Review Paper

Title: Perception, illusions and Bayesian inference

Running head: Perception, illusions and Bayesian inference

Authors: Matthew M Nour[1,2], Joseph M Nour[3]

Author Affiliation: [1]South London and Maudsley NHS Foundation Trust, London; [2]Institute of Psychiatry, Psychology and Neuroscience, Kings College London, London; [3]Oxford University Clinical Academic Graduate School, John Radcliffe Hospital, University of Oxford, Oxford.

Corresponding Author:

Dr Matthew M Nour

CT1, Postgraduate Centre, Maudsley Hospital, 111 Denmark Hill, London, SE5 8AZ

T: +44 20 3228 3840, F: +44 20 3228 5201

email: matthew.nour@kcl.ac.uk

Abstract word count: 103 **Text body word count:** 2166

This article contains 34 references, 0 tables and 1 figure.

Title is 47 characters long. Running head is 47 characters long.

Key Words: Illusion, perception, Bayes Theory, Predictive coding, perceptual inference.

Abstract

Descriptive psychopathology makes a distinction between veridical perception and illusory perception. In both cases a perception is tied to a sensory stimulus, but in illusions the perception is of a false object. This article re-examines this distinction in light of new work in theoretical and computational neurobiology, which views all perception as a form of Bayesian statistical inference that combines sensory signals with prior expectations. Bayesian perceptual inference can solve the 'inverse optics' problem of veridical perception, and provides a biologically plausible account of a number of illusory phenomena, suggesting that veridical and illusory perceptions are generated by precisely the same inferential mechanisms.

Introduction

In this article we will revisit the classical distinction between illusory perception and veridical perception, in light of recent work in theoretical neurobiology. We outline theories in contemporary perceptual neuroscience that view perception as an inferential process. These theories have primarily been applied to low-level illusory phenomena, but should also be of interest to clinicians and neuroscientists interested in psychopathological phenomena, as their application is readily extended to illusory phenomena that occur in the context of abnormal mental states. From this vantage point the classical qualitative distinction between true perceptions and illusions appears indefensible.

Perception as Inference

It is now widely acknowledged in the neurosciences that perception is an inferential process[1–3]. In the 18th century George Berkeley recognized that information collected by the sensory epithelia is mathematically insufficient to allow an unambiguous mapping back onto real-world sources[4,5]. The light hitting the retina, for example, forms a two-dimensional image, which has an infinite number of possible three-dimensional real-world sources. The image conflates information about object illumination, reflectance and transmittance[4,6]. The problem of inferring the 'state of the word' from sensory data alone is thus ill-posed[7] (i.e. mathematically degenerate).

Ill-posed problems can be made tractable by using contextual information to impose constraints on the interpretation of ambiguous data. Thus, although a certain sensory stimulus may support multiple conflicting interpretations if taken in isolation, one interpretation usually stands out as being 'optimal' once the context of the situation is taken into account. The computational details of precisely how this contextual information is acquired and used to guide perceptual inference remains the focus of lively debate in the field[4,8–11].

Perceptual Inference and Illusions

In recent years perceptual neuroscience has focussed on sensory illusions in an attempt to elucidate the mechanisms underlying perceptual inference. Illusions represent situations where our perception reliably and indomitably disagrees with the true nature of the raw sensory information[12]. Even with the knowledge that, say, the two lines are of equal length in the Müller-Lyer illusion (Fig 1A), or that flanking luminance values are equal in the Craik-O'Brien-Cornsweet effect (Fig1B), it is exceedingly difficult for most people to perceive the images in this 'veridical' manner.

Progress in the field of perceptual and computational neuroscience has shown that 'illusory' perceptual phenomena, far from representing failures in the perceptual apparatus, are explained by a theory of perceptual inference that uses contextual information and prior assumptions to constrain interpretations of sensory data[12,13,6]. Some of this work is briefly outlined here, with references to more comprehensive theoretical reviews.

Wholly Empirical Vision

In the 'wholly empirical' strategy, or Empirical Rank Theory (ERT), the subjective qualities of our perceptions are based on how the value of a given stimulus ranks against the cumulative distribution of previously experienced stimuli in the same context[6,14]. For example, it is well known that vertical lines of a given length appear longer than horizontal lines of the same length[15] (Fig 1C). ERT argues that this illusion arises because in the summed evidence from past experience of natural scenes horizontal lines are generally longer than vertical lines, so for two lines of a given length the horizontal line ranks lower with respect to all other perceived horizontal lines (and is thus perceived shorter) than the vertical line. Dale Purves and colleagues have used ERT to explain a number of visual illusions using both psychophysical paradigms[16,17] and computational modelling based on analysis of statistical regularities in natural scenes[18–20]. ERT has also been used to reproduce electrophysiological properties of early visual neurons in artificial neurons[21]. These studies have been nicely reviewed elsehwere[6,11,14]. Although it is natural for us to think of perception as making inferences about the 'real world', it is important to note that ERT makes no appeal to a world of 'hidden causes' that lies behind sensory data[4], and is hence said to be 'wholly empirical'[14].

Bayesian Perception

Alternatively, it has been argued that the brain *does* maintain and update a representation of the world, which is used to guide perceptual inference about the 'hidden causes' underlying sensory information[1,2,9]. This reasoning begins by acknowledging that perception is best explained by an inferential process that takes into account both current sensory data and prior knowledge (or expectation) about the state of the world, in a manner which follows the laws of Bayesian inference[3,22]. In other words, where the current sensory data and prior expectation (both represented as probability distributions over 'real world causes') are at odds, the Bayes-optimal perceptual inference is the precision-weighted combination of the competing evidence (where precision is the inverse-variance of the said distributions, and is roughly equivalent to the confidence the observer has in the information[9]). Bayesian perceptual inference thus uses prior expectations to constrain the interpretation of incoming sensory data.

Intriguingly, non-invasive functional neuroimaging experiments in humans support the notion that feature-specific perceptual expectations (set up by prior experience) can modify cortical responses to incoming sensory signals at multiple nodes of the sensory processing hierarchy[7,23–25]. Despite its appeal, exact Bayesian inference is computationally expensive[1], and a challenge has been to show how the brain may implement this strategy in a biologically plausible manner[11].

Hierarchical Predictive Coding

One popular and neurobiologically plausible implementation of Bayesian perceptual inference, outlined by Karl Friston and colleagues, is hierarchical predictive coding[1,9]. Here perception, based on empirical Bayesian inference, occurs in a distributed fashion in reciprocally-linked hierarchical sensory processing circuits.

At the heart of hierarchical predictive coding is the notion that brain maintains dynamic representations of the world, housed in the synaptic connections of hierarchical sensory processing circuits[1]. These representations become progressively more abstract at higher levels of the processing hierarchy, representing the 'real world causes' giving rise to sensory signals at the highest hierarchical levels[2]. Based on this internal world-model the brain is able to make on-line predictions about the state that the world is in, which are transformed into predicted incoming sensory signals by learned internal 'generative' models. Predicted sensory signals cascade down the neuronal hierarchy, acting as ‘prior’ sensory probabilities (or expectations) for lower regions, thus constraining the interpretation of incoming sensory signals.

If there is a good match between the ‘top down’ prior probability and the ‘bottom up’ sensory data, the current representation of the state of the world is reinforced. If there is a mismatch a ‘prediction error’ signal drives an updating of the brain’s current world model, which is subsequently re-tested against the real world data. The iterative process of matching ‘top-down’ predictions to ‘bottom-up’ sensory signals arrives at a multimodal internal representation of the state of the world that is internally coherent and contains representations not only of basic sensory properties but also of abstract representations about the state of the world[2]. This final representation (formally, a ‘posterior probability’), arrived at through an inversion of the generative model[9], is equivalent to a Bayes-optimal (perceptual) inference about the ‘real world’ source of the sensory data. Eloquent formal mathematical treatments of this proposal are outlined elsewhere[1,7,9].

Hierarchical predictive coding models of perception have been successfully used to explain visual[13] and somatosensory[26] illusions. Brown and Friston have recently applied the model to the Cornsweet illusion[13] (Fig 1B), in contrast to the wholly-empirical treatment of the illusion[16]. Object luminance, registered by the retina, is created by the interaction between object illumination and reflectance. In the Cornsweet illusion the luminance of the peripheral regions on the left and the right are equal, but the perceived brightness is not. Brown and Friston created a simulated Bayesian observer using a hierarchical implementation of empirical Bayes equipped with simple prior beliefs that relate to the ways illuminance and reflectance patterns dynamically change both spatially and temporally. Their simulated observer not only 'perceived' the illusory left-right brightness difference of the illusion, but also perceived Mach bands (illusory paracentral vertical bands). The simulated observer's perception of the Cornsweet and Mach band illusions was sensitive to changes in contrast precision in a manner that was qualitatively similar to human subjects.

Illusions and Psychopathology

So far we have argued that low-level perceptual illusions can shed light on the computational principles underlying everyday perception. These illusions, paradigmatic of those investigated by perceptual neuroscience, involve perceptual experience that reliably and systematically disagrees with the 'true' nature of the physical stimulus, despite the fact that the observer is calm and attentive. By contrast, the definition of the word 'illusion' in descriptive psychopathology, a cornerstone of clinical psychiatry[27], is subtly different.

Karl Jaspers, a founder of descriptive psychopathology, defined an illusion as a form of 'false perception', in which 'fresh, unreal objects' are perceived when external sensory stimuli are combined with 'transposing (or distorting) elements'.[28] Two illustrative examples, taken from Jaspers' canonical text, include 'illusions due to affect' (e.g. 'A melancholic patient, beset by fears of being killed [who] may take clothes hanging on the wall for a corpse') and 'illusions due to inattentiveness' (e.g. 'We overlook misprints in a book and complete the meaning correctly according to context')[28]. Here, the emphasis is on the false perception of meaningful 'objects' in the natural world (i.e. a corpse or a word), in contrast to the illusory stimuli of contemporary perceptual neuroscience, which are designed to elicit distorted perceptions of isolated basic features of the visual scene (e.g. luminance and length)[12].

The illusions described by descriptive psychopathology arise when an observer has a high prior expectation about the state of the world, and is confronted with noisy and ambiguous incoming sensory data. The observer's prior expectation about the state of the world may be informed by the semantic context of a situation (in what have come to be termed 'completion illusions'), the observer's current emotional state[29] (in 'affect illusions'), or active imaginative processes acting upon inherently ambiguous sensory data (in 'pareidolic illusions')[28,30–32]. Sensory data may be 'naturally ambiguous' (for example patterns of shadow in cloud formations or a poorly lit visual scene), or may be ambiguous due to inadequate deployment of attention by the observer.

Although this clinical subdivision of illusions[28,30–32] may seem alien to some, we argue that psychopathological illusions that arise in the context of high sensory noise (often caused by low attention) and high prior expectation can be easily accommodated by the hierarchical predictive coding model of perceptual inference. When the prior expectation about the state of the world disagrees with the incoming sensory data, the resulting posterior probability on the real world causes giving rise to sensory data lies between these hypotheses. In the case of psychopathological illusions, the prior expectation is held very confidently (formally, the prior distribution on causes has high precision), perhaps owing to a pathological mental state such as fear or paranoia, leading to a high subjective expectation of threatening objects. Conversely, the sensory data is noisy or poorly attended to (i.e. has low precision). Consequently the optimal perceptual inference about the real world causes underlying the sensory data (i.e. posterior probability) will lie closer to the prior expectation. If an observer subsequently allocates more attention towards the sensory information, the precision of incoming sensory data increases and the resulting posterior probability becomes more ‘veridical’. The Bayesian model, therefore, explicitly parameterizes expectation and attention in perceptual inference, and assigns them orthogonal roles[1,2,10]. The separable roles of expectation and attention in perceptual inference, along with discussions of their plausible neural correlates, have recently been reviewed elsewhere[10].

Conclusions

The account of Bayesian perceptual inference outlined above is thus capable of explaining a number of features of everyday perception, low-level perceptual illusions and the illusory phenomena outlined in descriptive psychopathology. In all cases perceptual inference aims to arrive at the most likely representation of the world by taking into account prior expectations, incoming sensory signals, and the precision that both sources of information contain. Thus, it is not just illusory perception, but *all* perception that results from a combination of mental imagery with sensory stimuli[33].

Despite the promising application of ERT to low level perception and perceptual illusions, it does not easily lend itself to an understanding of object perception[11], or how such perception is affected by states of low attention, heightened emotion or contextually-driven semantic expectations. As these features are necessary for psychopathological illusions to occur it is difficult to discuss these illusions within the context of wholly empirical strategies of perception. These illusions instead require an account of how the brain makes rich, context-dependent predictions about the state that the world is in, and tests these predictions against incoming sensory data.

We have argued that the Bayesian account of perception is well placed to do this. Moreover, in this account the context and meaning of the perceptual scene is given primacy in the process of perceptual inference. Multimodal nodes high in the cortical hierarchy postulate internally-coherent hierarchical representations of the state of the world, already pregnant with behavioral relevance and meaning for the organism, which are ready to be confirmed or refuted by incoming packets of sensory data[2]. This *generative* account of perception is consistent with the intuitions of Phenomenological philosophers like Merleau-Ponty, who argued that we do not construct meaning from meaningless sensory information, but that in perception '*the whole is prior to the parts'*.[34]

References

1. Friston K. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 2010;11(2):127-38.

2. Clark A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 2013;36(3):181-204.

3. Knill DC, Pouget A. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 2004;27(12):712-9.

4. Howe CQ, Lotto R, Purves D. Comparison of Bayesian and empirical ranking approaches to visual perception. *J. Theor. Biol.* 2006;241(4):866-75.

5. Berkeley G. *A New Theory of Vision*. Everyman's Library; 1976.

6. Purves D, Wojtach WT, Lotto RB. Understanding vision in wholly empirical terms. 2011;108(suppl. 3):15588-15595.

7. Friston K. A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 2005;360(1456):815-36.

8. Gold JI, Shadlen MN. The Neural Basis of Decision Making. *Annu. Rev. Neurosci.* 2007;30(1):535-574.

9. Friston K. Learning and inference in the brain. *Neural networks* 2003;16(9):1325-52.

10. Summerfield C, de Lange FP. Expectation in perceptual decision making: neural and computational mechanisms. *Nat. Rev. Neurosci.* 2014;15:745-756.

11. Purves D, Monson BB, Sundararajan J, Wojtach WT. How biological vision succeeds in the physical world. *Proc. Natl. Acad. Sci. U. S. A.* 2014;111(13):4750-4755.

12. Gregory RL. *Eye and Brain: The Psychology of Seeing (5th Ed.)*. Princeton, NJ, US: Princeton University Press; 1997.

13. Brown H, Friston KJ. Free-energy and illusions: The Cornsweet effect. *Front. Psychol.* 2012;3(43):1-13.

14. Purves D, Lotto RB, Williams SM, Nundy S, Yang Z. Why we see things the way we do: evidence for a wholly empirical strategy of vision. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 2001;356(1407):285-97.

15. Shipley WC, Nann BM, Penfield MJ. The apparent length of tilted lines. *J. Exp. Psychol.* 1949;39(4):548-551.

16. Purves D, Shimpi A, Lotto RB. An Empirical Explanation of the Cornsweet Effect. *J. Neurosci.* 1999;19(19):8542-8551.

17. Lotto RB, Purves D. An Empirical Explanation of the Chubb Illusion. *J. Cogn. Neurosci.* 2001;13(5):547-555.

18. Yang Z, Purves D. The statistical structure of natural light patterns determines perceived light intensity. *Proc. Natl. Acad. Sci. U. S. A.* 2004;101(23):8745-8750.

19. Howe CQ, Purves D. Natural-scene geometry predicts the perception of angles and line orientation. *Proc. Natl. Acad. Sci. U. S. A.* 2005;102(4):1228-1233.

20. Wojtach WT, Sung K, Truong S, Purves D. An empirical explanation of the flash-lag effect. *Proc. Natl. Acad. Sci.* 2008;105(42):16338-16343.

21. Morgenstern Y, Rukmini D V., Monson BB, Purves D. Properties of artificial neurons that report lightness based on accumulated experience with luminance. *Front. Comput. Neurosci.* 2014;8(134):1-11.

22. Kersten D, Mamassian P, Yuille A. Object perception as Bayesian inference. *Annu. Rev. Psychol.* 2004;55:271-304.

23. Kok P, Failing M, de Lange F. Prior Expectations Evoke Stimulus Templates in the Primary Visual Cortex. *J. Cogn. Neurosci.* 2014;26(7):1546-1554.

24. Egner T, Monti JM, Summerfield C. Expectation and Surprise Determine Neural Population Responses in the Ventral Visual Stream. *J. Neurosci.* 2010;30(49):16601-16608.

25. Meyer T, Olson CR. Statistical learning of visual transitions in monkey inferotemporal cortex. *Proc. Natl. Acad. Sci.* 2011;108(48):19401-19406.

26. Brown H, Adams RA, Parees I, Edwards M, Friston K. Active inference, sensory attenuation and illusions. *Cogn. Process.* 2013;14(4):411-27.

27. Stanghellini G, Broome MR. Psychopathology as the basic science of psychiatry. *Br. J. Psychiatry* 2014;205(3):169-170.

28. Jaspers K. *General Psychopathology*. Baltimore, Maryland: The Johns Hopkins University Press; 1997.

29. Barrett LF, Bar M. See it with feeling: affective predictions during object perception. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 2009;364(1521):1325-34.

30. Casey P, Kelly B. *Fish's Clinical Psychopathology: Signs and Symptoms in Psychiatry*. 3rd ed. London: Gaskell (R.C.Psych); 2007.

31. Oyebode F. *Sims' Symptoms in the Mind: An Introduction to Descriptive Psychopathology*. 4th ed. Edinburgh: Elsevier; 2008.

32. Semple D, Smyth R. *Oxford Handbook of Psychiatry*. 2nd ed. Oxford: Oxford University Press; 2013.

33. James W. *Psychology: The Briefer Course*. Mineola, NY: Dover Publications; 2001.

34. Merleau-Ponty M. *The Primacy of Perception: And Other Essays on Phenomenological Psychology, the Philosophy of Art, History and Politics (Studies in Phenomenology and Existential Philosophy)*. (Edie J, Cobb W, eds.).; 1964.

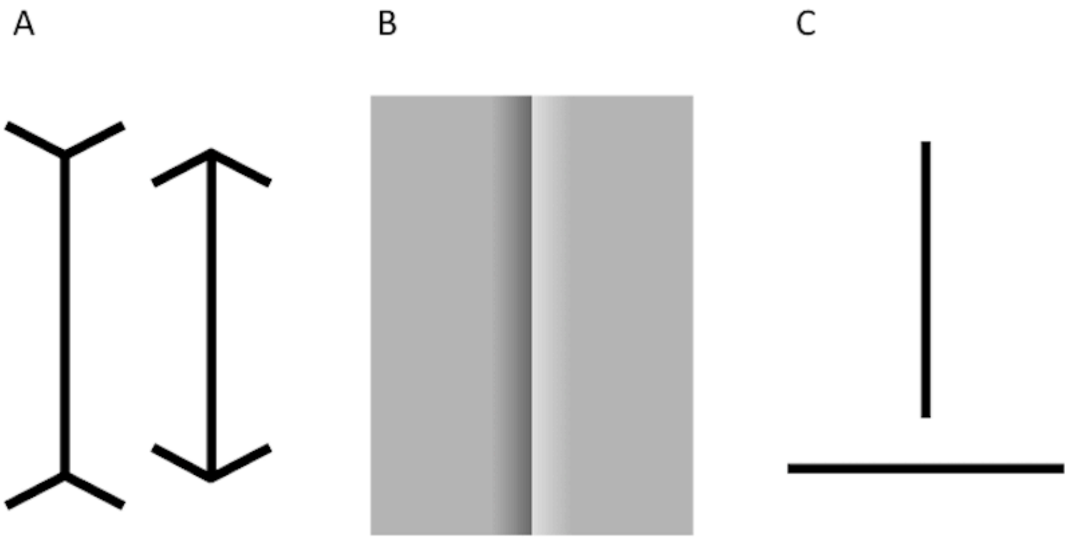

Figure 1. Simple visual illusions. A: Müller-Lyer illusion. The vertical lines are of equal length, yet the left line appears longer. **B:** Craik–O'Brien–Cornsweet illusion. The areas on the left and the right have the same luminance, yet the left shade appears brighter. **C:** Vertical and horizontal lines are the same length, but the vertical line appears longer.