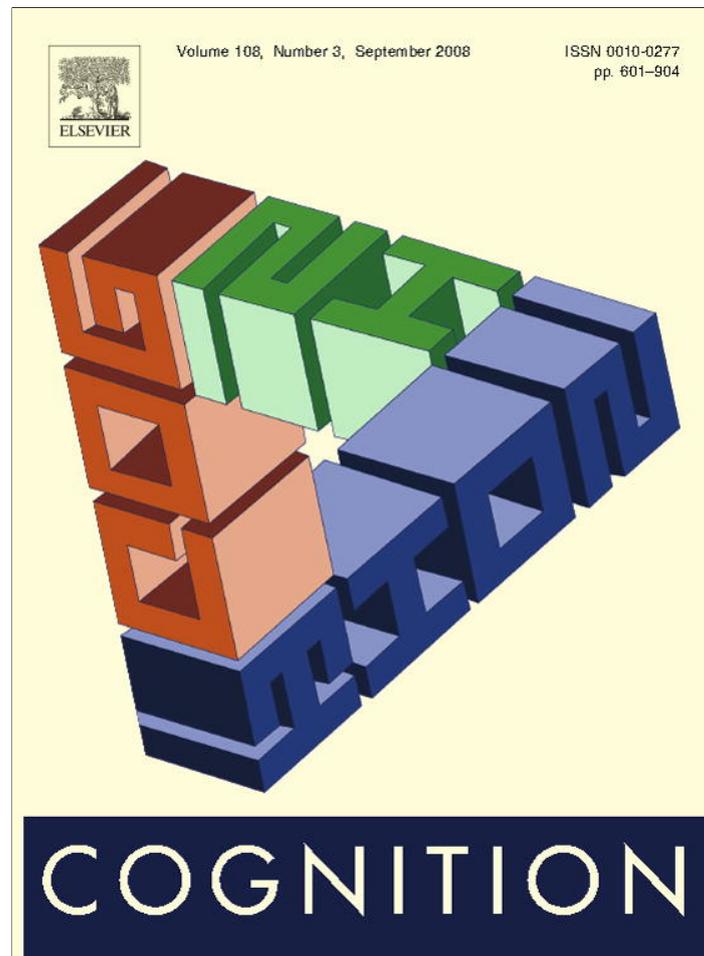


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Predictive coding explains binocular rivalry: An epistemological review

Jakob Hohwy^{a,*}, Andreas Roepstorff^{b,c}, Karl Friston^d

^a Department of Philosophy, Monash University, Australia

^b Department of Social Anthropology, University of Aarhus, Denmark

^c Danish National Research Foundation's Centre for Functionally Integrative Neuroscience, University of Aarhus, Denmark

^d The Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, UK

ARTICLE INFO

Article history:

Received 19 September 2007

Revised 18 April 2008

Accepted 22 May 2008

Keywords:

Binocular rivalry
Consciousness
Predictive coding
Empirical Bayes
Perceptual inference
Learning
Free-energy
Psychophysics
Neurophysiology

ABSTRACT

Binocular rivalry occurs when the eyes are presented with different stimuli and subjective perception alternates between them. Though recent years have seen a number of models of this phenomenon, the mechanisms behind binocular rivalry are still debated and we still lack a principled understanding of why a cognitive system such as the brain should exhibit this striking kind of behaviour. Furthermore, psychophysical and neurophysiological (single cell and imaging) studies of rivalry are not unequivocal and have proven difficult to reconcile within one framework. This review takes an epistemological approach to rivalry that considers the brain as engaged in probabilistic unconscious perceptual inference about the causes of its sensory input. We describe a simple empirical Bayesian framework, implemented with predictive coding, which seems capable of explaining binocular rivalry and reconciling many findings. The core of the explanation is that selection of one stimulus, and subsequent alternation between stimuli in rivalry occur when: (i) there is no single model or hypothesis about the causes in the environment that enjoys both high likelihood and high prior probability and (ii) when one stimulus dominates, the bottom-up, driving signal for that stimulus is explained away while, crucially, the bottom-up signal for the suppressed stimulus is not, and remains as an unexplained but explainable prediction error signal. This induces instability in perceptual dynamics that can give rise to perceptual transitions or alternations during rivalry.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

If one stimulus is shown to one eye and another stimulus to the other, then subjective experience alternates between them. For example, when an image of a house is presented to one eye and an image of a face to the other, then subjective experience alternates between the house and the face. This is known as binocular rivalry. Binocular rivalry is a challenge to our understanding of the visual system, and it is of special importance for studies of phenomenal consciousness in humans and monkeys, because the stimulus presented to subjects can be held constant

while the phenomenal percept changes (Frith, Perry, & Lumer, 1999; Koch, 2004).

There have been many empirical studies of binocular rivalry but the data they produce are conflicting and it is very difficult to give them an unequivocal interpretation. A number of proposals have been made but the neurocognitive mechanism that explains this striking visual effect remains unresolved (for reviews and overviews, see Alais & Blake, 2005; Blake & Logothetis, 2002; Leopold & Logothetis, 1999; Tong, Meng, & Blake, 2006). There are recent formal models that can explain a growing number of psychophysical findings and which fit with a range of neurophysiological facts (Koene, 2006; Moreno-Bote, Rinzal, & Rubin, 2007; Noest, van Ee, Nijs, & van Wezel, 2007; Wilson, 2007), and there is a general trend towards approaches that integrate top-down and bottom-up

* Corresponding author. Tel.: +61 3 9905 3208; fax: +61 3 9905 3206.
E-mail addresses: Jakob.Hohwy@arts.monash.edu.au, j.hohwy@gmail.com (J. Hohwy).

processes in the brain (Tong et al., 2006); however, we believe the study of binocular rivalry may benefit from a principled theoretical framework that can motivate these new developments. Most approaches to rivalry stress the role of inhibition, adaptation and stochastic noise. We take the approach of epistemology—the theory of knowledge—to go behind these approaches and ask the more fundamental theoretical question: “why should a *perceptual* system, such as the brain, have and exploit such mechanisms in the first place?” The motivation behind this approach is the idea that binocular rivalry is an epistemic response to a seemingly incompatible stimulus condition where two distinct objects occupy the same spatiotemporal location. This paves the way for the description of a principled and unified account of rivalling perceptions under dichoptic viewing conditions. Our intent is thus not to add new data to the burgeoning class of data already in hand concerning binocular rivalry but to describe a unifying framework for it.

There is growing support of the idea that the brain is an inference machine, or hypothesis tester, which approaches sensory data using principles similar to those that govern the interrogation of scientific data. In this view, perception is a type of unconscious inference. As Gregory states:

[P]erceptions are hypotheses, predicting unsensed characteristics of objects, and predicting in time, to compensate neural signalling delay (discovered by von Helmholtz in 1850), so ‘reaction time’ is generally avoided, as the present is predicted from delayed signals [...] Further time prediction frees higher animals from the tyranny of control by reflexes, to allow intelligent behaviour into anticipated futures (1997, p. 1122).

This view goes back at least to von Helmholtz (1860) and has been expressed with increasing finesse since that time (Gregory, 1998; MacKay, 1956; Neisser, 1967; Rock, 1983). More recently, it has been proposed that this intuitive idea can be captured in terms of hierarchical Bayesian inference, using generative models with predictive coding or free-energy minimisation; and that this is the main neurocomputational principle for the brain’s perception of the environment as well as its learning of new contingencies (Ballard, Hinton, & Sejnowski, 1983; Dayan, Hinton, Neal, & Zemel, 1995; Friston, 2002; Friston, 2003; Friston, 2005; Friston & Stephan, 2007; Kawato, Hayakawa, & Inui, 1993; Kersten, Mamassian, & Yuille, 2004; Knill & Pouget, 2004; Mumford, 1992; Murray, Schrater, & Kersten, 2004; Rao & Ballard, 1999).

Our proposal is that this general theoretical framework, in its more recent incarnations, provides the computational mechanism that best explains binocular rivalry and reconciles conflicting evidence. We set out some core properties of predictive coding, show how it explains binocular rivalry, and relate the explanation to a number of empirical neurophysiological, imaging and psychophysical findings concerning binocular rivalry. A Bayesian framework has been suggested recently for bistable perception (slant rivalry) (van Ee, 2003), however, though this framework is congenial to the account given here, it is not couched in terms of generative models, predictive coding and empirical Bayes. As we shall see, in its more complex version

Bayesian theory has great explanatory promise. Our account has more in common with an earlier model by Dayan (1998) that uses explicit generative models (A further recent study of bistable perception (monocular rivalry) by Knapen, Kanai, Brascamp, van Boxtel, & van Ee, 2007, seems to count against the use of generative models; we discuss this further in Section 6).

2. Core properties of predictive coding

A core task for the brain is to represent the environmental causes of its sensory input. This is computationally difficult; it is difficult to compute the causes when only the effects are known: as Hume (1739–40) reminded us, causes and effects are distinct existences and, in principle, many different environmental events could be causes of the same sensory effect. Conversely, the same environmental causes can occur in different contexts, so the same environmental event can be the cause of many different sensory effects.

Hierarchical Bayesian inference, using generative models, can manage or finesse these difficulties by harnessing the causal structure of sensory stimuli to furnish formal constraints on the mapping between cause and effect. Rather than trying to work backwards from sensory effects to environmental causes, neuronal computational systems work with models, or as we shall say *hypotheses*, that *predict* what the sensory input should be, if it were really caused by certain environmental events. The hypothesis that generates the best predictions then determines perceptual content. The hierarchical inversion of the generative models needed to finesse this inverse problem can be reduced to quite simple processes that, in principle, can be implemented by the brain. In fact, one can predict many anatomical and physiological aspects of the brain by assuming it is inverting a hierarchical model of its sensory input (e.g., Friston, 2003; Friston, 2005; Friston & Stephan, 2007). Below we give a simplified description, in basic Bayesian terms of *prior probability* and *likelihoods*, of some of the core properties of a system, employing predictive coding or free-energy minimisation, that is involved in solving Bayesian perceptual inference. (See Fig. 1).

2.1. Bayesian perceptual inference

According to this kind of Bayesian theory, the hypothesis with the highest posterior probability (i.e., most probable given the input) wins and gets to determine the perceptual content of the system. The posterior probability depends on the likelihood (i.e., how well the hypothesis predicts the input); and on the prior probability of the hypothesis (i.e., how probable the hypothesis was before the input) (Friston, 2002; Kersten et al., 2004; Murray, Kersten, Olshausen, Schrater, & Woods, 2002). These prior expectations are constructed hierarchically and are context-sensitive. For example, if the hypothesis is that visual input is caused by a box, then it is possible to predict, on the basis of that hypothesis, what the input is going to be as one moves around it. If the prediction turns out to be right, and if the presence of a box is otherwise probable,

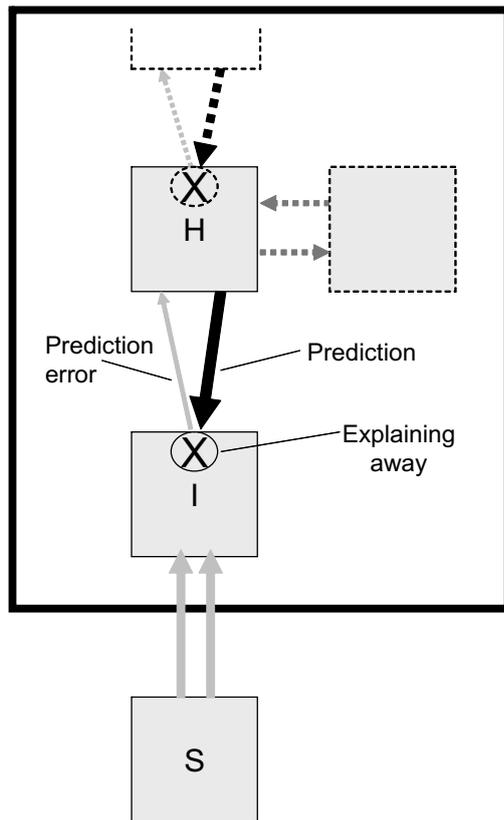


Fig. 1. Simplified schematic of a pair of cortical levels based on generative models. An open representational system like the brain (indicated with the black box) must perform perceptual inference about the environmental causes (S) of its sensory input (I). Higher level models or hypotheses (H) about the possible cause are used to generate top-down predictions (dark arrows) about the evolving input, which in turn explain away bottom-up sensory signal (light arrows), leaving only the prediction error as bottom-up signal to be explained away. Subsequent updating of H should further minimise prediction error.

then the probability for the hypothesis that it is a box goes up. If there are no better hypotheses in play, then this hypothesis wins and the perceptual inference will be that the environmental cause is indeed a box (These examples turn on visual perception; Bayesian frameworks are also often used in multisensory contexts, where one modality provides prior constraints for the other, e.g., Alais & Burr, 2004; Ernst & Banks, 2002).

2.2. Explaining away of bottom-up signal

The system tries to match bottom-up or driving signals, caused by objects and properties in the environment, with top-down predictions. If the predictions are good, then the bottom-up signal will be explained away such that only the discrepancies between prediction and driving signal – the prediction error signal – remains as a bottom-up signal. As predictions get better, there will be less error signal associated with a given stimulus at relatively lower levels in the neural system (Friston, 2005; Yuille & Kersten, 2006). This suppression of best predicted input will be central for the explanation of rivalry.

2.3. Hierarchy

The cognitive system is ordered hierarchically in levels. For any pair of levels, the higher level will have hypotheses that predict the driving bottom-up error signal from the lower level. The higher level will itself provide error signals for a yet higher level. The lower level of the pair will be higher level for a yet lower level. Priors come from higher levels, as in empirical Bayes (for a relevant predictive coding study of face perception, see Summerfield, Egner, Mangels, & Hirsch, 2005). Perceptual inference about different hierarchically organised attributes of the visual scene are made at different levels (Friston, 2005). It is not unusual for theories of visual processing and of binocular rivalry in particular (e.g., Freeman, 2005; Wilson, 2003) to be hierarchical, the crucial point here concerns the computational implications of hierarchical levels and the fact that these provide formal constraints on the generative models that make them empirical Bayes models.

2.4. Updating hypotheses/perceptual learning

In a hierarchical setting, that uses empirical Bayes, priors are not extracted directly from the natural scene statistics, nor are they free parameters. They emerge naturally on interaction with the world as learning suppresses prediction errors at all levels of a hierarchical model. The hierarchical nature of these models is central to empirical Bayes because priors on lower levels are themselves constrained by, and accountable to, higher levels. Empirical Bayes is a powerful and ubiquitous inference framework that arises in many contexts; ranging from the distinction between fixed and random effects models in statistical analysis of data to hierarchical models that have been proposed for perceptual inference in the brain.

The prediction error signal plays a crucial role in inference since it helps to update hypotheses at higher levels, such that better predictions can be issued and the prediction error continually minimised. Hypotheses also are context-sensitive, via modulation from other hypotheses at the same or higher levels. Thus predictions about sensory input can improve as more hypotheses about the context of the stimulus are generated.

2.5. Prediction error and free-energy minimisation

Terms like ‘predictions’ and ‘hypotheses’ sound rather intellectualist when it comes to basic perceptual inference. But at its heart the only processing aim of the system is simply to minimise prediction error or free-energy, and indeed, the talk of hypotheses and predictions can be translated into such a less anthropomorphic framework.

The notion of free-energy derives from statistical physics and is used widely in machine learning to convert difficult integration problems, inherent in inference, into easier optimisation problems. Free-energy is essentially a mathematical concept that is a function of probability distributions (like entropy, information or surprise) (a useful introduction can be found in Dayan & Abbott, 2001, Chap. 10). This optimisation or free-energy minimisation can, in principle, be implemented using relatively simple neuronal

infrastructures. The free-energy represents a bound on the surprise inherent in any exchange with the environment, under expectations encoded by its state or configuration. A system can minimise free-energy by changing its configuration to change the way it samples the environment, or to change its expectations. These changes correspond to action and perception, respectively, and lead to an adaptive exchange with the environment that is characteristic of biological systems (Friston & Stephan, 2007 contains numerous further references and discussion). In short, any change to the brain's state or connection parameters that reduces free-energy renders sensory input less surprising. If we discount uncertainty about the states, when optimising the parameters (and vice versa) it is fairly easy to show that the free-energy is the sum of squared prediction errors, weighted by their estimated precision (op.cit.).

2.6. Top-down and bottom-up

What ultimately determines the resulting conscious perception is the best hypothesis: the one that makes the best predictions and that, taking priors into consideration, is consequently assigned the highest posterior probability. The model is however interactionist (in the terminology of Tong, 2003) since it is essential to appreciate both activity at relatively higher levels where predictions are made and the nature of lower level activity in order to have a theoretical framework for understanding either. This accords with recent approaches to binocular rivalry that also stress the interactionist element (e.g., Blake & Logothetis, 2002; Nguyen, Freeman, & Alais, 2003; Tong et al., 2006). The interactionist perspective holds for any pair of levels that communicate with each other as top-down and bottom-up throughout the brain. We will assume that a percept corresponds to a prediction. It is important to note that, in a hierarchical setting, predictions exist at all levels of the hierarchy and, implicitly, all levels of perceptual detail. This suggests that the percept is encoded in a distributed way and accords with related notions of phenomenal perception and their neurophysiological underpinnings (e.g., Zeki, 2003).

3. Two problems concerning rivalry: Selection and alternation

In dichoptic viewing conditions, where one stimulus is shown to one eye and another to the other eye, binocular matching fails because two different objects seem to occupy the same spatiotemporal position (Blake & Boothroyd, 1985). The epistemological task for the system, given this incompatible or “un-ecological” condition is then to explain the combined bottom-up signal stemming from the two stimuli: it does this rather elegantly by selecting only one stimulus at a time and then alternating between them. To account for binocular rivalry, two things must then be explained (this important duality is also emphasised in Noest et al., 2007):

First, the *selection problem*: why is there a perceptual decision to select one stimulus for perception rather than the other, and, further why is one of the two stimuli

selected rather than some conjunction or blend of them? We propose a solution for this in Section 4.

Second, the *alternation problem*: why does perceptual inference alternate between the two stimuli rather than stick with the selected one? We propose a solution for this in Section 5.

3.1. The current rivalry debate

In the last decade or so there has been two main positions on binocular rivalry. One widely held view stressed low-level inter-ocular competition among monocular neurons in early visual cortex. Another view (to a large extent triggered by Logothetis, Leopold, & Sheinberg, 1996) stressed high-level competition among incompatible patterns. There are signs that these may merge in a view that stresses neural competition at multiple levels of the visual system (for review of this development, see Tong et al., 2006). A particular merger of top-down and bottom-up mechanisms in rivalry is central to our proposal too (for another detailed proposal, see Alais & Melcher, 2007).

The core of these various approaches to rivalry is that selection and alternation in rivalry must be explained in terms of two mechanisms: *inhibition* of the incoming signal from the stimulus which is not dominant, which is meant to explain selection; and *adaptation* of the inhibitory influence of the relevant neural populations, which is meant to explain alternation. Though the brain does exhibit both inhibition and adaptation, this is a rather *a priori* characterisation of the mechanism behind rivalry: any account of binocular rivalry will have elements of inhibition and adaptation; otherwise a pattern of dominance vs. non-dominance of perceptual content can hardly occur. Here, we take a more epistemological view and ask why a representational system such as the brain should have general computational and statistical properties such that it will exhibit rivalry in dichoptic viewing conditions? A model based on predictive coding provides a parsimonious and principled answer to this question, and explains in one move why there should be both inhibition and adaptation in dichoptic viewing.

4. A Bayesian approach to the selection problem

Assume the stimuli are a house and a face and that the percept currently experienced by the subject is the face. Then the question, from a Bayesian perspective, is why the face hypothesis (F) has the highest probability, given the conjoint evidence (I) of a house and a face. The question splits into two: (i) why is F favoured over the hypothesis that it is a house (H)? (ii) Why is F selected over some kind of conjunctive or blended hypothesis that it is a ‘house-face’ (F AND H)? (see Fig. 2).

Assuming the contents of the stimuli are independent, F and H explain the evidence equally well even though they each are unable to account for a large part of it. That is to say, they are roughly equally likely. Given equal likelihood, the perceptual inference will tend to depend on the prior probability of the hypotheses. If, for some reason, F has a

Selection as driven by priors	
Input: I	
Hypotheses	F+H: "It's a face-house" H: "It's a house" F: "It's a face"
Likelihoods	$P(I/F) = P(I/H) < P(I/F+H)$
Priors	$P(F) > P(H) \gg P(F+H)$
Perceptual inference	$P(F/I) > P(H/I) > P(F+H/I)$ 

Fig. 2. Simplified Bayesian account of the selection of one stimulus rather than the other and rather than a blend: no hypothesis enjoys both high likelihood and high prior probability, hence the hypothesis with the highest prior can win (as long as the conjunctive hypothesis does have very low likelihood).

higher prior, then it will be selected for perceptual dominance.

For the conjoint hypothesis *F AND H* the situation is the opposite. It has a higher likelihood than *F* or than *H* because, in principle, it can predict much more of the evidence. But *F AND H* has a much lower prior than both *F* and *H*: it is *a priori* very improbable that what is seen is really a “house-face” and it is difficult to think of interactions with the environment that could have induced a prior for this hypothesis. As long as the low prior off-sets the likelihood advantage for *F AND H*, over *F* and *H*, it will not be selected over *F* or over *H*. A prediction follows from this: rivalry will be extinguished if the blended hypothesis happens to have a high prior (this may describe what happens in perceptual grouping under rivalry, see below).

Perceptual selection of a unitary hypothesis follows naturally from the predictive coding. This is because, *a priori*, the brain has learnt that there can be only one cause of sensory input at the same place and time. This generic prior constraint (a “hyperprior”) reflects the way we sample the visual world; binocular vision, in primates, rests upon both eyes foveating the same part of visual space. We have therefore learned that the explanation for binocular visual input is unitary (i.e., has just one cause). In other species, such as reptiles (whose eyes point in different directions) it is possible that a house and face could be perceived conjointly in different places. However, for us, this is *a priori* highly unlikely. In other words, the prior probability of both a house and face being co-localised in time and space is extremely small, to the extent it is almost impossible for us to support this representation or percept. The neuronal mechanisms mediating this selection are probably very similar to those mediating lateral inhibition

in the early visual system. These lateral interactions induce ‘winner-takes-all’ or ‘biased competition’ (Desimone, 1998) dynamics and may represent a fundamental mechanism in Bayesian inference.

In a recent model (Noest et al., 2007), it is argued that selection and alternation are the results of two fundamentally different mechanisms. Noest et al. accordingly model selection (what they call ‘percept choice’) using subthreshold facilitation, and alternation with adaptation, inhibition and noise. We agree in principle that these two questions are separate but the account we describe explains them in a unified manner as different facets of the same empirical Bayesian mechanism.

5. Solving the alternation problem

The theoretical challenge is to explain why the system, having selected one stimulus for perception, after a few seconds decides to de-select it in favour of the alternative stimulus. It is clear *a priori* that some kind of reciprocal inhibition must be involved but inhibition cannot be the whole story, if alternation is to be explained. There must be a dynamic evolution of inhibition and activity to ensure alternation. Traditionally, one appeals to adaptation, which allows disinhibition; this seems to be a fundament of any theory of rivalry; so we explain in epistemological terms why the visual system should exhibit adaptation.

The predictive coding framework posits a hierarchical inversion of generative models of how inputs are caused:

At the higher, hypothesis-generating level only the currently best hypothesis is allowed to generate predictions. It seems plausible that inhibition will be lateral, in relation to other hypotheses at the same level. This gives high activity for the winning hypothesis with the highest posterior and thus for the dominant percept, and lower activity for other hypotheses at that level.

At the lower level there is the opposite pattern: the bottom-up driving signal for the dominating percept is explained away by good predictions, meaning the prediction error for the dominant hypothesis is suppressed. Conversely, the bottom-up error signal for the currently suppressed stimulus is not.

In our example, there will be predictive activity creating the top-down signal for the *dominant* face stimulus and much driving activity in the bottom-up prediction error signal for the *suppressed* house stimulus. The key point is that even though *F* successfully explains the face-signal, there remains a large error signal, stemming from the house-stimulus (see Fig. 3).

This unexplained prediction error renders perceptual inference unstable. It is this instability that causes perceptual alternations. It is probably easiest to understand the mechanisms of perceptual transitions in terms of a bi-stable system: if the brain is trying to minimise prediction error or free-energy, we can associate a free-energy or potential with every brain state, for a fixed stimulus. In bi-stable systems the resulting energy landscape corresponds to a double well (see Fig. 4). The state of the brain will try to minimise the free-energy by occupying one of

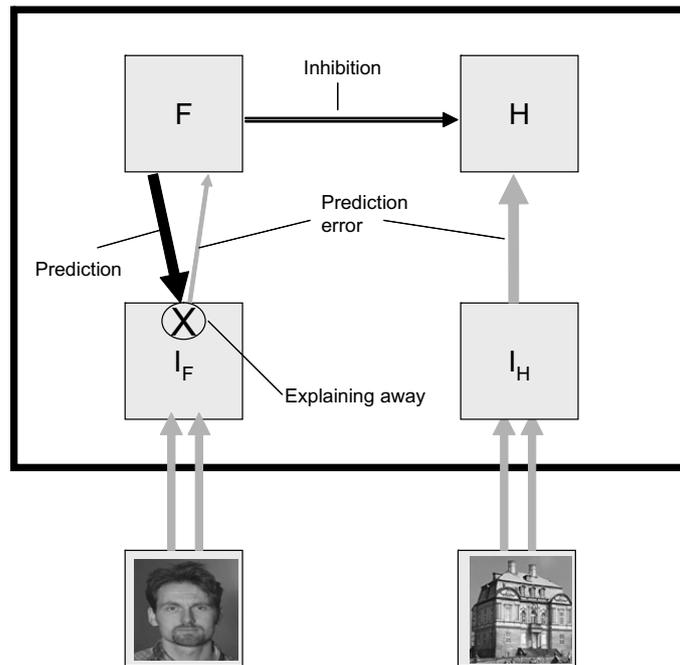


Fig. 3. Simplified schematic of rivalry using generative models and predictive coding for a system consisting of just one pair of levels: even though one hypothesis (F) about the environmental cause leaves only little prediction error from that stimulus (thin light arrow from I_F), a large unexplained signal is left unexplained from the other stimulus (thick light arrow from I_H).

the two (face or house) wells. Theoretically, there are two mechanisms that can cause the state of the brain to switch from one well to the other (i.e., cause perceptual alternations). These are dynamical and structural in nature and can be understood in terms of free-energy minimisation:

5.1. Structural instability and adaptation

This mechanism rests on changes in the free-energy landscape that make the occupied well unstable. Put sim-

ply, the well that is currently occupied increases its free-energy so that the brain's state is expelled to the other well (i.e., perceptual state). This involves a structural change to the landscape that makes the current state structurally unstable. The reason the current state becomes unstable could be that there is a strong (hyper-)prior that the world changes. A static hypothesis will quickly lose its clout in a changing world. There are many instances of this in terms of neuronal dynamics such as spike-rate adaptation and other adaptation phenomena observed neurophysiologically. In terms of predictive coding, the current hypothesis will always have a decreasing prior probability. In neuronal terms this would be mediated by adaptation of the corresponding neuronal representation. As this hypothesis shows adaptation, it fails to suppress prediction error and the free-energy of that state increases. This means that the occupied energy well becomes unstable and the prediction error associated with the competing hypotheses will eventually supervene and cause the percept to switch to the other energy minimum. See Fig. 4a for a schematic summary of this adaptation hypothesis.

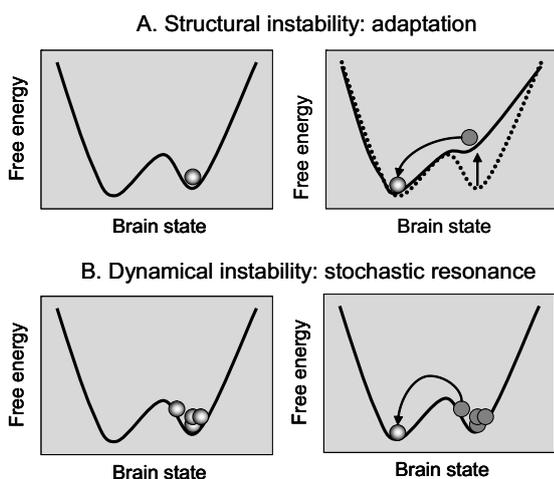


Fig. 4. Schematic summaries of: (A) Structural instability and adaptation. A hyperprior that makes the system expect change in the environment diminishes the energy well for the current perceptual inference. (B) Dynamical instability and stochastic resonance. Stochastic resonance refers to the same mechanism by which random fluctuations in a system's state enables it to move over energy barriers and explore multi-stable landscapes. See main text for further explanation.

5.2. Dynamical instability and stochastic resonance

Structural instability can be mediated using deterministic mechanisms. Another possible mechanism for perceptual alternations relies on stochastic or random effects (for recent discussions, see Brascamp, van Ee, Noest, Jacobs, & van den Berg, 2006; Kim, Grabowecky, & Suzuki, 2006; Moreno-Bote et al., 2007). Because the brain is trying to minimise its free-energy, it has to explore the free-energy landscape. A generic scheme for this exploration relies on random or stochastic effects (cf., random mutations in evolutionary selection or random noise in simulated

annealing). In multi-stable dynamical systems, this can be expressed as stochastic resonance. Put simply, random changes, due to neuronal noise, in the brain's state can occasionally push it over the free-energy barrier separating the house and face wells. This mechanism does not involve changes in, or adaptation of, the free-energy landscape but rests on dynamical instability introduced by random fluctuations in the brain's state. See Fig. 4b for a schematic illustration of this mechanism. There is good evidence that stochastic resonance plays such a role in rivalry. Kim et al. (2006) subjected rival stimuli to weak periodic contrast modulations and observed dominance peaks predicted by stochastic resonance. Generally, stochastic resonance occurs when the signal-to-noise ratio of a nonlinear system is maximized at a moderate level of noise. It occurs in bistable and excitable systems with sub-threshold inputs. Usually, the inputs constitute a weak periodic signal, which have a greater effect when noise enables the input to surpass threshold. However, in our case we are not deal-

ing with input–output characteristics but the dynamics of a system that is trying to optimise perception. Here, we use stochastic resonance to refer to the same mechanism by which random fluctuations in a system's state enables it to move over energy barriers and explore multi-stable landscapes.

In short, either structural or dynamic mechanisms of predictive coding, or a combination, can explain perceptual alternation. Alternation ensues in rivalry conditions specifically where there is a large unexplained but explainable error signal. In Bayesian terms, in this situation no one hypothesis has both high likelihood *and* high prior, and inference becomes unstable. See Fig. 5 and Dayan (1998) who provides modeling evidence “that alternation can be generated by competition between top–down cortical explanations for the inputs, rather than by direct competition between the inputs”.

For some relatively compatible pairs of stimuli, conjoint hypotheses may have a relatively high prior, which would

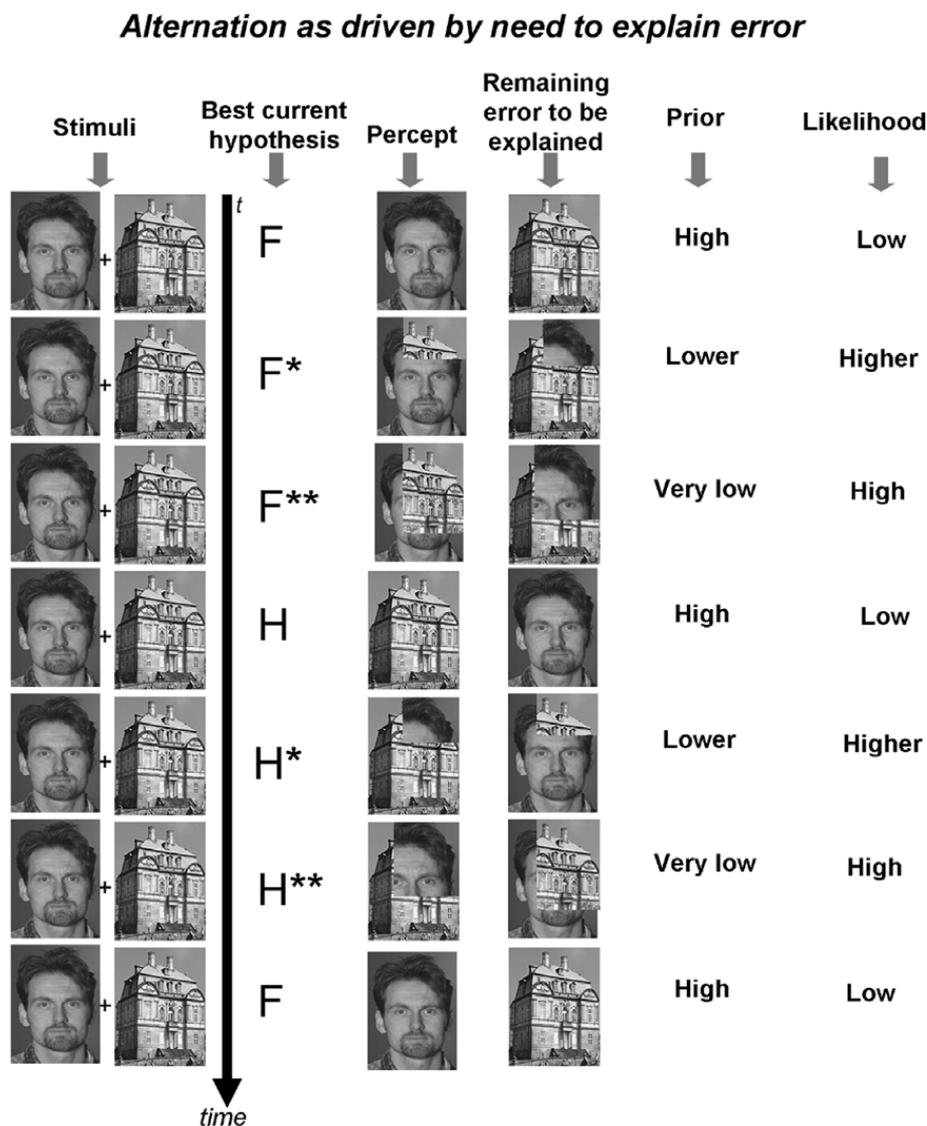


Fig. 5. Simplified Bayesian scheme for the alternation of stimuli in rivalry. When one stimulus achieves dominance and there are diminishing returns for predictions regarding it, the system must consider the best explanation of the unexplained error signal stemming from the currently suppressed stimulus (Starred hypotheses signify explorations of the free-energy landscape).

slow down alternation by creating a longer transitional phase (i.e., by adding a third well or attractor). Large differences in prior for F and H may make the system try to revert to F rather than shift to H (see Brascamp et al., 2006, for such transition returns); however, the system will not be stable as long as a predictable but as yet unexplained error signal from the house stimulus remains. These properties of the system correspond well to the often hesitant alternations for various combinations of stimuli (we say more about the psychophysics in the next section).

5.3. Summary

In this framework, the *inhibition* is not of the bottom-up, incoming signal *per se*. Rather it is inhibition of the competing high-level hypotheses that could explain away the sensory signal. In other words, inhibition decreases top-down predictions of the suppressed stimulus. The epistemological motivation for this is that the best performing hypothesis dominates perceptual content. The explanation for competition among high level explanations (*c.f.*, biased competition; Desimone, 1998) is simple; our experience of the world tells us that only one object can exist in the same place at the same time. This hyperprior is learnt and engrained in our neuronal circuits as an empirical prior.

The effect of this hyperprior is that bottom-up signals from the suppressed stimulus are not cancelled by top-down predictions; this increases the free-energy of the system and makes it more unstable. Mechanistically, this may be mediated by a reduction in the strength of lateral connections in the cortex that encode the uncertainty about, or precision of, visual signals (Friston, 2003). These changes may be enacted by modulatory neurotransmission (*c.f.*, Yu & Dayan, 2005) or possibly fast synchronised oscillations (*c.f.*, Womelsdorf & Fries, 2007).

The ensuing instability helps explain why random fluctuations may play a significant role in rivalry. In contrast, approaches to rivalry that do not employ predictive coding or free-energy minimisation will tend to view inhibition as decreasing the strength of the bottom-up signal; this stabilises the system and thus makes it harder to see why alternation should occur in the first place. Also, a predictive coding scheme fits particularly well with a system that exploits stochastic effects, since both the effect of the noise and the occurrence of attractors is explained in terms of the brain's free-energy landscape. There are many examples of this interplay in both the physical sciences (e.g., Yang, Onuchic, & Levine, 2006) and neurobiology (e.g., Winterer et al., 1999).

With empirical Bayes we can see why there is *adaptation*: if the system has learned that the world always changes – that there is variability in the environment – then even initially adequate hypotheses will have decreasing posteriors over time; as it will be more and more probable that there will be portions of sensory evidence that it fails to explain away. Notice that a prior for change is not something that the system will be able to extract from static visual scene statistics; instead, it comes down as a hyperprior in an empirical Bayes framework.

Once structural and dynamical instability have done their jobs, and the perceptual state has shifted from F to H , the mechanisms kick in again, and the system will then adapt to H , and eventually shift back to F . There is no psychophysical evidence that rivalry can be extinguished altogether except for very weak stimuli (Liu, Tyler, & Schor, 1992) (this seems in contrast to perception of ordinary bistable stimuli) so even though one stimulus may enjoy a high prior probability and be highly variable, and the other a low prior and not be variable, alternation will eventually occur. We explain this by appealing to the fact that incompatible hypotheses (like F and H) will each have low but roughly equal likelihoods, which will always leave an attractor for the non-dominant hypothesis. Given stochastic effects, the system will eventually come to occupy this state too, with probability one.

In sum, the proposal therefore motivates inhibition and adaptation in a more principled way than non-epistemological accounts, and thus explains rivalry as an unavoidable and emergent outcome of representational systems like the brain. It rests on the recurrent dynamics required by hierarchical inference and positions itself in direct opposition to conventional heuristics that frame perception in terms of feedforward dynamics; e.g., like the following from Lee, Blake, and Heeger (2005) “competition between two rival stimuli involves neural circuits in V1, and attention is crucial for the consequences of this neural competition to advance to higher visual areas and promote perceptual waves.”

6. Integrating psychophysical evidence under the predictive coding framework

6.1. Less rivalry for consistent stimuli

As noted by Blake (1989) rivalry tends to occur when there is an increasing incompatibility between the stimuli presented to the two eyes. More consistent stimuli will tend to fuse. This fits within the predictive coding framework because it is a case where the conjoint hypothesis does have high prior. That is, were the stimuli a mouth-less face and a mouth, then the updated, dominant hypothesis F^* (“it’s a face with a mouth”) would have a substantial prior. Fusion would then be allowed since the most likely hypothesis will have a high prior and the system will settle in a deep third well.

6.2. Patchy break-through of suppressed percept

Often, there is no clear-cut shift between percepts in binocular rivalry. Dominance breaks through in small patches of the visual field and gradually spreads before completely or partially suppressing the competing image (Lee et al., 2005; Meenes, 1930; Wheatstone, 1838). So there are periods where the subject experiences some of the face and some of the house. This is explained by the attempts to update the currently dominating hypothesis by exploring the free-energy landscape in response to the prediction error signal. The system does not stabilise with these patches because much prediction error still is unac-

counted for and because the conjoint hypothesis has a very low prior.

This part of the phenomenology may be influenced by perceptual inference for local stimulus attributes at low levels of the visual hierarchy. For example, for a particular area of visual space, where parts of the house and the face do not have much overlap, there may be a good perceptual inference to the occurrence of, say, the elemental features of a nose such that error for that area is efficiently explained away. Solutions for this area may therefore be a starting point for perceptual inference for the whole stimulus. However, even though the local posterior probability of the occurrence of a nose is high, it will decrease when considered in a more global context where ‘nose-houses’ have very low prior.

6.3. Inter-ocular grouping

Subjects may also experience rivalry where they perform visual grouping of items presented to both eyes. For example, if there is an image of half a face and half a house presented to one eye, and an image of the other halves of the face and of the house presented to the other eye, then there may be perceptual rivalry between a house and a face; the two halves of the two images have been grouped together, and it is the re-grouped percepts that are rivaling, not the original segmented images (Díaz-Caneja, 1928). This can also be done with patchy rivalry sets, and the effect is less stable than with non-patchy, conventional rivalry stimuli (Alais & Blake, 1999a; Kovacs, Papathomas, Yang, & Feher, 1996; Lee & Blake, 2004). This can be explained by a higher prior probability for the grouped stimuli than for the divided stimuli. That is, F and H will each have higher priors or stronger attractors than the hypothesis that it is a half face-half house. As F begins to dominate, the face signal from each eye will be explained away, leaving a coherent whole-house signal unaccounted for as the prediction error. This effect would be more top-down or prior driven than when no grouping occurs (since the likelihoods of the competing hypotheses will be similar) and indeed this effect requires some learning and is harder to sustain.

For interocular grouping of Díaz-Caneja stimuli (dichoptic viewing of two half-fields of concentric circles and vertical lines) perception alternates between rivalry between the half-fields and rivalry between the coherent stimuli (Ngo, Miller, Liu, & Pettigrew, 2000), and dominance times of the coherent but not the half-field percepts can be modulated with caloric vestibular stimulation (Ngo, Liu, Tilley, Pettigrew, & Miller, 2007). This suggests two pairs of attractors, at different hierarchical levels, that rival within each pair and among pairs (“meta-rivalry”, Ngo et al., 2007). This complex energy landscape could be explained by pairwise adaptation of attractors: once the attractors for half-fields have both been occupied the system begins to expect change in the environment away from half-fields. Since a pair of higher level attractors is available the state settles there until they in turn adapt.

Lee and Blake (2004) explored the role of patchiness in inter-ocular grouping and found that eye-specific processing may have a role to play in this type of inter-ocular

grouping. They propose that dominance patterns comprise local eye-based zones of dominance that are in turn subject to more global grouping forces. This is in fact consistent with the hierarchical nature of predictive coding that allows local solutions to bias priors in favour of one or the other global stimulus. The exact course of dominance and grouping will depend strongly on the choice of rivaling stimuli and the nature of the patchiness since rivalry for local stimulus attributes will depend on the balance of priors and likelihoods for the patch as well as the concurrent updating of probabilities for neighbouring patches.

6.4. Flicker and swap rivalry

When flickering stimuli are swapped rapidly between the eyes, normal dominance patterns of rivalry still occur such that one flickering percept may dominate for several seconds, during which period each eye is actually presented with each stimulus numerous times (Logothetis et al., 1996). This is surprising because we would expect, perhaps, that there would be rivalry between difficult-to-distinguish trains of flickering, shifting percepts. We again explain this in terms of relatively high priors or attractors for the distinct hypotheses F and H , relative to the conjoint hypothesis.

Under these conditions, there is, in fact, change in the world since each eye channel is presented with changing stimuli. This might suggest that the hyperprior for change is satisfied, and that adaptation therefore should not take place. However, the prior probability that two objects could move sufficiently fast between locations to reproduce flicker-stimuli is very low. Indeed such hyperpriors, that flickering stimuli are caused by the motion of a single stimulus, are the cornerstone of many psychophysical and electrophysiological studies of apparent motion (e.g., Billock & Tsou, 2007).

The unexpectedly normal pattern of rivalry, where the swapping is not perceived, can be explained by appeal to hierarchical processing, which entrains early visual cortex, before binocular convergence, and imposes higher level constraints on the percept. This has been modelled successfully in a hierarchical neural model (Wilson, 2003). In short, it may be that the hyperprior for a variable environment favours slow change over rapid change (cf. apparent motion). A good case for such a prior for slow change has also been made in relation to tactile perception in a model of the cutaneous rabbit illusion; this is a condition with much noise due to poor tactile acuity, which therefore allows priors to play a pivotal role for perceptual inference (Goldreich, 2007).

In a different type of paradigm Blake and colleagues (Blake, Westendorf, & Overton, 1980; Lee & Blake, 2004) allowed one stimulus to achieve dominance before they gradually decreased the intensity of the stimuli and swapped them. When the suppressed stimulus is swapped to the eye of the dominant stimulus, it becomes dominant, suggesting a role for eye-dominance rather than pattern competition in rivalry. This is also consistent with predictive coding. With respect to processing for the dominant eye stimulus it is a situation where there is successful prediction of (gradual or non-rapid) changes in the world. On

the assumption that there is less adaptation effect for such changing stimuli the system should remain relatively stable, and one should expect the dominant eye to continue its domination.

6.5. Percept selection repetition (rivalry memory)

If viewing of bistable stimuli is interrupted for long periods of time (5 s), then the selected percept post-interruption will tend to be a repeat of the last seen stimulus (Orbach, Ehrlich, & Heath, 1963). At shorter interruptions rivalry alternation is not disturbed in the same way. This has recently been modelled with subthreshold facilitation (Noest et al., 2007), or excitatory synaptic facilitation followed by depression (Wilson, 2007). This phenomenon can also be accommodated within the present framework. Interruptions are changes in the environment and thus, given a hyperprior for change in the environment, something that relieves the need for continued adaptation (i.e., decreased prior) for the dominant percept. Thus when the stimuli are shown again, short-term changes in synaptic efficacy established by the last percept confer an advantage in terms of perceptual inference. Such changes in synaptic efficacy (Noest et al., 2007) are entirely consistent with perceptual learning under empirical Bayes (Friston & Stephan, 2007) and may mediate sensory learning in the auditory domain, when stimuli are repeated (Garrido, Kilner, Kiebel, Stephan, & Friston, 2007). The reason the repetition effect is weaker after shorter interruptions may be that the hyperprior for change (possibly mediated by synaptic depression) is still active.

6.6. Monocular rivalry

Rivalry can also occur for a single stimulus presented to one or both eyes (Andrews & Purves, 1997; Breese 1899; Campbell, Gilinsky, Howell, Riggs, & Atkinson, 1973). The experience of monocular rivalry is less stable than in binocular rivalry and seems to occur mostly for fairly rudimentary stimuli such as a mesh of blurred green and red gratings. We think this reflects dynamically stable priors for the hypothesis that the environment has line segments of different distinct orientations. In other words, this is something the visual system is always expecting (Hubel & Wiesel, 1962; Kenet, Bibitchkov, Tsodyks, Grinvald, & Arieli, 2003). Therefore, when no higher level hypotheses are involved and when there is enough uncertainty or noise in the system (e.g., blurring), the system will try to predict the scene for line segments rather than meshes. With these kinds of meshes such predictions will be successful and rivalry will then occur. This would also help explain why there is rivalry for orthogonal gratings presented dichoptically even though the prior for the conjunctive hypothesis is not very low.

A recent study of monocular rivalry (Knapen et al., 2007) shows that increased depth perception of blurred orthogonal gratings (such that one is perceived to be behind the other) does not decrease rivalry even though it then is less likely that they are incompatibly occupying the same spatiotemporal location. This supports the view that suppression is “determined by a distance in a low-le-

vel neurally represented space subtended by features such as orientation” rather than by estimation of likelihoods based on the parameters of fully elaborated object representations in internal models. Given the hierarchical nature of our empirical Bayes framework, different stimulus attributes are each processed at distinct levels in the hierarchy where the priors of each model will be influenced from levels above. The question is then why the priors of this low level model do not give the predicted role to depth cues. The answer, as above, is that this may be tied to the specific stimulus used, which is very basic and noisy. It follows that higher-level bistable stimuli should decrease rivalry as incompatibility is lessened, e.g., if depth and context cues allow us to interpret the faces in Rubin’s vase as behind the vase, then rivalry should decrease, contrary to the findings for monocular rivalry gratings in Knapen et al.’s study.

6.7. Levelt’s Second Proposition (Levelt, 1965)

This is the key finding that contrast change in one eye (the “variable” eye) primarily causes changes in dominance durations in the other, “fixed” eye, rather than in the variable eye itself (the variable eye still has some dominance change (Bossink, Stalmeier, & De Weert, 1993; Mueller & Blake, 1989). Our account of Levelt’s Second Proposition is that, when the fixed eye stimulus is dominant, changes in the unexplained prediction error from the suppressed stimulus in the variable eye induces changes in the overall energy landscape, such that the perceptual decision for the fixed stimulus is brought away from or towards transitions over the free-energy barrier. For example, the probability that the fixed stimulus is the cause decreases as the variable stimulus is strengthened because the hypothesis for the fixed stimulus then predicts less of the total bottom-up signal. The fixed stimulus attractor is then evacuated earlier than if the variable stimulus had not been changed. This decreases its suppression periods, as the Second Proposition says (and *vice versa* when the variable stimulus is weakened). On the other hand, when the variable stimulus is itself dominant there is not this additional increased prediction error to destabilize the system. In that case, there is only the normal structural and stochastic dynamics in play.

The above account of Levelt’s Second Proposition is tied to conditions where the fixed stimulus contrast is relatively high. It has recently been found (Brascamp et al., 2006) that when the fixed stimulus contrast is very low, the Proposition is reversed such that the dominance of the variable stimulus is mainly modulated by changes in itself. Under the current account, this may be because the attractor for the fixed stimulus is then already quite shallow and thus already more susceptible to transitions associated with an unstable energy landscape.

Levelt’s Fourth Proposition is that, when the strength of both stimuli are increased, suppression periods for both will be shortened (Levelt, 1965). Again, the explanation is that when perception of both stimuli is associated with stronger, unpredicted error signals from the other perceptually suppressed stimulus, then the system will be impelled to explore the free-energy landscape earlier.

These considerations also help explain why changes in the suppressed stimulus will be noticed, and thrust the stimulus into dominance, if accompanied by abrupt increases in stimulus strength (Blake, Yu, Lokey, & Norman, 1998). On the other hand, when such probes have less abrupt onsets they tend to go unnoticed irrespective of whether they are congruent or not with the stimulus (Blake & Camisa, 1979). This coheres with the predictive coding account since probes that are congruent with the dominant stimulus are already predicted and probes that are not congruent are not predicted and just adds to the already large error signal.

6.8. Modulation of dominance duration

(i) When one stimulus is viewed in a congruent context and the other in a non-congruent context, the dominance duration of the former tend to increase (Alais & Blake, 1999). Introducing a congruent context does not increase bottom-up error signal strength when suppressed, so the predictive coding framework can explain why context modulation does not give shorter dominance periods for the non-congruent stimulus. On the other hand, context increases the prior for a congruent stimulus relative to a non-congruent stimulus, so it would take longer for the posterior for the dominant, updated hypothesis to be destabilised. This would explain the increased dominance periods. (ii) With practice, voluntary (endogenous) attention can prolong dominance periods for the attended stimulus without however being able to extinguish rivalry (on the other hand, endogenous attention to properties of the suppressed stimulus will not bring that stimulus out of suppression) (for discussion, see Blake & Logothetis, 2002; Leopold & Logothetis, 1999). This is an example of a top-down process modulating dominance. In the predictive coding framework we can view endogenous selective attention as increasing or enforcing priors for a certain hypothesis. That would make the system sensitive to what the hypothesis “wants” to see, which could prolong dominance. This admittedly schematic proposal is also consistent with findings that attention to a cue can determine onset predominance (Mitchell, Stoner, & Reynolds, 2004) and that removing attention from the stimuli slows down rivalry alternations (Paffen, Alais, & Verstraten, 2006) (for more on the relationship between attention and free-energy, see Friston & Stephan, 2007). It also makes sense that this cannot halt rivalry; if the system is using inappropriate priors they will not be sustained, because these priors are themselves subject to top-down influences. This is what happens in rivalry conditions where F and H are in effect bad hypotheses due to the large error signal they must leave unpredicted. On the assumption that the high level hypothesis for the suppressed stimulus is inhibited, it seems plausible that activity in it cannot be artificially maintained, which explains why endogenous attention to the suppressed stimulus will not bring it to dominance. (iii) Whereas endogenous attention has some effect on the dominant stimulus, exogenous attention (attention “grabbing”) in the suppressed stimulus will bring it out of suppression (Fox & Check, 1968). Attention grabbing thus decreases the dominance period of the other stimulus

and is explained by an increase in strength of the error signal (this seems consistent with findings on continuous flash suppression; Tsuchiya, Koch, Gilroy, & Blake, 2006).

7. Accounting for conflicting neurophysiological and imaging evidence

Empirical findings on rivalry using single unit recordings and fMRI seem to be in conflict and are difficult to unify under a single theoretical framework. However, it is important to remember that neuronal implementations of predictive coding require both the representation of the prediction and the prediction error in hierarchically ordered pairs of levels in the brain. It is the hierarchal deployment of reciprocal changes among these that will offer an explanation for diverse empirical findings.

Single unit studies in monkeys, yield the following consistent picture. Starting with the LGN, there seems to be no evidence of rivalry related changes in the geniculostriate system (Lehky & Maunsell, 1996). Successive stages of the visual cortex show increasing levels of activity in phase with the animal's reports of dominance: at low levels (V1) only few units selective for a given stimulus will fire in phase with dominance and suppression. At middle levels (V4, MT) more will, but the picture is somewhat mixed with some cells more active than other, almost no cells completely suppressed and some cells even active when their preferred stimulus were suppressed. This suggests that single unit recording can selectively sample either the predicting neurons or the prediction error neurons. At high visual levels in the temporal lobe there is good correspondence between rivalry alternation and physical stimulus alternation with most units firing only when their preferred stimulus were perceptually dominating and not firing when it was suppressed (Leopold & Logothetis, 1996; Logothetis & Schall, 1989; Logothetis & Sheinberg, 1996; Logothetis et al., 1996; Sheinberg & Logothetis, 1997). This would be expected because this is where high-level predictions are formed.

In general, fMRI studies in humans furnish a different picture. These studies have found that activity during rivalry corresponds to activity during physical alternations of stimuli over a large posterior portion of the brain ranging from temporal (fusiform and parahippocampal) areas (Tong, Nakayama, Vaughan, & Kanwisher, 1998), over V1 (Lee & Blake, 2002; Polonsky, Blake, Braun, & Heeger, 2000), including monocular areas such as the blind spot representation (Tong & Engel, 2001) and extending all the way to the lateral geniculate nucleus (Haynes, Deichmann, & Rees, 2005; Wunderlich, Schneider, & Kastner, 2005). Thus, in these areas of the brain, fMRI activity during dominance is comparable to activity during monocular viewing and activity during suppression is comparable to when the stimulus is not presented to the subject.

Perceptual rivalry presents a particular challenge to interpreting fMRI results in terms of predictive coding. This is because low-level areas that represent the elemental features of both stimuli will always express prediction error, because only one set of sensory signals can be explained away at any time. This means that there may be no difference in fMRI signals between the two perceptual

states in these regions. We would only expect fMRI differences at the first hierarchical levels, encoding one of the perceptual attributes showing rivalry. This signal might reflect the activity of deep pyramidal cells sending predictions to the lower levels or their post-synaptic effects in the subordinate level. The differential signals in the fusiform and parahippocampal areas are easy to understand because these areas show category-specific responses; but what about lower visual areas like V1? Interestingly Tong and Engel found rivalry effects for elemental features (gating orientation) that are encoded in V1. In their study the difference between orientations (which are all represented in V1) was observed in the monocular region corresponding to the blind spot. This region represents or predicts the input from only one eye and can therefore show perceptual differences that are not confounded by predictions or prediction error from the other eye. Although these authors framed their explanation in terms of lateral interactions within V1, their conclusions were based on the same constructivist arguments used by predictive coding. The study by Polonsky et al. (2000) elected difference by using stimuli with differing contrast, another elemental feature encoded by V1. The findings in the LGN are consistent with prolific top-down influences from V1 (backwards connections from visual cortex are an order of magnitude greater in number than forward afferents). These results suggest that fMRI signals reflect the postsynaptic effects of top-down afferents and the inherent predictions these projections convey to lower areas. Physiologically, this is sensible because hemodynamic signals are thought to be driven by pre-synaptic discharges causing depolarization in both target neurons and glial cells (c.f., Logothetis & Pfeuffer, 2004). This depolarization does not necessarily cause the neurons to fire.

Exactly the same dissociation between single-unit recordings and fMRI signals has been observed with top-down attentional effects, which are seen with fMRI but not in terms of single unit firing (Somers, Dale, Seiffert, & Tootell, 1999). In short, fMRI correlates of rivalry may be driven by top-down predictions, whereas electrophysiological responses may reflect predictions or prediction error, depending on which population or unit is recorded.

Irrespective of these considerations, the highest prediction error (free-energy and BOLD signal) would be anticipated during perceptual transitions, when neither stimulus is explained away. This is exactly what was found in one of the first studies of rivalry using fMRI (Lumer, Friston, & Rees, 1998). In summary, generative models and predictive coding therefore provide a framework that is capable of unifying the apparently conflicting findings on binocular rivalry.

8. Discussion

Under the account described here, an empirical Bayes framework with generative models and implemented with predictive coding or free-energy minimisation explains many aspects of binocular rivalry; because dichoptic viewing of mutually inconsistent stimuli creates a situation where no hypothesis about the environmental causes of

the incoming sensory signal has both a high prior *and* high likelihood. The system therefore settles into a rhythm, where at any time the hypothesis with the highest posterior probability determines perceptual content but at the cost of leaving a large unexplained but explainable error signal. In the attempt to account for this error signal, the posterior probability for the winning hypothesis is driven down below the free-energy for the alternative hypothesis that therefore begins to dominate.

8.1. Recent models of binocular rivalry

A number of formal models have been developed which are able to reproduce aspects of binocular rivalry (Dayan, 1998; Grossberg & Mingolla, 1985; Kalarickal, 2000; Kawamoto & Anderson, 1985; Koene, 2006; Laing & Chow, 2002; Lehky, 1988; Lumer, 1998; Matsuoka, 1984; Noest et al., 2007; Wilson, 2003; Wilson, 2007; Zhou, Gao, White, Merk, & Yao, 2004). The majority of these models analyse binocular rivalry *a priori* as a phenomenon driven by adaptation of the winning neural population, and lateral and/or top down inhibition of the competing population. Since successive dominance durations are stochastic system noise is often added.

Two very recent and impressive models exemplify this. Noest et al. (2007) construct a very simple low-level dynamic model that has terms for adaptation, lateral inhibition and noise. Wilson's (2007) model also has the virtue of simplicity: it too has terms for adaptation, inhibition, and can incorporate noise. These models differ mainly in mathematical complexity, and much of their behaviour can be described in terms of a double well energy landscape. Without appealing to high-level decision-making or memory, they can both reproduce phenomena such as rivalry, Levelt's Second and Fourth Propositions as well as percept choice repetition (explained differently in the two models; Wilson also incorporates a top-down role for attentional bias). This could appear to contrast with the present account that does appeal essentially to top-down processes. However, our account is based on hierarchical Bayes: each pair of levels throughout the cortical hierarchy forms a dynamic, functional unit that exercises perceptual inference. So perceptual inference is not driven directly by "decision-making" levels very high in the cortical hierarchy, though within each pair of levels rivalry is the outcome of dynamics ensuing from the upper level's attempt to explain the lower level's activity (for some evidence of decision-making processes in rivalry, see Einhauser, Stout, Koch, & Carter, 2008). A further, key difference is that, on our account, the dynamics for the structural and stochastic destabilization requires that there be no direct inhibition, between the lower level populations of any pair of levels, of the prediction error from the suppressed stimulus. The framework thus describes a different functional role for lateral inhibition on which explanation of phenomena such as Levelt's Second Proposition comes out as more surprising than it does for models that are built to directly inhibit competing populations. For these reasons Dayan's (1998) model, which explicitly uses top-down explaining-away, remains the model closest to ours. His model, however, does not motivate the adaptation (or 'fatigue'

function) in terms of hyperpriors and does not give the prediction error a core role in explaining the dynamics. Our framework is based on a principled story about overall brain function that has some biological and epistemic plausibility (Friston, 2002; Friston, 2003; Friston, 2005).

8.2. *The role of noise*

Successive dominance durations are unpredictable (Levelt, 1965) so most models of rivalry operate with added system noise. Recently, there has been an increased focus on the role of noise and stochastic resonance (e.g., Brascamp et al., 2006; Freeman, 2005; Kim et al., 2006; Moreno-Bote et al., 2007). Brascamp et al. (2006) show that the noise that is normally allowed in simple low-level models cannot account for the length of transition periods nor for transition returns. They congenially point to the possibility of a functional role of noise that “continuously reorganize sensory input to reach a perceptual solution. [N]oise may act to destabilize the present organization and prevent the brain from getting trapped in a single interpretation while others may have more survival value.” (1250).

If the minimization of free-energy is the overall processing aim of the brain, then we probably should not expect that the system produces blanket noise to disperse perceptual states randomly. However, hyperpriors, such as the expectation of change in the environment, may decrease prior probabilities of specific hypotheses even though they are successfully predicting the input at the moment. This destabilizes the energy landscape in a directed manner. The effect is, as described by Brascamp et al., that the system begins to explore the free-energy landscape. In addition, our account suggests a pivotal role for increased free-energy that is specific for rivalry conditions, namely the destabilizing effect of the unexplained prediction error from the suppressed stimulus. This is not merely added system noise but “noise” that occurs as an intrinsic aspect of the basic Bayesian framework.

8.3. *Other kinds of bistable perception: the role of attention*

Our account of binocular rivalry is principled and based on a general view of brain function, and we believe that it may also apply to other kinds of bistable perception, such as Rubin's vase and the Necker cube, but with less involvement at lower cortical areas (a shared type of mechanism is also suggested by the similarity of time courses for different kinds of bistable perception, van Ee, 2005). The difference in levels may be related to an epistemic difference: dichoptic viewing tends to present two different objects with different properties in the same spatiotemporal position (e.g., a red house vs. a green face). This *never* happens at any level in the causal hierarchy mapped by the cortex, so the hyperprior against it is global and as such encoded at low cortical levels. In contrast, the incompatibility for the dioptically presented bistable stimuli is less severe: e.g., one object that can be interpreted as having different properties (i.e., the Necker cube where transparency creates a situation with matched likelihoods). It is rare but not impossible to be confronted with such scenarios of

equally poised likelihoods and priors, so the hyperpriors that influence perceptual inference in these unusual circumstances will probably be less strong and encoded at higher hierarchical levels (since it takes analysis of deeper causal levels to grasp phenomena such as transparency). This leads us to expect that non-binocular rivalry bypasses lower levels.

A difference between binocular and non-binocular rivalry is that dominance durations in the former are harder to modulate with selective attention, while on the other hand non-selective attention can increase the alternation rate similarly in both types of rivalry (Meng & Tong, 2004). It is fundamental for the epistemic framework we are describing that there are distinct causes in the environment, so the hyperprior that prohibits spatiotemporal co-occupancy is global: without it the system would always have to consider whether an entirely distinct cause was at the same place and time, which would not be conducive to adaptive behaviour. Selective attention to one rivaling stimulus at the expense of the other is an attempt to get the brain to ignore prediction error at this very fundamental level, and it is thus not surprising that selective attention cannot modulate binocular rivalry so strongly. This does not hold for other kinds of bistable percepts where the stimulus incompatibility is less strong. Non-selective attention, in contrast, can be seen as relaxing the much less fundamental hyperprior that change in the environment is mostly relatively slow, thus speeding up alternation rate. Priors embody our expectations and whereas we can be in a context where unusually fast changes are common we cannot be in context where numerous distinct objects share spatiotemporal location.

8.4. *Predictions*

In this review, we have been most concerned with providing an epistemological framework for the conflicting data on rivalry rather than generating new data. However, a number of predictions can be made. (i) Semantic or subliminal priming will increase priors and thereby facilitate inter-ocular grouping and bias “meta-rivalry”. (ii) Conjoint hypotheses with a high likelihood will tend to facilitate fusion rather than rivalry (e.g., two transparent images on a shared background). (iii) A moving stimulus (e.g., gratings with a chaotic time pattern) will dominate over more predictable moving stimuli since new predictions will continually be needed and there will be less adaptation. (iv) LGN and blind spot representation activity measured with fMRI will not suggest that rivalry is resolved before binocular convergence, if deprived of backwards signals from areas above binocular convergence.

9. *Conclusions*

Core properties of a theoretical framework for perceptual inference in the brain based on generative models and predictive coding can be described in fairly basic probabilistic terms. The framework can explain and unify many aspects of binocular rivalry, in particular why one stimulus

is selected for perception and why there is alternation between stimuli. The framework also accommodates many of the major psychophysical findings on rivalry and provides a unified interpretation of the apparently conflicting single-unit and fMRI studies of rivalry.

The predictive coding explanation of binocular rivalry allows a principled, theoretically motivated combination of top-down and bottom-up processes. This further integrates the debate on how the primary findings on rivalry should be interpreted (Blake, 1989; Leopold & Logothetis, 1999; Tong et al., 2006). It does this by describing one unifying mechanism – prediction error minimisation – rather than a variety of different mechanisms (Blake & Logothetis, 2002). The framework relates to recent computational models (Noest et al., 2007; Wilson, 2007) in as far as it builds on inhibition, adaptation and noise but it gives these notions a distinct functional role and hierarchical dynamics (Dayan, 1998).

Acknowledgements

This research was supported by the Danish Research Council for Communication and Culture, The Danish National Research Foundation, The Wellcome Trust, and a Monash Arts/IT Grant.

References

- Alais, D., & Blake, R. (1999). Grouping visual features during binocular rivalry. *Vision Research*, 39, 4341–4353.
- Alais, D., & Blake, R. (Eds.). (2005). *Binocular rivalry*. Cambridge, Mass: MIT Press.
- Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, 14, 257.
- Alais, D., & Melcher, D. (2007). Strength and coherence of binocular rivalry depends on shared stimulus complexity. *Vision Research*, 47(2), 269–279.
- Andrews, T. J., & Purves, D. (1997). Similarities in normal and binocularly rivalrous viewing. *Proceedings of the National Academy of Sciences of the United States of America*, 94, 9905–9908.
- Ballard, D. H., Hinton, G. E., & Sejnowski, T. J. (1983). Parallel visual computation. *Nature*, 306(5938), 21.
- Billock, V. A., & Tsou, B. H. (2007). Neural interactions between flicker-induced self-organized visual hallucinations and physical stimuli. *Proceedings of the National Academy of Sciences of the United States of America*, 104(20), 8490.
- Blake, R. (1989). A neural theory of binocular rivalry. *Psychological Reviews*, 96, 145–167.
- Blake, R., & Boothroyd, K. (1985). The precedence of binocular fusion over binocular rivalry. *Perception & Psychophysics*, 37, 114.
- Blake, R., & Camisa, J. (1979). The inhibitory nature of binocular rivalry suppression. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 315.
- Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience*, 3, 13–21.
- Blake, R., Westendorf, D., & Overton, R. (1980). What is suppressed during binocular rivalry? *Perception*, 9, 223.
- Blake, R., Yu, K., Lokey, M., & Norman, H. (1998). Binocular rivalry and motion perception. *Journal of Cognitive Neuroscience*, 10(1), 46–60.
- Bossink, C. J., Stalmeier, P. F., & De Weert, C. M. (1993). A test of Levelt's second proposition for binocular rivalry. *Vision Research*, 33(10), 1413–1419.
- Brascamp, J. W., van Ee, R., Noest, A. J., Jacobs, R. H., & van den Berg, A. V. (2006). The time course of binocular rivalry reveals a fundamental role of noise. *Journal of Vision*, 6(11), 1244–1256.
- Breese, B. B. (1899). On inhibition. *Psychological Monographs*, 3, 1–65.
- Campbell, F. W., Gilinsky, A. S., Howell, E. R., Riggs, L. A., & Atkinson, J. (1973). The dependence of monocular rivalry on orientation. *Perception*, 2, 123.
- Dayan, P. (1998). A hierarchical model of binocular rivalry. *Neural Computation*, 10(5), 1119–1135.
- Dayan, P., & Abbott, L. F. (2001). *Theoretical neuroscience*. Cambridge, Mass: MIT Press.
- Dayan, P., Hinton, G. E., Neal, R. M., & Zemel, R. S. (1995). The Helmholtz machine. *Neural Computation*, 7(5), 889–904.
- Desimone, R. (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical Transactions of the Royal Society of London B*, 353, 1245.
- Diaz-Caneja, E. (1928). Sur l'alternance binoculaire. *Annales de Oculistique (Paris)*, 721.
- Einhauser, W., Stout, J., Koch, C., & Carter, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proceedings of the National Academy of Sciences of the United States of America*, 105(5), 1704–1709.
- Ernst, M., & Banks, M. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415, 429ff.
- Fox, R., & Check, R. (1968). Detection of motion during binocular rivalry suppression. *Journal of Experimental Psychology*, 78, 388.
- Freeman, A. W. (2005). Multistage model for binocular rivalry. *Journal of Neurophysiology*, 94(6), 4412–4420.
- Friston, K. (2002). Functional integration and inference in the brain. *Progress in Neurobiology*, 68, 113–143.
- Friston, K. (2003). Learning and inference in the brain. *Neural Networks*, 16(9), 1325–1352.
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions: Biological Sciences*, 369(1456), 815–836.
- Friston, K., & Stephan, K. (2007). Free energy and the brain. *Synthese*, 159(3), 417–458.
- Frith, C., Perry, R., & Lumer, E. (1999). The neural correlates of conscious experience: An experimental framework. *Trends in Cognitive Sciences*, 3(3), 105.
- Garrido, M. I., Kilner, J. M., Kiebel, S. J., Stephan, K., & Friston, K. (2007). Dynamic causal modelling of evoked potentials: A reproducibility study. *Neuroimage*, 36(3), 571–580.
- Goldreich, D. (2007). A Bayesian perceptual model replicates the cutaneous rabbit and other tactile spatiotemporal illusions. *PLoS One*, 2(3), e333.
- Gregory, R. L. (1997). Knowledge in perception and illusion. *Philosophical Transactions of the Royal Society B*, 352(1358), 1121–1127.
- Gregory, R. L. (1998). *Eye and brain* (5th ed.). Oxford: Oxford University Press.
- Grossberg, S., & Mingolla, E. (1985). Neural dynamics of form perception: boundary completion, illusory figures, and neon color spreading. *Psychological Reviews*, 92, 173.
- Haynes, J. D., Deichmann, R., & Rees, G. (2005). Eye-specific effects of binocular rivalry in the human lateral geniculate nucleus. *Nature*, 438, 496.
- Helmholtz, H. v. (1860). *Treatise on physiological optics* (J. P. C. Southall, Trans. 1962 ed., Vol. 3). New York: Dover.
- Hubel, D., & Wiesel, T. (1962). Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *Journal of Physiology (London)*, 160, 106–154.
- Hume, D. (1739–40). *A treatise of human nature*. Oxford: Oxford Clarendon Press.
- Kalarickal, G. J. (2000). Neural model of temporal and stochastic properties of binocular rivalry. *Neurocomputing*, 32(33), 843.
- Kawamoto, A. A. H., & Anderson, J. J. A. (1985). A neural network model of multistable perception. *Acta Psychologica*, 59(1), 35–65.
- Kawato, M., Hayakawa, H., & Inui, T. (1993). A forward-inverse optics model of reciprocal connections between visual cortical areas. *Network: Computation in Neural Systems*, 4, 415–422.
- Kenet, T., Bibitchkov, D., Tsodyks, M., Grinvald, A., & Arieli, A. (2003). Spontaneously emerging cortical representations of visual attributes. *Nature*, 425, 954–956.
- Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology*, 55(1), 271–304.
- Kim, Y.-J., Grabowecy, M., & Suzuki, S. (2006). Stochastic resonance in binocular rivalry. *Vision Research*, 46(3), 392.
- Knapen, T., Kanai, R., Brascamp, J., van Boxtel, J., & van Ee, R. (2007). Distance in feature space determines exclusivity in visual rivalry. *Vision Research*, 47(26), 3269–3275.
- Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27(12), 712–719.
- Koch, C. (2004). *The quest for consciousness: A neurobiological approach*. Englewood, Colorado: Robert and Company Publishers.
- Koene, A. (2006). A model for perceptual averaging and stochastic bistable behavior and the role of voluntary control. *Neural Computing*, 18(12), 3069–3096.

- Kovacs, I., Papathomas, T. V., Yang, M., & Feher, A. (1996). When the brain changes its mind: Igrouping during binocular rivalry. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 15508.
- Laing, C. R., & Chow, C. C. (2002). A spiking neuron model for binocular rivalry. *Journal of Computational Neuroscience*, 12(1), 39.
- Lee, S., & Blake, R. (2004). A fresh look at interocular grouping during binocular rivalry. *Vision Research*, 44, 983.
- Lee, S. H., & Blake, R. (2002). V1 activity is reduced during binocular rivalry. *Journal of Vision*, 2, 618–626.
- Lee, S. H., Blake, R., & Heeger, D. J. (2005). Traveling waves of activity in primary visual cortex during binocular rivalry. *Nature Neuroscience*, 8, 22.
- Lehky, S. R. (1988). An astable multivibrator model of binocular rivalry. *Perception*, 17, 215–228.
- Lehky, S. R., & Maunsell, J. H. (1996). No binocular rivalry in the LGN of alert macaque monkeys. *Vision Research*, 36, 1225.
- Leopold, D., & Logothetis, N. (1999). Multistable phenomena: Changing views in perception. *Trends in Cognitive Sciences*, 3, 254–264.
- Leopold, D. A., & Logothetis, N. K. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature*, 379, 549–553.
- Levelt, W. (1965). *On binocular rivalry*. Assen, Netherlands: Royal Van Gorcum.
- Liu, L., Tyler, C. W., & Schor, C. M. (1992). Failure of rivalry at low contrast: Evidence of a suprathreshold binocular summation process. *Vision Research*, 32(8), 1471–1479.
- Logothetis, N. K., Leopold, D. A., & Sheinberg, D. L. (1996). What is rivalling during binocular rivalry? *Nature*, 380, 621–624.
- Logothetis, N. K., & Schall, J. D. (1989). Neuronal correlates of subjective visual perception. *Science*, 245, 761–763.
- Logothetis, N. K., & Sheinberg, D. L. (1996). Visual object recognition. *Annual Reviews of Neuroscience*, 19, 577–621.
- Logothetis, N. K., & Pfeuffer, J. (2004). On the nature of the BOLD fMRI contrast mechanism. *Magnetic Resonance Imaging*, 22, 1517–1531.
- Lumer, E. D. (1998). A neural model of binocular integration and rivalry based on the coordination of action-potential timing in primary visual cortex. *Cerebral Cortex*, 8, 553.
- Lumer, E. D., Friston, K., & Rees, G. (1998). Neural correlates of perceptual rivalry in the human brain. *Science*, 280, 1930.
- MacKay, D. M. (1956). The epistemological problem for automata. In C. E. Shannon & J. McCarthy (Eds.), *Automata studies* (pp. 235–251). Princeton: Princeton University Press.
- Matsuoka, K. (1984). The dynamic model of binocular rivalry. *Biological Cybernetics*, 49(3), 201–208.
- Meenes, M. (1930). A phenomenological description of retinal rivalry. *American Journal of Psychology*, 42, 260–269.
- Meng, M., & Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision*, 4(7), 539–551.
- Mitchell, J. F., Stoner, G. R., & Reynolds, J. H. (2004). Object-based attention determines dominance in binocular rivalry. *Nature*, 429(6990), 410–413.
- Moreno-Bote, R., Rinzal, J., & Rubin, N. (2007). Noise-induced alternations in an attractor network model of perceptual bistability. *Journal of Neurophysiology*, 98(3), 1125–1139.
- Mueller, T. J., & Blake, R. (1989). A fresh look at the temporal dynamics of binocular rivalry. *Biological Cybernetics*, 61, 223.
- Mumford, D. (1992). On the computational architecture of the neocortex. *Biological Cybernetics*, 66(3), 241.
- Murray, S. O., Kersten, D., Olshausen, B. A., Schrater, P., & Woods, D. L. (2002). Shape perception reduces activity in human primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 99(23), 15164–15169.
- Murray, S. O., Schrater, P., & Kersten, D. (2004). Perceptual grouping and the interactions between visual cortical areas. *Neural Networks*, 17(5–6), 695–705.
- Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.
- Ngo, T. T., Liu, G. B., Tilley, A. J., Pettigrew, J. D., & Miller, S. M. (2007). Caloric vestibular stimulation reveals discrete neural mechanisms for coherence rivalry and eye rivalry: A meta-rivalry model. *Vision Research*, 47(21), 2685–2699.
- Ngo, T. T., Miller, S. M., Liu, G. B., & Pettigrew, J. D. (2000). Binocular rivalry and perceptual coherence. *Current Biology*, 10(4), 134–136.
- Nguyen, V., Freeman, A., & Alais, D. (2003). Increasing depth of binocular rivalry suppression along two visual pathways. *Vision Research*, 43(19), 2003–2008.
- Noest, A. A. J., van Ee, R. R., Nijs, M. M. M., & van Wezel, R. R. J. A. (2007). Percept-choice sequences driven by interrupted ambiguous stimuli: A low-level neural model. *Journal of Vision*, 7(8), 10.
- Orbach, J., Ehrlich, D., & Heath, H. (1963). Reversibility of the Necker cube. I. An examination of the concept of satiation of orientation. *Perception and Motor Skills*, 17, 439–458.
- Paffen, C. L. E., Alais, D., & Verstraten, F. A. J. (2006). Attention speeds binocular rivalry. *Psychological Science*, 17(9), 752–756.
- Polonsky, A., Blake, R., Braun, J., & Heeger, D. (2000). Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. *Nature Neuroscience*, 3, 1153.
- Rao, R., & Ballard, D. (1999). Predictive coding in the visual cortex. *Nature Neuroscience*, 2, 79.
- Rock, I. (1983). *The logic of perception*. Cambridge, Mass: MIT Press.
- Sheinberg, D. L., & Logothetis, N. K. (1997). The role of temporal cortical areas in perceptual organization. *Proceedings of the National Academy of Sciences of the United States of America*, 94(7), 3408–3413.
- Somers, D. C., Dale, A. M., Seiffert, A. E., & Tootell, R. B. (1999). Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 1663–1668.
- Summerfield, C., Egnér, T., Mangels, J., & Hirsch, J. (2005). Mistaking a house for a face: neural correlates of misperception in healthy humans. *Cerebral Cortex*, 16, 500–508.
- Tong, F. (2003). Primary visual cortex and visual awareness. *Nature Reviews Neuroscience*, 4(3), 219–229.
- Tong, F., & Engel, S. A. (2001). Interocular rivalry revealed in the human cortical blind-spot representation. *Nature*, 411, 195–199.
- Tong, F., Meng, M., & Blake, R. (2006). Neural bases of binocular rivalry. *Trends in Cognitive Sciences*, 10(11), 502.
- Tong, F., Nakayama, K., Vaughan, J. T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron*, 21, 753–759.
- Tsuchiya, N., Koch, C., Gilroy, L. A., & Blake, R. (2006). Depth of interocular suppression associated with continuous flash suppression, flash suppression, and binocular rivalry. *Journal of Vision*, 6, 1068–1078.
- van Ee, R. (2003). Bayesian modeling of cue interaction: bistability in stereoscopic slant perception. *Journal of the Optical Society of America B*, 20(7), 1398.
- van Ee, R. (2005). Dynamics of perceptual bi-stability for stereoscopic slant rivalry and a comparison with grating, house-face, and Necker cube rivalry. *Vision Research*, 45(1), 29–40.
- Wheatstone, C. (1838). Contributions to the physiology of vision. Part I. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society (London) B*, 128, 371–394.
- Wilson, H. R. (2003). Computational evidence for a rivalry hierarchy in vision. *Proceedings of the National Academy of Sciences of the United States of America*, 100(24), 14499–14503.
- Wilson, H. R. (2007). Minimal physiological conditions for binocular rivalry and rivalry memory. *Vision Research*, 47(21), 2741–2750.
- Winterer, G., Ziller, M., Dorn, H., Frick, K., Mulert, C., Dahhan, N., et al (1999). Cortical activation, signal-to-noise ratio and stochastic resonance during information processing in man. *Clinical Neurophysiology*, 110(7), 1193–1203.
- Womelsdorf, T., & Fries, P. (2007). The role of neuronal synchronization in selective attention. *Current Opinion in Neurobiology*, 17(2), 154–160.
- Wunderlich, K., Schneider, K., & Kastner, S. (2005). Neural correlates of binocular rivalry in the human lateral geniculate nucleus. *Nature Neuroscience*, 8(11), 1595.
- Yang, S., Onuchic, J., & Levine, H. (2006). Effective stochastic dynamics on a protein folding energy landscape. *Journal of Chemical Physics*, 125, 054910.
- Yu, A. J., & Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron*, 46(4), 681.
- Yuille, A., & Kersten, D. (2006). Vision as Bayesian inference: Analysis by synthesis? *Trends in Cognitive Sciences*, 10(7), 301.
- Zeki, S. (2003). The disunity of consciousness. *Trends in Cognitive Sciences*, 7(5), 214–218.
- Zhou, Y. H., Gao, J. B., White, K. D., Merk, I., & Yao, K. (2004). Perceptual dominance time distributions in multistable visual perception. *Biological Cybernetics*, 90(4), 256–263.