Oxford Handbooks Online

Bayesian Perceptual Psychology

Michael Rescorla

The Oxford Handbook of Philosophy of Perception
Edited by Mohan Matthen

Print Publication Date: Jul 2015 Subject: Philosophy, Epistemology, Philosophy of Science
Online Publication Date: Sep 2013 DOI: 10.1093/oxfordhb/9780199600472.013.010

Abstract and Keywords

Contemporary perceptual psychology uses Bayesian decision theory to develop Helmholtz's view that perception involves 'unconscious inference'. The science provides mathematically rigorous, empirically well-confirmed explanations for diverse perceptual constancies and illusions. The explanations assign a central role to mental representation. This article highlights the explanatory centrality of representation within current Bayesian perceptual models. The article also discusses how Bayesian perceptual psychology bears upon several prominent philosophical topics, including: eliminativism about representation (defended by Churchland, Field, Quine, and Stich); relationalism about perception (endorsed by Brewer, Campbell, Martin, and Travis); phenomenal content (postulated by Chalmers, Horgan and Tienson, and Thompson); and the computational theory of mind (espoused by Fodor and many other philosophers).

Keywords: Bayesian decision theory, unconscious inference, perceptual psychology, perceptual constancies, perceptual illusions, mental representation, eliminativism, relationalism, phenomenal content, computational theory of mind

BAYESIAN decision theory is a mathematical framework that models reasoning and decision-making under uncertainty. Around 1990, perceptual psychologists began constructing detailed Bayesian models of perception.[1] This research program has proved enormously fruitful. As two leading perceptual psychologists put it, 'Bayesian concepts are transforming perception research by providing a rigorous mathematical framework for representing the physical and statistical properties of the environment, describing the tasks that perceptual systems are trying to perform, and deriving appropriate computational theories of how to perform those tasks' (Geisler and Kersten, 2002, 508). To understand perception, one must acquire detailed knowledge of Bayesian perceptual psychology. Or so I hope to convince you.

1 Perception as Unconscious Inference

Perception solves an *underdetermination problem*. The perceptual system estimates environmental conditions, such as the shapes, sizes, colours, and locations of distal objects. It does so based upon *proximal stimulations* of sensory organs. The proximal stimulations underdetermine their environmental causes. For instance, a convex object under normal lighting generates retinal stimulations ambiguous between at least two possibilities: that the object is convex and that light comes from overhead; or that the object is concave and that light comes from below. Similarly, light reflected from a surface generates retinal stimulations consistent with various colours (e.g. the surface may be red and bathed in daylight, or the surface may be white and bathed in red light). In general, then, retinal input underdetermines possible states of the distal environment. We cannot yet programme a (p. 695) computer that solves this underdetermination problem. The perceptual system solves it quickly, effortlessly, automatically, and reliably. How?

Helmholtz (1867) proposed that the perceptual system executes an ‘unconscious inference’ from sensory stimulations to hypotheses about the environment. The inference reflects ‘implicit assumptions’ concerning the environment or the interaction between environment and perceiver. For instance, the visual system deploys an ‘implicit assumption’ that light comes from overhead. Helmholtz’s approach, now called *constructivism,* helps explain two notable phenomena: *perceptual constancies* and *illusions*.

Perceptual constancies are capacities to represent properties or entities as the same despite large variation in proximal stimulation. To varying degrees, human vision displays constancies for numerous properties, including size, shape, location, colour, depth, and motion. How does the perceptual system achieve constancies? By using ‘implicit assumptions’ to discount variations in proximal stimulation. *Colour constancy* provides a good illustration. This is the capacity to perceive surface colour as constant despite large variation in viewing conditions, including background illumination. To estimate surface colour, the perceptual system first deploys various ‘implicit assumptions’ (such as that the light source is fairly uniform, or that certain surface colours are likelier than others) to estimate background illumination based upon overall retinal stimulation. The perceptual system then deploys this background illumination estimate so as to estimate a surface’s colour based upon retinal stimulation caused by that surface. As Helmholtz famously put it, the perceptual system ‘discounts the illuminant’.

Perceptual constancies are reliable but fallible, as demonstrated by *illusions*. Consider again the assumption that light comes from overhead. The assumption is correct in normal cases, so it usually supports an inference to an accurate percept. When the assumption fails, the resulting percept is inaccurate. For instance, lighting a concave object from below generates an illusory percept as of a convex object. Constructivists explain the mistaken shape-estimate by isolating its source: the implicit assumption that light comes from overhead. Similarly, a red spotlight directed upon a single object

violates the implicit assumption of a fairly uniform illuminant, thereby inducing an illusory colour percept. These examples illustrate constructivism's template for explaining illusions: isolate an implicit assumption deployed during perceptual inference; show how failure of the assumption can induce an inaccurate percept.

Perceptual processes are subpersonal and inaccessible to the thinker. There is no good sense in which *the thinker herself,* as opposed to *her perceptual system,* executes perceptual inferences. For instance, a normal perceiver simply sees a surface as having a certain colour. Even if she notices the light spectrum reaching her eye, as a painter might, she cannot access the perceptual system's inference from retinal stimulations to surface colour.[2]

The twentieth century produced various rivals to constructivism, including Gibson's *direct perception* framework. Gibson (1979) denied that perception involves complex psychological activity, inferential or otherwise. He held that the perceptual system directly 'picks up' certain distal properties by 'resonating' to them. Gibson's work yielded many invaluable insights, such as the importance of optic flow, which can be incorporated into constructivism. Viewed as an *alternative* to constructivism, Gibson's direct perception (p. 696) framework has difficulty explaining the vast bulk of constancies and illusions (Fodor and Pylyshyn, 1981). That is why the direct perception framework remains marginal within perceptual psychology.

A satisfactory development of constructivism must answer three questions:

a. In what sense does the perceptual system execute 'inferences'?
b. In what sense do the inferences 'reflect' various 'implicit assumptions'?
c. In what sense does perceptual inference yield the 'best' hypothesis?

Different versions of constructivism answer these questions differently. For instance, some constructivists regard 'implicit assumptions' as stored premises fit to participate in unconscious deductive, inductive, or abductive inferences (Rock, 1983, 272–282). Bayesian perceptual psychology develops constructivism in a different direction, as I will now explain.

2 Perception as Unconscious Statistical Inference

The perceptual system operates under conditions of uncertainty, stemming from at least three sources:

1. *Ambiguity* of sensory input, as described above.
2. *Noisiness* of perceptual organs and neural mechanisms: that is, their vulnerability to corruption by random errors.

3. *Potential conflict* between sensory modalities (e.g. visual versus auditory cues to an object's location) or between cues within a modality (e.g. binocular disparity cues to depth versus monocular linear perspective cues to depth).

It therefore seems natural to formalize constructivism through Bayesian decision theory, which models decision-making under uncertainty.

The core notion underlying Bayesian decision theory is *subjective probability*. Subjective probabilities reflect psychological facets of the individual or her subsystems, rather than 'objective' features of reality. To formalize probabilities, we introduce a *hypothesis space* H containing various hypotheses h. Each hypothesis h reflects a possible state of the world (e.g. a possible shape of some distal object; or a possible colour of some distal surface; or a possible assignment of distal objects to spatial locations). A probability function p maps each hypothesis h to a real number $p(h)$, reflecting the agent's subjective probabilities.[3]

Bayesian decision theory dictates how to update subjective probabilities based on new evidence. *Bayes's Theorem* states that:

$$p(h|e) \propto p(e|h)p(h)$$ (p. 697)

meaning that the left-hand side is proportional to the right-hand side. $p(h \mid e)$ and $p(e \mid h)$ are *conditional probabilities*. For instance, $p(e \mid h)$ is the probability of e, conditional on h. *Bayes's Rule* states that, when one receives evidence e, one should update $p(h)$ by replacing it with $p(h \mid e)$. To execute Bayes's Rule, one multiplies the *prior probability* $p(h)$ by the *prior likelihood* $p(e \mid h)$. One then normalizes so that all probabilities sum to 1. Finally, one adopts the resulting *posterior probability* $p(h \mid e)$ as a revised probability assignment for h. Thus, the new probability of h is proportional to its original probability, multiplied by the likelihood of evidence e given h.[4]

Bayesian perceptual psychologists use this framework to model perceptual inference (Knill and Richards, 1996). On a Bayesian approach, the perceptual system entertains hypotheses drawn from a hypothesis space H. The perceptual system assigns prior probabilities to hypotheses h and prior likelihoods to (e, h) pairs, where each e corresponds to some possible sensory input. After receiving input e, the perceptual system reallocates probabilities across the hypothesis space, in rough accord with Bayes's Rule.

To illustrate, consider the extraction of *shape from shading* (Mamassian, Landy, and Maloney, 2002). Let s reflect possible shapes, θ reflect possible lighting directions, and e reflect possible patterns of retinal illumination. The visual system encodes:

A prior probability $p(s)$, which assigns higher probability to certain distal shapes than others (e.g. it may assign higher probability to convex shapes).

A prior probability $p(\theta)$, which assigns higher probability to an overhead lighting direction than to alternative lighting directions.

A prior likelihood $p(e \mid s, \theta)$, which assigns a higher probability to an (e, s, θ) triplet if distal shape s and lighting direction θ are likely to cause retinal illumination e.

Upon receiving retinal illumination e, the perceptual system redistributes probabilities over shape-estimates, yielding a posterior $p(s \mid e)$. Depending on the case, the posterior might assign a much higher probability to convexity than concavity. For details, see Stone (2011).

Perception normally yields a determinate percept. For instance, one sees an object as having a determinate shape, not a spectrum of more or less probable shapes. Accordingly, Bayesian models explain how the perceptual system selects a single hypothesis h based on the posterior $p(h \mid e)$. Typical models invoke *expected utility maximization.* The 'action' is selection of h. The *utility function,* which is task-dependent, reflects the penalty for an incorrect answer. If the utility function has a suitable shape, then expected utility maximization reduces to a much simpler decision rule, such as selecting the mean or the mode of the posterior probability.

As another example, Bayesian models of surface colour perception proceed roughly as follows. A surface has reflectance $R(\lambda)$, specifying the fraction of incident light that the surface reflects at each wavelength λ.[5] The illuminant has spectral power (p. 698) distribution $I(\lambda)$: the light power at each wavelength. The retina receives light spectrum $C(\lambda) = I(\lambda)\, R(\lambda)$ from the surface. The visual system seeks to estimate surface reflectance $R(\lambda)$. This estimation problem is underdetermined, since $C(\lambda)$ is consistent with numerous $I(\lambda)$-$R(\lambda)$ pairs. Typical Bayesian models posit that two surfaces have the same colour appearance for a perceiver when her perceptual system estimates the same reflectance for each surface. To estimate $R(\lambda)$, the visual system estimates $I(\lambda)$. It does so through a Bayesian inference, based upon overall retinal stimulation, that deploys a prior probability over possible illuminants and possible surface reflectances. To a first approximation, the illuminant prior assigns higher probability to illuminants that resemble natural daylight, while the surface prior assigns higher probability to surface reflectances that occur more commonly in the natural environment. This framework can explain both the success and the failure of human colour constancy under various conditions. For details, see Brainard (2009).[6]

We can schematize a typical Bayesian model through the template depicted in Figure 37.1. Note that this template does not require perception to *represent* Bayesian norms. There is no evidence that the perceptual system explicitly represents Bayes's Rule or expected utility maximization. The perceptual system simply proceeds *in rough accord with* Bayesian norms.

A typical Bayesian model dictates a unique outcome given four factors: prior probabilities, prior likelihoods, sensory input, and the utility function.[7] In that sense, the model is deterministic. Of course, the model's generalizations are *ceteris paribus*.

Perceptual malfunction, external interference, or corruption by internal noise can induce exceptions.

Most Bayesian models conform roughly to the foregoing template. But some models vary the template. For instance, some models augment the template by incorporating *motor efference copy*.[8] Other models replace expected utility maximization with *probability matching,* a non-deterministic process whereby the probability that the perceptual system selects some hypothesis matches the posterior probability assigned to that hypothesis (Mamassian, Landy, and Maloney, 2002). One phenomenon sometimes analyzed through non-deterministic Bayesian modelling is *multistable perception* (such as the Necker cube). During multistable perception, experience fluctuates between distinct percepts, rather than yielding a unique percept.

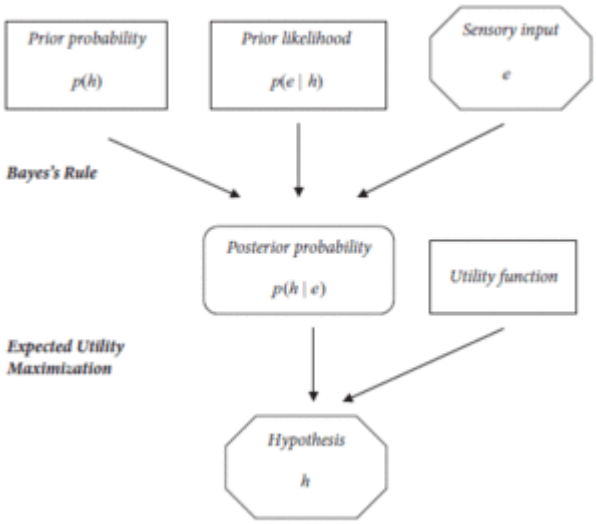

Fig. 37.1 A template for Bayesian models of perception.

One can construe Bayesian models of perception in two different ways (Kersten and Mamassian, 2009). On the first construal, a Bayesian model describes how an 'ideal (p. 699) observer' would estimate environmental conditions based upon sensory input. We construct the model only so as compare human performance with an ideal benchmark. On the second construal, a Bayesian model approximately describes *actual* mental processes. The model seeks to describe, perhaps in an idealized way, how the perceptual system *actually* transits from sensory input to perceptual estimates. Both construals figure in the scientific literature. I emphasize the second construal. I am discussing Bayesian models as empirical theories of actual human perception.

Many Bayesian models are fairly unrealistic. For example, the hypothesis space is often uncountable. In general, Bayesian inference over an uncountable hypothesis space is computationally intractable. So I think that we should regard most Bayesian perceptual models as *idealizations,* akin to models from physics that postulate frictionless surfaces or infinite wires. Of course, we eventually want less idealized descriptions. However, I see no principled problem here. Artificial Intelligence (AI) offers numerous tools for constructing computationally tractable approximations to idealized Bayesian computation. No doubt we will eventually supplement or replace current perceptual models with computationally tractable approximations.

Bayesian perceptual psychology provides detailed answers to the three questions (a), (b), and (c) posed at the end of the previous section:

a. Transitions among perceptual states approximately conform to norms of Bayesian inference. *In that sense,* the transitions are statistical inferences.
b. Bayesian models replace talk about 'implicit assumptions' with talk about prior probabilities and likelihoods. The models thereby depart substantially from many earlier versions of constructivism. On Rock's approach, for example, an 'implicit (p. 700) assumption' that light comes from overhead corresponds to a single stored premise whose content is that light comes from overhead. Bayesians instead posit a prior assignment of probabilities to possible lighting directions. This prior figures not as a premise but rather as input to Bayesian reallocation of subjective probabilities over shape-estimates.
c. The perceptual system produces an estimate that is 'best' or 'optimal' insofar as it conforms to rational norms of Bayesian decision theory. In this manner, Bayesian models depict numerous perceptual illusions as natural by-products of a near-optimal process that infers environmental conditions from ambiguous sensory stimulations.

Hence, the Bayesian framework converts talk about 'implicit assumptions' and 'unconscious inferences' into mathematically rigorous, quantitatively precise psychological models.

Where do the prior probabilities and prior likelihoods come from? The human visual system evolved over millennia in a fairly stable environment. Accordingly, one might expect certain lawlike or statistical environmental regularities to be 'encoded in the genes'. Nevertheless, Bayesian perceptual priors do not simply reflect innate programming. For instance, even the 'light-from-overhead' prior reflects a complex interplay between nature and nurture. It gathers considerable strength during early childhood (Stone, 2011), and it changes rapidly upon adult exposure to deviant environments (Adams, Graf, and Ernst, 2004). At present, we do not know how genetic endowment and individual experience jointly determine Bayesian priors. Current research mainly tries to identify the priors, not to explain the aetiology of the priors.[9]

Ultimately, we want detailed theories explaining how Bayesian priors originate and develop. Even lacking such theories, we can cite the priors to explain constancies and illusions. In this connection, I stress that the priors postulated by Bayesian perceptual psychology are not *ad hoc*. Admittedly, a precise quantitative match usually requires some 'curve fitting'. Qualitatively, though, the priors typically reflect antecedently motivated claims about lawlike or statistical properties of our environment. It is plausible that the perceptual system acquires these priors through some combination of nature and nurture, even if we do not yet know how.[10]

How can we legitimately postulate Bayesian priors, lacking a developed theory of their aetiology? Because Bayesian priors generate the *unifying predictive power* characteristic of good explanation. To illustrate, consider *motion perception*.[11] The visual system can (p. 701) directly measure local retinal image velocities, which underdetermine the distal

motions that cause them. The visual system must estimate distal motion based upon local retinal image velocities. It does so fairly well but not perfectly, as illustrated by the fact that low-contrast stimuli appear to move more slowly than high-contrast stimuli. (This may explain why drivers accelerate in the fog—they underestimate relative velocities.) Weiss, Simoncelli, and Adelson (2002) offer a Bayesian motion perception model with two features:

The prior probability favours slow distal motions.
The visual system treats low-contrast retinal images as less reliable.[12]

This model explains why vision underestimates velocity under low-contrast conditions: namely, because the slow-motion prior exerts more influence over the velocity-estimate. The model also explains other motion illusions, including the following: a fat rhombus moving horizontally appears to move horizontally, but a thin rhombus seems to move diagonally at low contrasts and horizontally at high contrasts. (Readers can experience this effect at <www.cs.huji.ac.il/~yweiss/Rhombus/rhombus.html>.) Thus, a single Bayesian model explains diverse illusions that otherwise resist unified treatment. Subsequent models have elaborated the Bayesian approach to motion perception in increasingly sophisticated ways (Ernst, 2010).

Bayesian perceptual psychology offers illuminating, rigorous explanations for numerous constancies and illusions. It is our best current science of perception. We should carefully consider how it bears upon contemporary philosophy of mind—a task to which I now turn.

3 Estimation and Representation

A natural view holds that perceptual states are evaluable as *accurate* or *inaccurate*. For instance, suppose I perceive a concave object that appears convex due to misleading lighting. It seems natural to say that my percept is inaccurate. To say this, we must ascribe truth, accuracy, or veridicality conditions to the percept. Some philosophers distinguish among ‘truth’, ‘accuracy’, and ‘veridicality’ (Burge, 2010), but I remain neutral on this issue. Call the view that perceptual states have veridicality-conditions *representationalism*. Burge (2005, 2010, 2011) argues that current perceptual psychology supports representationalism. I will now defend the same conclusion by examining Bayesian models of perception.[13]

On the Bayesian approach, perceptual inference reallocates probabilities over a hypothesis space and then selects a favoured hypothesis. This favoured hypothesis is incorporated into the final percept, whose accuracy depends upon whether the hypothesis is accurate. (p. 702) To illustrate, consider Bayesian models of shape perception. The perceptual system assigns prior probabilities to estimates *of specific distal shapes*. After receiving sensory input, perceptual inference revises the probability assignment and selects a favoured estimate *of a specific distal shape*. The resulting

percept incorporates this favoured shape-estimate. The percept may also incorporate various size-estimates, motion-estimates, and so on. Accuracy of the percept depends upon accuracy of the individual estimates. By describing perceptual inference in this way, we type-identify perceptual states representationally. We individuate perceptual states partly through environmental conditions that must obtain for the states to be accurate.

What exactly are the accuracy-conditions of percepts? According to Davies (1992), a percept involves something like *existential quantification*. The percept is accurate when *there exist* objects with properties represented by the percept. An opposing view, espoused by Burge (2005), holds that perceptual accuracy-conditions are *object-dependent*. A percept represents environmental *particulars,* such as physical bodies or events. The percept attributes properties to those particulars. It is accurate only if *those particulars* have the represented properties. I remain neutral between these two views. I emphasize a shared presupposition underlying both views: that perceptual states *have* accuracy-conditions. This presupposition is integral to perceptual psychology. The science seeks to explain how the perceptual system generates a percept that estimates specific environmental conditions. Estimates can be either accurate or inaccurate.

Following standard philosophical usage, I say that a mental state has *representational content* when it has a veridicality-condition. On this usage, perceptual states have representational content. I do not assume a specific theory of representational content. One might gloss perceptual contents as sets of possible worlds, or Russellian propositions, or Fregean senses. There are many other options.[14] The key point for us is that the science routinely individuates perceptual states through their representational import.

Bayesian models individuate both *explananda* and *explanantia* in representational terms. The science explains perceptual states *under representational descriptions,* and it does so by citing other mental states *under representational descriptions*. For instance, Bayesian models of shape from shading assume prior probabilities over hypotheses *about specific distal shapes* and *about specific lighting directions*. The models articulate generalizations describing how retinal input, combined with these priors, causes a revised probability assignment to hypotheses *about specific distal shapes,* subsequently inducing a unique estimate *of a specific distal shape*. The generalizations type-identify perceptual states *as* estimates *of* specific distal shapes. Similarly, Bayesian models of surface colour perception type-identify perceptual states *as* estimates *of* specific surface reflectances. Thus, the science assigns representation a central role within its explanatory generalizations. The generalizations describe how mental states *that bear certain representational relations to the environment* combine with sensory input to cause mental states *that bear certain representational relations to the environment*.

In what follows, I develop my analysis by examining various philosophical theories that either reject representationalism or else downplay the importance of representational content.

(p. 703) 4 The Relational View of Perception

Brewer (2007), Campbell (2010), Martin (2004), and Travis (2004) espouse a *relational* view of perception. Relationalists eschew all talk about perceptual representation. They treat perceptual states as relations not to representational contents but rather to objects or properties in the perceived environment. For instance, Campbell (2010, 202) holds that 'the content of visual experience is constituted by the objects and properties in the scene perceived', rather than by anything resembling an accuracy-condition. He cautions that we should not 'think of experience itself as already a representational state' (ibid.). The relational approach is sometimes allied with Gibsonian direct perception, sometimes not.

To illustrate, consider two counterfactual situations A and B in which I perceive the same object O, yielding qualitatively indistinguishable percepts P_A and P_B:

In situation A, O is convex and looks convex.
In situation B, O is concave but looks convex through misleading lighting.

Representational taxonomization type-identifies P_A and P_B by correlating them with the same accuracy-condition. In particular, both percepts estimate the same distal shape: convexity. In situation A, the estimate is correct. In situation B, the estimate is incorrect. By contrast, Campbell's relational taxonomization treats P_A and P_B as type-distinct. Campbell type-identifies the first percept through its relation to a distal property (convexity) to which the second percept is not appropriately related.

Bayesian perceptual psychology supports representationalism over relationalism.

A core postulate underlying the science is that perception produces an estimate of environmental conditions, where the estimate may be either accurate or inaccurate. Consider Figure 37.1. If we neglect noise, malfunction, and external interference, then a unique percept-type is determined by four factors: the prior probability, the prior likelihood, proximal sensory input, and the utility function. We may stipulate that all four factors are the same in situations A and B. It follows that percepts P_A and P_B are type-identical *from the perspective of the Bayesian model*. In both cases, the final percept incorporates a convexity-estimate. The perceptual system produces a convexity-estimate *whether or not the perceived object is convex*. (Cf. Burge, 2005, 22–25; 2010, 362–364.) An appropriately modified diagnosis applies to non-deterministic Bayesian models, such as models that replace expected utility maximization with probability matching. For such models, the probability that situation A yields a convexity-estimate equals the probability that situation B yields a convexity-estimate. Thus, explanatory generalizations of Bayesian perceptual psychology enshrine a representational, non-relational taxonomic scheme. The generalizations type-identify percepts by specifying environmental conditions that must obtain for a given percept to be accurate.

Campbell (2010) suggests that we can interpret perceptual science in relational terms. This suggestion seems unpromising, because the Bayesian explanation of illusion relies essentially upon non-relational taxonomization. The central idea is that the perceptual system estimates some environmental state, *which may or may not obtain*. Bayesian modelling seeks to explain the environmental state estimate, regardless of whether the (p. 704) estimate is veridical. Contrary to Campbell's relationalist strictures, the science routinely type-identifies veridical and non-veridical percepts. Of course, there is a difference between the veridical and the non-veridical percept. Perceptual psychologists acknowledge this difference. Yet they also emphasize fundamental representational commonalities between the two percepts. Those commonalities play a key individuative role within Bayesian explanatory generalizations. So a relational, non-representational taxonomic scheme flouts explanatory practice within perceptual psychology.

Brewer (2007, 173) seeks to accommodate illusions inside a relational framework. He concedes that there can be a 'visually relevant similarity' between a veridical and a non-veridical percept. He compares: (i) a red surface in daylight; and (ii) a white surface surreptitiously bathed in red light. He acknowledges that the surface in scenario (ii) looks red. He says that 'this consists in the fact that [the surface] has visually relevant similarities with paradigm red objects: the light reflected from it is like that reflected from such paradigms in normal viewing conditions' (ibid.).

Naturally, I agree that (i) and (ii) emit similar light spectra. However, merely noting this commonality does not capture the fact that both surfaces look red. A surface that emits the same light spectrum under different viewing conditions may not look red. A surface that emits a radically different light spectrum under different viewing conditions can still look red. Thus, we must reject Brewer's proposed analysis of *looks red*. In contrast, representationalists can say that a surface looks red when one's percept represents the surface as red.

Brewer's account omits crucial *scientifically relevant* commonalities between the two percepts. A key scientifically relevant commonality is that both percepts result from perceptual estimation of a single surface reflectance $R(\lambda)$. The estimate is correct in (i), incorrect in (ii). We do not capture this key commonality between the percepts simply by noting that (i) and (ii) emit similar light. The perceptual system can estimate reflectance $R(\lambda)$ despite large variation in the light spectrum $C(\lambda)$ emitted by a surface. Moreover, depending on the perceptual system's estimate of illumination $I(\lambda)$, it may not estimate $R(\lambda)$ even when the surface emits the same light spectrum $C(\lambda)$. Capturing the scientifically relevant commonalities between (i) and (ii) requires us to cite perceptual estimation (and hence perceptual representation) of surface reflectance. Yet relationalists eschew all talk about perceptual representation.

There are delicate issues here surrounding the relation between colours and surface reflectances. According to current science, a percept that represents a surface as red is

caused by perceptual activity that represents reflectance. But does the final percept itself represent reflectance? There are at least three salient options:

a. The percept represents colour but not reflectance.
b. The percept represents reflectance and *separately* represents colour.
c. The percept represents reflectance and *thereby* represents colour.

The choice between (a), (b), and (c) depends upon other matters, including the metaphysics of colour (cf. note 6). We need not choose among (a)–(c) here. The crucial point is that relationalists must reject all three options. Relationalists do not countenance perceptual representation of colour, reflectance, or any other distal property.

In summary, relationalism cannot accommodate a core postulate underlying contemporary perceptual psychology: that perception produces an estimate of environmental conditions, where the estimate may be either accurate or inaccurate.

(p. 705) 5 Eliminativism, Instrumentalism, and Realism

Beginning with Quine (1960), various philosophers have argued that *intentionality* (or *representationality*) deserves no place in serious scientific discourse. They have argued that we should replace intentional psychology with some alternative framework, such as Skinnerian behaviourism (Quine, 1960) or neuroscience (Churchland, 1981). This *eliminativist* position concedes that representational locutions are instrumentally useful in everyday life. It denies that they offer literally true descriptions. Dennett (1987) advocates a broadly instrumentalist position intermediate between intentional realism and eliminativism. He acknowledges that the ‘intentional stance’ is instrumentally useful for scientific psychology, but he questions whether mental states *really* have representational content.

I assume a broadly scientific realist perspective: explanatory success is a *prima facie* guide to truth. From a scientific realist perspective, the explanatory success of Bayesian perceptual psychology provides *prima facie* reason to attribute representational content to perceptual states. The science is empirically successful and mathematically rigorous. It routinely individuates perceptual states through representational relations to the environment. We have no clue how to preserve the resulting explanatory benefits without employing representational locutions. Thus, current perceptual psychology strongly supports intentional realism over eliminativism and instrumentalism. We should no more adopt an eliminativist or instrumentalist posture towards intentionality than we should adopt an eliminativist or instrumentalist posture towards electrons. The famous Quinean criticisms of intentional psychology are notably less rigorous and compelling than the science they purport to undermine. Philosophers who reject intentionality as spooky,

obscure, or otherwise unscientific are in fact opposing our current best science of perception.

One might greet my argument by proposing an instrumentalist interpretation of perceptual psychology. In this vein, McDowell insists that appeals to representational content within perceptual psychology are ‘metaphorical’ (2010, 250). On his analysis, perceptual psychologists do not *literally* claim that perception represents. They claim only that perception proceeds *as if* it represents. Representational talk is mere heuristic.

McDowell’s proposal misinterprets perceptual psychology. (Cf. Burge, 2011, 67–68.) A fundamental idea underlying how the science treats illusion is that a perceptual estimate can be *inaccurate*. An estimate is accurate only if the environmental conditions that it estimates actually obtain. Thus, intentional attribution is embedded within the foundations of the science. Representational locutions do not play a metaphorical role within Bayesian perceptual psychology. They are not heuristic chitchat. They reflect the central, explicit goal of the science: to describe how the perceptual system estimates environmental conditions. Instrumentalism is no more justified toward Bayesian perceptual psychology than toward any other science.

Even readers who reject full-blown instrumentalism may contemplate a *moderate instrumentalist* agenda: construe representational description literally when applied to *explananda* but metaphorically when applied to *explanantia*. Consider again Figure 37.1. Moderate instrumentalism adopts a realist stance towards sensory input *e* and the output hypothesis *h* but an instrumentalist stance towards the priors, posterior, and utility function. On this approach, the priors, posterior, and utility function are simply useful tools (p. 706) for predicting how certain sensory inputs cause certain perceptual states. The perceptual system transits from retinal input to perceptual estimates *as if* it encodes Bayesian priors. Moderate instrumentalism concedes that the perceptual system implements a mapping from sensory inputs to perceptual estimates, but it remains neutral regarding how the perceptual system implements that mapping. For defence of moderate instrumentalism regarding Bayesian perceptual psychology, see Colombo and Seriès (2012).

Moderate instrumentalism does not flout the science as blatantly as full-blown instrumentalism. Nevertheless, it strikes me as unsatisfactory. A key point here is that experience can alter the mapping from proximal input to perceptual estimates. For example, Adams, Graf, and Ernst (2004) manipulated the light-from-overhead prior by exposing subjects to deviant haptic feedback regarding shape. The new prior caused altered shape-estimates. Moreover, the new prior *transferred* to a different task that required subjects to estimate which side of an oriented bar was lighter than the other. Realists can offer a principled,unified explanation for the altered shape-estimates and lightness-estimates: namely, that there is a change in the prior over lighting directions. Moderate instrumentalists seem unable to offer a comparably satisfying explanation. Moderate instrumentalists must simply say that the mapping from retinal input to shape-estimates changes *and* that the mapping from retinal input to lightness-estimates

changes, without offering any underlying explanation for why the mappings change as they do. In this case, at least, realism seems more explanatorily fruitful than moderate instrumentalism.[15]

We must exercise care in stating the realist position. As already noted, current Bayesian models are highly idealized. When the hypothesis space is large enough, the perceptual system may only *approximately encode* the priors and the posterior. What does it mean to 'approximately encode' a probability assignment? What is the difference between saying that the mind *approximately* implements Bayesian inference and saying that the mind merely behaves *as if* it implements Bayesian inference?[16] These questions—which lie at the intersection of philosophy, AI, and empirical psychology—merit extensive further study.

6 Phenomenal Content

Relatively few philosophers reject representationalism. However, many popular philosophical theories downplay perceptual representation of the distal environment. Most of these theories are *consistent with but unsupported by* contemporary science. I will now illustrate by considering *phenomenal content*, as postulated by Chalmers (2006), Horgan and Tienson (2002), Thompson (2010), and various other philosophers.

A distinguishing feature of phenomenal content is that it supervenes upon phenomenal aspects of experience. For example, suppose that a normal perceiver Nonvert observes (p. 707) a red object and experiences a perceptual state with a certain phenomenological character. Suppose that a spectrally inverted perceiver Invert observes a green object and experiences a phenomenally indistinguishable perceptual state. Chalmers and Thompson hold that, in both cases, the resulting percept is veridical. Nonvert's percept correctly attributes redness, while Invert's percept correctly attributes greenness. Chalmers and Thompson also hold that the two percepts share a uniform phenomenal content. The content represents red *as used by Nonvert* and green *as used by Invert*. Similarly, Chalmers and Thompson hold that a single phenomenal content might represent circularity *as used by one perceiver* and non-circular ellipticality *as used by a phenomenological twin suitably embedded in a sufficiently different environment*.

There may be many good reasons for positing phenomenal contents. However, Bayesian perceptual psychology makes no use of such contents. The science delineates explanatory generalizations dictating how mental states *that represent certain environmental properties* induce other mental states *that represent certain environmental properties*. Bayesian models describe how the perceiver, exercising standing capacities to represent *specific environmental properties*, executes perceptual inferences yielding estimates of *specific environmental properties*. To illustrate, let us follow Thompson (2010) by considering phenomenological twins embedded in such different environments that one twin's percept *P* represents circularity while the other twin's qualitatively

indistinguishable percept P^* represents non-circular ellipticality. There may be many worthy explanatory projects that type-identify P and P^*. But Bayesian perceptual psychology does not type-identify the two percepts. The science studies perceptual estimation of environmental conditions. P and P^* estimate radically different environmental conditions: P estimates circularity, while P^* estimates non-circular ellipticality. The science features no explanatory generalizations that assimilate these two percepts, because the relevant generalizations are tailored to specific shapes. Phenomenological overlap per se is irrelevant to the current science. What matters is representational overlap.

Similarly, suppose that Nonvert observes a red object, while spectrally inverted Invert observes a green object. Chalmers and Thompson associate the resulting qualitatively indistinguishable percepts with a shared phenomenal content. In contrast, Bayesian perceptual psychology does not type-identify the percepts. Bayesian models treat surface colour perception as involving estimation of reflectance. Explanatory generalizations cite representational relations to specific reflectances. Current Bayesian models of Nonvert describe how retinal illumination $C(\lambda)$ induces an estimate of illuminant $I(\lambda)$, subsequently inducing an estimate of reflectance $R(\lambda)$. Current Bayesian models of Invert describe how different retinal illumination $C^*(\lambda)$ induces an estimate of a different reflectance $R^*(\lambda)$. Reflectance-estimate $R(\lambda)$ as used by Nonvert and reflectance-estimate $R^*(\lambda)$ as used by Invert may be associated with the same phenomenology. But this phenomenological overlap is irrelevant to the science. No explanatory generalizations type-identify the relevant perceptual processes. At no level of description does current science assimilate Nonvert's colour perception and Invert's colour perception.[17] (p. 708)

Current perceptual psychology individuates perceptual states by citing representational relations to specific environmental properties.[18] Taxonomization through phenomenal content ignores these representational relations. I conclude that phenomenal content is an armchair construct with no grounding inside contemporary science. Readers must judge for themselves whether philosophical energy is better expended studying this armchair construct or analyzing our current best science of perception.

7 The Computational Theory of Mind

I now want to consider the relation between Bayesian perceptual psychology and the popular philosophical view that mental activity involves *computation* over formal syntactic types in a *language of thought* (Field, 2001), (Fodor, 2008), (Stich, 1983). The paradigm here is a Turing machine manipulating formal syntactic items, such as stroke marks, inscribed in memory locations. A formal syntactic type may *have* a meaning. But it could have had a *different* meaning, just as the English word 'cat' could have denoted dogs. Depending on the perceiver's causal or evolutionary history, a formal syntactic type that represents some distal property could just as easily have represented some other distal

property. Formal syntactic manipulation is not sensitive to such changes in meaning. Transition rules governing mental computation allude solely to ‘local’ syntactic properties of mental states, without citing representational relations to the external environment.

Field (2001) and Stich (1983) combine the formal syntactic picture with *eliminativism*. They urge scientific psychology to eschew any talk about representational content. Fodor (2008) combines the formal syntactic picture with *intentional realism*. In particular, he urges scientific psychology to delineate causal laws that cite representational content. He holds that intentional laws are *implemented* by syntactic mechanisms. So Fodor assigns a central role to representational content *in addition to* formal syntactic manipulation.

Egan (1992) argues that perceptual psychology postulates formal syntactic manipulation. She defends her conclusion by analyzing the writings of Marr (1982). I set aside whether Egan correctly describes Marr’s work, which was historically important but is now outdated.[19] I claim that the formal syntactic picture finds no support within *current* perceptual psychology, as epitomized by Bayesian modelling. Current perceptual psychology individuates mental computations in representational rather than formal syntactic (p. 709) terms (Burge, 2010, 95–101). For instance, Bayesian models of shape perception describe a computation whereby the visual system reallocates probabilities over hypotheses *about distal shape*. Each hypothesis is individuated partly by its representational relation to a specific distal shape. Transition rules governing the computation derive from Bayesian norms. Of course, the transition rules characterize *initial sensory inputs* (such as retinal inputs) physiologically rather than representationally. Crucially, though, the rules use representational vocabulary to characterize the perceptual states caused by initial sensory inputs. The rules do not cite formal syntax when characterizing sensory inputs (which are described physiologically) or ensuing perceptual states (which are described representationally). Bayesian models do not cite formal syntactic items divested of representational import.[20]

A complete science of perception must illuminate the *neural mechanisms* that implement Bayesian computation.[21] Thus, a complete theory should include non-representational *neural* descriptions. But should it include non-representational *syntactic* descriptions? Syntax is supposed to be *multiply realizable*, in the sense that systems with wildly different intrinsic physical constitutions can satisfy the same syntactic description (Fodor, 2008, 91). Systems may be homogeneous under syntactic description but heterogeneous under neural description. Should a good theory posit formal syntactic types that are multiply realizable *and* that underdetermine representational content? There may be many good reasons for positing formal syntactic types with these features. Yet no such types figure in current perceptual psychology. The science does not employ computational descriptions that prescind from both representational and neural details. Eliminativist versions of the formal syntactic picture *conflict* with current perceptual psychology. Intentional realist versions of the formal syntactic picture are *consistent with but unsupported by* current perceptual psychology.

A common rejoinder is that we can reinterpret intentional explanations in formal syntactic terms, without explanatory loss. In this vein, Field (2001, 72–82, 153–156) proposes a version of Bayesian modelling on which subjective probabilities attach to formal syntactic items individuated without regard to meaning or content. He claims that this framework can preserve any alleged explanatory benefits offered by intentional explanation.

Field's proposal is *revisionary* regarding contemporary psychology. Current science individuates perceptual states representationally. Field proposes an alternative scientific framework that individuates perceptual states in formal syntactic terms. Whether an alternative hypothesis subserves equally good explanations is not a question to be settled *a priori*. Proponents must first develop the alternative hypothesis in rigorous mathematical and empirical detail. Field must reconstruct current science, expunging any apparent reference to representation. Yet he does not indicate how to execute the needed reconstruction for a single real case study. He does not demonstrate through a single real example that his approach can replicate the explanatory benefits offered by intentional explanation within Bayesian psychology. Thus, Field's proposal amounts to an unsupported conjecture that we can gut perceptual psychology of a central theoretical construct without explanatory loss. We have no reason to believe this conjecture, absent detailed confirmation.[22] (p. 710) Generally speaking, we cannot radically alter how a science individuates its subject matter while preserving the science's explanatory shape. We should not expect that we can transfigure the taxonomic scheme employed by current Bayesian models while retaining the explanatory benefits provided by those models.

In her later writings, Egan (2010) avoids talk about formal syntactic manipulation. Instead, she claims that computational models of perception offer "abstract mathematical descriptions" that ignore representational properties of perceptual states. This new account shares a crucial feature with the formal syntactic picture. Both accounts prioritize non-intentional, non-neural computational descriptions. As I have argued, no such descriptions figure in Bayesian perceptual psychology.

Philosophers motivate non-intentional computational modelling through various arguments. One popular argument emphasizes *explanatory generality* (Egan, 2010; Stich, 1983, 160–170). Following Egan (2010), consider a creature Visua whose perceptual states represent some environmental property (such as depth). Imagine a neurophysiological duplicate Twin Visua embedded in such a radically different environment that its corresponding perceptual states do not represent the same property.[23] A non-intentional computational description can type-identify the doppelgangers. We cannot type-identify the doppelgangers if we classify perceptual states through representational relations to the environment. Shouldn't we prefer the more general theory?

Assessing the merits of this argument is a large task that lies beyond our main focus. The key point for present purposes is that Bayesian perceptual psychology does not type-identify Egan's putative neurophysiological twins. The science explains how perceptual

systems of terrestrial animals transit from sensory input to hypotheses that represent specific environmental properties. It studies terrestrial animals endowed with standing capacities to represent specific environmental properties. Its scope is not intergalactic. It does not seek to accommodate chimerical creatures imagined by philosophers. Whatever the putative explanatory benefits of non-intentional computational modelling, our actual best science of perception individuates perceptual states partly through representational relations to specific environmental properties.

8 An Abstract Mathematical Description?

To bolster my assessment, I will now examine more carefully the role that probability theory plays within Bayesian modelling. Interested readers can consult any standard probability-theory textbook for the technical background to my discussion.

Fig. 37.2 The probability density function for a Normal distribution.

Probability theory, as axiomatized by Kolmogorov, posits a *sample space* Ω whose elements are possible ‘outcomes’. Kolmogorov’s axioms place no restrictions on elements of Ω. (p. 711) If Ω is discrete, then we can assign probabilities directly to its elements. If Ω is continuous, then we instead assign probabilities to privileged *subsets* of Ω. We introduce a *σ-algebra* over Ω (i.e. a set of subsets of Ω that contains Ω and is closed under countable union and complementation in Ω). A *probability measure* assigns a probability (a real number) to each element of the σ-algebra.

A *random variable* is a measurable function from Ω to the real numbers ℝ.[24] A probability measure and a random variable jointly induce a *probability distribution*: an assignment of probabilities to privileged subsets of ℝ. Intuitively, the random variable lets us transform a probability assignment involving Ω into a probability assignment involving ℝ.[25] The probability distribution exists entirely within the realm of abstract mathematical entities. By citing the random variable and the probability distribution, we vastly increase the elegance and utility of our mathematical formalism. In particular, we can now apply real analysis to probabilistic modelling.

When Ω is continuous, we can often introduce a *probability density function* (pdf), which carries each element of ℝ to a *probability density* (also drawn from ℝ). A famous example is the *Normal* (or *Gaussian*) *distribution*, whose associated probability density function is depicted in Figure 37.2. The probability that a random variable attains a value within some region is found by integrating the pdf over that region. In other words, the

probability assigned by the *probability distribution* to a region equals the integral of the pdf over that region.[26] A pdf is a purely mathematical entity, just like a probability distribution.

To apply probability theory to psychological modelling, we must specify the nature of the underlying sample space Ω. When we seek to model perception, we should construe Ω's elements as perceptual estimates or hypotheses. For instance, if we are modelling depth perception, then we should construe each element of Ω as a perceptual estimate of some particular depth. One might gloss 'perceptual estimates' as mental representations, or Russellian propositions, or Fregean senses, or sets of possible worlds, and so on. The key point is that we individuate perceptual estimates at least partly through the environmental properties that the estimates represent. As I have argued, this is how the science (p. 712) typically individuates perceptual estimates. Once we have introduced an underlying sample space, we can also introduce appropriate random variables. To illustrate, suppose that Ω contains depth-estimates. Then we can introduce a random variable D that maps each depth-estimate h to a real number $D(h)$. Depending on our choice of D, the real number $D(h)$ might be the depth estimated by h as measured in metres, or as measured in feet, and so on.

In practice, Bayesian perceptual psychologists rarely highlight the underlying sample space Ω. Typical models, including all the models described in this chapter, instead emphasize probability distributions or pdfs. For instance, Jacobs (1999) posits a pdf for a random variable corresponding to depth. A pdf is a purely mathematical entity. By specifying it, we do not specify a unique sample space Ω. The pdf is consistent with numerous sample spaces.

At first blush, the scientific emphasis on probability distributions and pdfs may seem to undermine my representationalist interpretation of Bayesian perceptual psychology. Consider once again Visua, whose perceptual states represent depth, and doppelganger Twin Visua, whose corresponding states do not represent depth. According to Egan, explanatory generalizations of perceptual psychology should and do apply uniformly to Visua and Twin Visua. We can *supplement* the generalizations by specifying the environmental properties represented by Visua or Twin Visua. But the generalizations themselves ignore environmental *representata*. The generalizations constitute an 'abstract mathematical description' equally consistent with diverse distal interpretations (Egan, 2010, 256). Initially, Bayesian models may seem to offer precisely what Egan demands: 'abstract mathematical descriptions' that prescind from environmental *representata*. After all, Bayesian models emphasize pdfs, and a pdf is a purely mathematical entity: a function from real numbers to real numbers. Shouldn't we conclude that Bayesian models of depth perception describe Twin Visua just as well as Visua?

Any such conclusion would be mistaken. I concede that a Bayesian perceptual model has an abstract mathematical form. I concede that, in principle, this abstract form encompasses diverse chimerical creatures. Nevertheless, the model describes statistical

inferences over perceptual hypotheses, which it individuates partly through representational relations to specific environmental properties. Bayesian perceptual psychology does not pursue explanatory generalizations framed at an abstract mathematical level. Just as physics uses abstract mathematics to articulate generalizations over physical state-types, perceptual psychology uses abstract mathematics to articulate generalizations over representational mental state-types.

The central issue here is the notion of *random variable*. A random variable is a function from a sample space Ω to the real numbers ℝ. Thus, a random variable is defined only given a sample space. Ultimately, any Bayesian perceptual model featuring a random variable presupposes an appropriate sample space Ω. Perceptual models cite random variables only so as to illuminate probability assignments to environmental state estimates. The goal is to describe a statistical inference over *estimates about the perceiver's environment*. The random variable is a valuable device for describing this statistical inference. But it is simply a tool for formulating rigorous, elegant explanatory generalizations *concerning perceptual estimates*.

As evidence for my position, I cite *alternative measurement units*. Our mapping from depth-estimates to real numbers depends upon our choice of units. The metric system (p. 713) yields one random variable. The British imperial system yields another. Our choice of random variable reflects our measurement units. Thus, the specific mathematical parameters enshrined by a random variable are mere artefacts of our measurement system. The parameters lack any explanatory significance for scientific psychology. *We* may use metric units to measure depth, but *the perceptual system* almost certainly does not. Psychological significance resides in the state estimate, not the mathematical entities through which we parameterize state estimates. Our ultimate concern is the probability measure *over environmental state estimates*, not the probability distribution *over mathematical parameters*. To privilege the latter over the former is to read *our own* idiosyncratic measurement system into the psychological phenomena. We must not conflate our measurement units with the environmental states that we use the units to measure.

I conclude that Bayesian perceptual psychology offers intentional generalizations governing probability assignments to environmental state estimates. We articulate the generalizations by citing probability distributions and pdfs over mathematical entities. But these purely mathematical functions are artefacts of our measurement units. They reflect our idiosyncratic measurement conventions, not the underlying psychological reality. They do not yield any explanatorily significant level of non-representational psychological description. They are tools for describing how the perceptual system allocates probabilities over a hypothesis space whose elements are individuated representationally. A Bayesian perceptual model has an abstract mathematical form, but this form does not secure explanatorily significant non-representational descriptions of perceptual states.

What if we identify the privileged measurement units used by the perceptual system? Can't we assign explanatory priority to a pdf defined over *those* units? And won't the resulting theory be non-representational?

One problem with this suggestion is that the perceptual system may not employ measurement units. In Peacocke's (1992) terminology, perceptual representation may be 'unit-free'. As far as we know, for example, the visual system may form a depth-estimate without denominating that estimate in feet, metres, or any other measurement units (although *we* use units to describe the estimate's accuracy-condition). Admittedly, we may eventually discover that the perceptual system employs measurement units. It is difficult to anticipate how such a discovery might impact perceptual psychology. At present, the matter is speculative. All we can say for sure is that *current* Bayesian models do not attribute measurement units to the perceptual system. Current science posits probabilistic updating over perceptual hypotheses. It individuates the hypotheses partly through the specific environmental properties they represent.

9 Open Questions

Bayesian perceptual psychology raises numerous further questions, many on the border between philosophy and science. A few examples:

- What neural mechanisms implement, or approximately implement, the computations posited by Bayesian models?
- Does the Bayesian paradigm generalize from perception to cognition? (p. 714)
- Can Bayesian models illuminate the relation between normativity and intentionality?
- Can Bayesian models illuminate *what it is* to represent the external world?

Philosophers who pursue these questions will discover an imposing scientific literature that rewards intensive foundational analysis.[27]

References

Adams, W., Graf, E., and Ernst, M. (2004). 'Experience Can Change the "Light-From-Above" Prior'. *Nature Neuroscience*, 7, 1,057–1,058.

Beierholm, U., Quartz, S., and Shams, L. (2009). 'Bayesian Priors Are Encoded Independently from Likelihoods in Human Multisensory Perception'. *Journal of Vision, 9*, 1–9.

Bradley, P. (2008). 'Constancy, Categories, and Bayes'. *Philosophical Psychology, 21*, 601–627.

Brainard, D. (2009). ‘Bayesian Approaches to Color Vision’. In M. Gazzaniga (ed.), *The Visual Neurosciences*, 4th edn (pp. 395–408). Cambridge, MA: MIT Press.

Brewer, B. (2007). ‘How to Account for Illusion’. In F. Macpherson and A. Haddock (eds), *Disjunctivism: Perception, Action, Knowledge* (pp. 168–180). Oxford: Oxford University Press.

Burge, T. (2011). ‘Disjunctivism Again’. *Philosophical Explorations, 14*, 43–80.

Burge, T. (2005). ‘Disjunctivism and Perceptual Psychology’. *Philosophical Topics, 33*, 1–78.

Burge, T. (2010). *Origins of Objectivity*. Oxford: Oxford University Press.

Campbell, J. (2010). ‘Demonstrative Reference, the Relational View of Experience, and the Proximality Principle’. In R. Jeshion (ed.), *New Essays on Singular Thought* (pp. 193–212). Oxford: Oxford University Press.

Chalmers, D. (2006). ‘Perception and the Fall from Eden’. In T. Gendler and J. Hawthorne (eds), *Perceptual Experience* (pp. 49–125). Oxford: Oxford University Press.

Churchland, P. (1981). ‘Eliminative Materialism and the Propositional Attitudes’. *Journal of Philosophy, 78*, 67–90.

Clark, A. (2013). ‘Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science’. *Behavioral and Brain Sciences, 36*, 181–204.

Colombo, M. and Seriès, P. (2012). ‘Bayes in the Brain—On Bayesian Modeling in Neuroscience’. *British Journal for the Philosophy of Science, 63*, 697–723

Davies, M. (1992). ‘Perceptual Content and Local Supervenience’. *Proceedings of the Aristotelian Society, 92*, 21–45.

Dennett, D. (1987). *The Intentional Stance*. Cambridge, MA: MIT Press.

Egan, F. (2010). ‘Computational Models: A Modest Role for Content’. *Studies in History and Philosophy of Science, 41*, 253–259.

Egan, F. (1992). ‘Individualism, Computation, and Perceptual Content’. *Mind, 101*, 443–459.

Egan, F. (2009). ‘Vision’. In E. Craig (ed.), *Routledge Encyclopedia of Philosophy Online*. **<www.rep.routledge.com>**.

Ernst, M. (2010). ‘Eye Movements: Illusions in Slow Motion’. *Current Biology, 20*, R357–R359.

Field, H. (2001). *Truth and the Absence of Fact*. Oxford: Oxford University Press. (p. 715)

Fodor, J. (2008). *LOT 2*. Oxford: Oxford University Press.

Fodor, J. and Pylyshyn, Z. (1981). 'How Direct is Visual Perception? Some Reflections on Gibson's "Ecological Approach" '. *Cognition, 9*, 139–196.

Geisler, W. (2008). 'Visual Perception and the Statistical Properties of Natural Scenes'. *Annual Review of Psychology, 59*, 167–192.

Geisler, W. and Kersten, D. (2002). 'Illusions, Perception, and Bayes'. *Nature Neuroscience, 5*, 508–510.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.

Helmholtz, H. von. (1867). *Handbuch der Physiologischen Optik*. Leipzig: Voss.

Horgan, T. and J. Tienson (2002). 'The Intentionality of Phenomenology and the Phenomenology of Intentionality'. In D. Chalmers (ed.), *Philosophy of Mind: Classical and Contemporary Readings* (pp. 520–533). Oxford: Oxford University Press.

Jacobs, R. (1999). 'Optimal Integration of Texture and Motion Cues to Depth'. *Vision Research 39*, 3,621–3,629.

Kersten, D. and Mamassian, P. (2009). 'Ideal Observer Theory'. iIn L. Squire (ed.), *Encyclopedia of Neuroscience*, vol. 5 (pp. 89–95). Oxford: Oxford University Press.

Kersten, D. and Schrater, P. (2002). 'Pattern Inference Theory: A Probabilistic Approach to Vision'. In D. Heyer and R. Mausfeld (eds)*, Perception and the Physical World* (pp. 191–228). New York: Wiley and Sons.

Knill, D. (2007). 'Learning Bayesian Priors for Depth Perception'. *Journal of Vision, 7*, 1–20.

Knill, D. and Pouget, A. (2004). 'The Bayesian Brain: The Role of Uncertainty in Neural Coding and Computation'. *Trends in Neuroscience, 27*, 712–719.

Knill, D. and Richards, W. (eds) (1996). *Perception as Bayesian Inference*. Cambridge: Cambridge University Press.

Maloney, L. and Mamassian, P. (2009). 'Bayesian Decision Theory as a Model of Human Visual Perception: Testing Bayesian Transfer'. *Visual Neuroscience , 26*, 147–155.

Mamassian, P., Landy, M., and Maloney, L. (2002). 'Bayesian Modeling of Visual Perception'. In R. Rao, B. Olshausen, and M. Lewicki (eds), *Probabilistic Models of the Brain* (pp. 13–36). Cambridge, MA: MIT Press.

Marr, D. (1982). *Vision*. New York: Freeman.

Martin, M. (2004). 'The Limits of Self-Awareness'. *Philosophical Studies, 120*, 37–89.

Matthen, M. (2005). *Seeing, Doing, and Knowing*. Oxford: Clarendon Press.

McDowell, J. (2010). 'Tyler Burge on Disjunctivism'. *Philosophical Explorations, 13*, 243–255.

Peacocke, C. (1992). 'Scenarios, Concepts and Perception'. In T. Crane (ed.), *The Contents of Experience* (pp. 105–135). Cambridge: Cambridge University Press.

Quine, W. V. (1960). *Word and Object*. Cambridge, MA: MIT Press.

Rescorla, M. (2012). 'How to Integrate Representation into Computational Modeling, and Why We Should'. *Journal of Cognitive Science, 13*, 1–38.

Rock, I. (1983). *The Logic of Perception*. Cambridge, MA: MIT Press.

Segal, G. (1991). 'In Defense of a Reasonable Individualism'. *Mind, 100*, 485–494.

Seydell, A., Knill, D., and Trommershäuser, J. (2011). 'Prior and Learning in Cue Integration'. In J. Trommershäuser, K. Körding, and M. Landy (eds), *Sensory Cue Integration* (pp. 155–172). Oxford: Oxford University Press.

Siegel, S. (2011). 'The Contents of Perception'. In E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. **<plato.stanford.edu/archives/spr2011/entries/perception-contents>**.

Silverberg, A. (2006). 'Chomsky and Egan on Computational Theories of Vision'. *Minds and Machines, 16*, 495–524. (p. 716)

Stich, S. (1983). *From Folk Psychology to Cognitive Science*. Cambridge, MA: MIT Press.

Stone, J. (2011). 'Footprints Sticking Out of the Sand, Part 2: Children's Bayesian Priors for Shape and Lighting Direction'. *Perception, 40*, 175–190.

Thompson, B. (2010). 'The Spatial Content of Experience'. *Philosophy and Phenomenological Research, 81*, 146–184.

Travis, C. (2004). 'The Silence of the Senses'. *Mind, 113*, 57–94.

Trommershäuser, J., Körding, K., and Landy, M. (eds) (2011). *Sensory Cue Integration*. Oxford: Oxford University Press.

Weiss, Y., Simoncelli, E., and Adelson, E. (2002). 'Motion Illusions as Optimal Percepts'. *Nature Neuroscience, 5*, 598–604.

Wolpert, D. (2007). 'Probabilistic Models in Human Sensorimotor Control'. *Human Movement Science, 26*, 511–524.

Notes:

([1]) Bayesian perceptual psychology generalizes *signal detection theory*, which was developed in the 1950s. For comparison of the two frameworks, see Kersten and Schrater (2002, 193–199).

([2]) On the distinction between the perceiver and her perceptual system, see Burge (2010, 23–24; 2011, 68–69).

([3]) When the hypothesis space is continuous, $p(h)$ is a *probability density function*. See below for details. For ease of exposition, I often blur the distinction between probability and probability density.

([4]) There is an unfortunate tendency among scientists and even some philosophers to conflate Bayes's *Theorem* and Bayes's *Rule*. The former is an easily provable mathematical theorem. The latter is a *prescriptive norm* that dictates how to reallocate probabilities in light of new evidence.

([5]) The models described in this paragraph assume *diffusely illuminated flat matte surfaces*. To handle other viewing conditions, we must replace $R(\lambda)$ with a more complicated surface reflectance property, such as a *bidirectional reflectance distribution*.

([6]) Current models describe perception of *surface* colour. As Matthen (2005, 176) emphasizes, colour perception also responds to transmitted colour (e.g. stained-glass windows) and coloured light sources. Thus, we should not *identify* colours with surface reflectance properties. Should we identify colours with other, possibly disjunctive, physical properties? Maybe. But the Bayesian models I am describing do not presuppose a physicalist reduction of colour. One might combine those models with various metaphysical views of colour, such as that colours are dispositions to cause sensations in normal human perceivers, or such as Matthen's (2005) *pluralistic realism*. Current Bayesian models assume no particular metaphysics of colour. They simply assume that human surface colour perception involves estimation of surface reflectance, as informed by an estimate of background illumination.

([7]) Cf. Burge's 'Proximality Principle' (2005).

([8]) Motor efference copy figures most prominently in Bayesian models of *sensorimotor control* (Wolpert, 2007).

([9]) There are exceptions, such as Knill (2007).

([10]) In some cases, the priors reflect non-obvious statistical regularities about the environment (Geisler, 2008). In other cases, a satisfying explanation awaits discovery. An example: somewhat mysteriously, the perceptual system assumes that the light source is located overhead *and slightly to the left* (Mamassian, Landy, and Maloney, 2002). One question in this area concerns *informational encapsulation*: to what extent can cognition influence the priors?

([11]) *Cue combination* provides another good illustration. The perceptual system typically receives multiple cues, often through different sensory modalities, regarding a single environmental variable. Bayesian perceptual psychology offers a unified framework for explaining diverse cases of intermodal and intramodal sensory fusion: visual and auditory cues to location; visual and proprioceptive cues to limb position; conflicting visual cues to depth; and so on. See Trommershäuser, Körding, and Landy (2011) for an overview.

([12]) More technically: the prior likelihood $p(e \mid h)$, considered as a function of h for fixed e, has higher variance when the retinal image e has lower contrast.

([13]) Burge discusses several Bayesian perceptual models, but he does not discuss their specifically Bayesian features. Bradley (2008) defends representationalism by citing Bayesian models of colour perception.

([14]) For a survey of philosophical approaches to perceptual content, see Siegel (2011).

([15]) There are additional phenomena in a similar vein that favour realism towards prior probabilities and likelihoods (Seydell, Knill, and Trommershäuser, 2011; Beierholm, Quartz, and Shams, 2009). Realism towards the utility function seems well-supported for Bayesian models of bodily motion (Maloney and Mamassian, 2009). I am less sure about the utility functions that figure in Bayesian models of perception. Moderate instrumentalism may be more promising for that case.

([16]) Clark (2013) raises the same worry.

([17]) As noted above, one might hold that the final percept represents colour but not reflectance. However, this suggestion provides no support for phenomenal content. If one perceives a surface as a specific colour, then one's percept is veridical only if the surface has that colour. Since Invert's percept is veridical, and since the perceived surface is green, Invert does not perceive the surface as red. So Nonvert perceives a surface as red, while Invert does not perceive a surface as red. There is no basis here for type-identifying the relevant percepts.

([18]) One can individuate perceptual states through the environmental *properties* they represent without individuating them through the environmental *particulars* they represent. Burge (2010) introduces an individuative scheme for perceptual content along these lines. To illustrate, suppose that a percept attributes convexity to object O. According to Burge, any percept expressing the same content must also represent convexity. But a percept might express that same content while attributing convexity to a distinct object O^*. Or a percept expressing that same content may involve a referential illusion, in which case it does not successfully attribute convexity to any object.

([19]) Silverberg (2006) argues that Egan misinterprets Marr. Egan (2009) discusses Bayesian models of perception but does not discuss how they bear upon her views regarding non-intentional computational modelling.

([20]) Rescorla (2012) relates these points to the computational models employed within CS and AI.

([21]) For discussion of possible neural mechanisms, see Clark (2013) and Knill and Pouget (2004).

([22]) The details of Field's discussion raise further doubts about the conjecture. He claims that there is no viable interpersonal notion of type-identity for mental representation tokens (2001, 75, fn. 3). In other words, Field's favoured taxonomic scheme cannot type-identify the mental states of distinct creatures. This result is incompatible with current perceptual psychology, which routinely type-identifies the perceptual states of distinct creatures. How could any serious science of perception do otherwise?

([23]) Not everyone accepts that there exist creatures Visua and Twin-Visua satisfying these assumptions. In particular, Segal (1991) denies that perceptual states of neurophysiological twins can represent different environmental properties. For the sake of argument, I grant Egan's description of the thought experiment.

([24]) A function X: $\Omega \rightarrow \mathbb{R}$ is *measurable* just in case, for every Borel set $B \subseteq \mathbb{R}$, $X^{-1}(B)$ belongs to the σ-algebra. One can generalize the definition of *random variable* to include functions from Ω to mathematical structures besides the real numbers. For ease of exposition, I focus on real-valued random variables. Consideration of generalized random variables would not alter my main conclusions.

([25]) Let P be a probability measure, let X: $\Omega \rightarrow \mathbb{R}$ be a random variable, and let $B \subseteq \mathbb{R}$ be a Borel set. Then we define a probability distribution P_X by $P_X(B) = P(X^{-1}(B))$.

([26]) If P is a probability distribution, and if $\rho(x)$ is an associated pdf, then $P([a, b]) = \int_a^b \rho(x)dx$.

([27]) I am indebted to Mohan Matthen and Susanna Siegel for comments that vastly improved this entry. I have also benefited from discussion of these issues with Tyler Burge, John Campbell, Kevin Falvey, Ian Nance, Christopher Peacocke, and Tamar Weber.

Michael Rescorla

Michael Rescorla, University of California, Santa Barbara