The role of the parahippocampal cortex in cognition. *Trends Cogn. Sci.* **17**: 379–390.
34. Sakai, K. 2008. Task set and prefrontal cortex. *Annu. Rev. Neurosci.* **31**: 219–245.
35. Sakagami, M. & K. Tsutsui. 1999. The hierarchical organization of decision making in the primate prefrontal cortex. *Neurosci. Res.* **34**: 79–89.
36. Sakai, K. & R.E. Passingham. 2003. Prefrontal interactions reflect future task operations. *Nat. Neurosci.* **6**: 75–81.
37. Barbas, H. 2000. Connections underlying the synthesis of cognition, memory, and emotion in primate prefrontal cortices. *Brain Res. Bull.* **52**: 319–330.
38. Cavada, C., T. Company, J. Tejedor, *et al.* 2000. The anatomical connections of the macaque monkey orbitofrontal cortex. A review. *Cereb. Cortex* **10**: 220–242.
39. Kringelbach, M.L. & E.T. Rolls. 2004. The functional neuroanatomy of the human orbitofrontal cortex: Evidence from neuroimaging and neuropsychology. *Progress Neurobiol.* **72**: 341–372.
40. Dikker, S. & L. Pyllkkänen. 2013. Predicting language: MEG evidence for lexical preactivation. *Brain Lang.* **127**: 55–64.
41. Freeman, J.B., Y. Ma, M. Barth, *et al.* 2015. The neural basis of contextual influences on face categorization. *Cereb. Cortex* Advanced Online Publication. **25**: 415–422.
42. Oliva, A. & A. Torralba. 2001. Modeling the shape of a scene: a holistic representation of the spatial envelope. *Int. J. Comput. Vision* **42**: 145–175.
43. Kihara, K. & Y. Takeda. 2010. Time course of the integration of spatial frequency-based information in natural scenes. *Vision Res.* **50**: 2158–2162.
44. De Cesarei, A. & G.R. Loftus. 2011. Global and local vision in natural scene identification. *Psychon. Bull. Rev.* **18**: 840–847.
45. Mu, T. & S. Li. 2013. The neural signature of spatial frequency-based information integration in scene perception. *Exp. Brain Res.* **227**: 367–377.
46. Musel, B., C. Bordie, M. Dojat, *et al.* 2013. Retinotopic and lateralized processing of spatial frequencies in human visual cortex during scene categorization. *J. Cogn. Neurosci.* **25**: 1315–1331.
47. Schyns, P.G. & A. Oliva. 1994. From blobs to boundary edges: evidence for time- and spatial- dependent scene recognition. *Psychol. Sci.* **5**: 195–200.
48. Bar, M., K.S. Kassam, A.S. Ghuman, *et al.* 2006. Top-down facilitation of visual recognition. *Proc. Natl. Acad. Sci. U.S.A.* **103**: 449–454.
49. Kveraga, K., J. Boshyan & M. Bar. 2007. Magnocellular projections as the trigger of top-down facilitation in recognition. *J. Neurosci.* **27**: 13232–13240.

50. Kaplan, E. 2004. "The M, P, and K pathways of the primate visual system." In *The Visual Neuroscience*. L. M. Chalupa & J. S. Werner, Eds.: 481–494. Cambridge, MA: The MIT Press.
51. Kaplan, E. & R.M. Shapley. 1986. The primate retina contains two types of ganglion cells, with high and low contrast sensitivity. *Proc. Natl. Acad. Sci. U.S.A.* **83**: 2755–2757.
52. Wiesel, T.N. & D.H. Hubel. 1966. Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. *J. Neurophysiol.* **29**: 1115–1156.
53. Wassle, H. & B.B. Boycott. 1991. Functional architecture of the mammalian retina. *Physiol. Rev.* **71**: 447–480.
54. Chaumon, M., K. Kveraga, L.F. Barrett & M. Bar. 2014. Visual predictions in the orbitofrontal cortex rely on associative content. *Cereb. Cortex* **24**: 2899–2907.
55. Li, F.F., R. VanRullen, C. Koch & P. Perona. 2002. Rapid natural scene categorization in the near absence of attention. *Proc. Natl. Acad. Sci. U.S.A.* **99**: 9596–9601.
56. Thorpe, S.J., D. Fize & C. Marlot. 1996. Speed of processing in the human visual system. *Nature* **381**: 520–522.
57. Breitmeyer, B.G. 1975. Simple reaction time as a measure of the temporal response properties of the transient and sustained channels. *Vision Res.* **15**: 1411–1412.
58. Bullier, J. 2001. Integrated model of visual processing. *Brain Res. Rev.* **36**: 96–107.
59. Van Essen, D.C. & E.A. Deyoe. 1995. "Concurrent processing in the primate visual cortex." In *The Cognitive Neurosciences* M. Gazzaniga, Ed.: 383–400. Cambridge: Bradford Book.
60. Bar, M., R. Tootell, D. Schacter, *et al.* 2001. Cortical mechanisms of explicit visual object recognition. *Neuron* **29**: 529–535.
61. Shenhav, A., L.F. Barrett & M. Bar. 2012. Affective value and associative processing share a cortical substrate. *Cogn. Affect. Behav. Neurosci.* **13**: 46–59.
62. Gottfried, J.A., J. O'Doherty & J.R. Dolan. 2003. Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science* **301**: 1104–1107.
63. Kahnt, K., J. Heinzle, S.Q. Park & J.D. Haynes. 2010. The neural code of reward anticipation in human orbitofrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* **107**: 6010–6015.
64. Takahashi, Y.K., C.Y. Chang, F. Lucantonio, *et al.* 2013. Neural estimates of imagined outcomes in the orbitofrontal cortex drive behavior and learning. *Neuron* **16**: 507–518.
65. O'Shea, J. & V. Walsh. 2006. Cognitive Neuroscience: trickle-down theories of vision. *Curr. Biol.* **166**: 206–209.
66. Horr, N.K., C. Braun & K.G. Volz. 2014. Feeling before knowing why: the role of the orbitofrontal cortex in intuitive judgments—an MEG study. *Cogn. Affect. Behav. Neurosci.* **14**: 1271–1285.
67. Hsu, M., R. Bhatt, D. Adolphs, *et al.* 2005. Neural systems responding to degrees of uncertainty in human decision making. *Science* **210**: 1680–1683.
68. Volz, K.G., R. Rübsemann & D.Y. von Cramon. 2008. Cortical regions activated by the subjective sense of perceptual coherence of environmental sounds: a proposal for a neuroscience of intuition. *Cogn. Affect. Behav. Neurosci.* **8**: 318–328.
69. Wilson, R.C., Y.K. Takahashi, G. Schoenbaum & Y. Niv. 2014. Orbitofrontal cortex as a cognitive map of task space. *Neuron* **81**: 267–279.
70. Schoenbaum, G. & B. Setlow. 2001. Integrating orbitofrontal cortex into prefrontal theory: common processing themes across species and subdivision. *Learn Mem.* **8**: 134–147.
71. Duncan, J. & A.M. Owen. 2000. Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci.* **23**: 475–483.
72. Miller, E.K. & J.D. Cohen. 2001. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* **24**: 167–202.
73. Kveraga, K., A.S. Ghuman, K.S. Kassam, *et al.* 2011. Early onset of neural synchronization in the contextual associations network. *Proc. Natl. Acad. Sci. U.S.A.* **108**: 3389–3394.
74. Friston, K. 2005. A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **360**: 815–836.