

# Innateness and (Bayesian) Visual Perception

## *Reconciling Nativism and Development*

### 1 A Research Strategy

Because innateness is such a complex and controversial issue when applied to higher level cognition, it can be useful to explore how nature and nurture interact in simpler, less controversial contexts. One such context is the study of certain aspects of visual perception—where especially rigorous models are possible, and where it is less controversial to claim that certain aspects of the visual system are in part innately specified. The hope is that scrutiny of these simpler contexts might yield lessons that can then be applied to debates about the possible innateness of other aspects of the mind. This chapter will explore a particular way in which visual processing may involve innate constraints and will attempt to show how such processing overcomes one enduring challenge to nativism. In particular, many challenges to nativist theories in other areas of cognitive psychology (e.g., “theory of mind,” infant cognition) have focused on the later development of such abilities, and have argued that such development is in conflict with innate origins (since those origins would have to be somehow changed or overwritten). Innateness, in these contexts, is seen as antidevelopmental, associated instead with static processes and principles. In contrast, certain perceptual models demonstrate how the very same mental processes can both be innately specified and yet develop richly in response to experience with the environment. In fact, this process is entirely unmysterious, as is made clear in certain formal theories of visual perception, including those that appeal to spontaneous endogenous stimulation, and those based on Bayesian inference.

For helpful conversation and comments on earlier drafts, I thank Paul Bloom, Peter Carruthers, Erik Cheries, Frank Keil, Koleen McCrink, and Steve Mitroff. The preparation of this chapter was supported by National Institute of Mental Health grant number R03-MH63808-01 and National Science Foundation grant number BCS-0132444.

### 1.1 *Innateness in Cognitive Science*

One of the most persistent and important themes in cognitive science is the issue of whether and how various cognitive mechanisms, processes, abilities, and concepts may in some sense be innate. This debate is far older than cognitive science or even the modern incarnation of psychology (see Diamond, 1974; Samet, 1999), and (as the other essays in this volume attest) it remains a lively research topic today. In a sense, the continuing debate about innateness has been incredibly unifying, if only because it draws together researchers from so many different disciplines, who study so many distinct aspects of the mind. (An online search of the *MIT Encyclopedia of Cognitive Science* turns up discussions of innateness in almost every imaginable corner, in fields ranging from psychology and linguistics to ethology and neuroscience, and in specific topics ranging from imitation and ethics to numeracy and phantom limbs.)

Of course, questions of innateness have also long been among the most divisive and controversial issues in cognitive science. Indeed, one can hardly mention the yin of nativism without the yang of empiricism, and the combative tension between these two traditions is alive and well in all of the aforementioned areas of cognitive science. The suggestion of a possible nativist origin for part of almost any cognitive process always seems to evoke an academic quarrel, and such debates have sometimes polarized entire fields. Such debates will often even spill over into the popular press and the general public—especially when questions of innateness are raised about issues such as the nature of our emotions, or our mate-choice preferences (see Pinker, 2002).

### 1.2 “Barely Worth Mentioning”

Because issues surrounding innateness are liable to become so complex and controversial when applied to human cognition, some researchers have suggested that a useful strategy is to study how such issues play out in simpler and possibly less controversial contexts. Gallistel (2000), for instance, has attempted to study nativism (along with other issues such as computation and modularity) in circumscribed domains encountered by simpler organisms. As a case study, he has studied the nature of certain forms of wayfinding in insects. Bees, for example, are able to communicate the direction and distance of food to their hive mates, using the location of the sun in the sky as a reference angle. Their ability to do this year-round and even on cloudy days, moreover, indicates that they represent the solar ephemeris function: they know where the sun is in the sky as a function of the time of day and the day of the year—a function that varies depending on where the bees live. Studies of this process have gradually revealed a detailed picture of how innate and learned components contribute to this ability (e.g., Dyer & Dickinson, 1996). Even without ever experiencing the sun, bees still in some sense assume that it exists, and that it rises in the east and sets in the west—the one thing that is true of the ephemeris function anytime and anywhere—and they initially assume that it traverses this path with a discrete jump around noon. The shape and timing of this path are then gradually tuned by early visual experiences, in ways that are beginning to be

understood, to match the actual ephemeris function for the sun at that location (Dyer & Dickinson, 1994). In few areas of human cognition are we able to obtain such a detailed and rigorous picture of how nature and nurture interact.

Gallistel (2000) stresses several specific lessons of this research for cognitive science, but perhaps the most important lesson is methodological: It is possible to gain insight into the ways nature and nurture cooperate by studying simpler cases that are more amenable to rigorous study and critical experiments, and at the same time are far less controversial. Indeed, Gallistel has stressed that claims of innate specialized structure are “barely worth mentioning” in certain parts of biology, in that they are assumed to be the norm. In this way, then, it is possible to study innateness in simpler contexts than higher level human cognition, and perhaps learn some important things about how nature and nurture interact in situations that are divorced from the abstractness, concern, and controversy that normally attends such projects. This chapter explores a similar strategy, but with a focus on what are perhaps simpler cognitive processes rather than simpler organisms. In particular, the aim of this chapter is to make a few specific points about how nature and nurture can interact in the context of certain theories of visual perception.

The study of vision is certainly one of the most successful projects of cognitive science: it has arguably enjoyed the most rigorous theories, the most developed computer models, and perhaps the tightest coupling so far of psychology and neuroscience. Some have argued that this is no accident. Neuroscientists have stressed that more than half of cortex is devoted to vision (at least in monkeys), and Fodor (1983) has famously argued that the relatively modular and encapsulated structure of “input systems” such as vision is bound to translate into greater empirical tractability. Innateness has certainly been a persistent theme in the study of visual perception, and in parts of vision research (e.g., face perception) it is just as controversial as in the study of higher level cognition (e.g., Gauthier & Nelson, 2001; Kanwisher & Moscovitch, 2000). However, whereas many scientists are hesitant to grant the existence of significant innate components for large swaths of human cognition, it is essentially uncontroversial among vision scientists that at least some portions of (early) visual perception are (in part) innately specified.<sup>1</sup>

The hope is that by studying the interaction of nature and nurture in visual perception, we might gain some useful insights that can then be applied to debates about nativism in other areas of cognitive science. This chapter is one small part of this project, and it attempts to dissolve apparent conflicts between innate structure and later development by focusing on how the very same visual processes can both

1. For the rest of this chapter, I will often drop caveats such as “in part.” I take it as given that no mental process—perceptual or otherwise—is *entirely* innately specified, any more than any such process is entirely learned. The question of innateness is taken here to be whether any significant parts of the mind can develop without bona fide *learning* (see Pylyshyn, 1985, for careful definition and discussion of this distinction), taking for granted (1) that other mental abilities are largely learned via experience with the environment, and (2) that environmental interaction is also always likely to play a critical role in revealing innate structure.

be innately structured and yet themselves develop richly in response to experience with the environment.

## 2 Is Nativism Antidevelopmental?

Most researchers today recognize that the mind must develop as a result of some mixture of innateness and learning, but it is not always clear how this cooperation is to be forged. While some writers emphasize that even innate structure itself is designed to learn (e.g., Marler, 1991; Pinker, 1997a), others seem to assume that there is some inherent tension between the two: some processes may be (mostly) innate and others may be (mostly) learned, but there is a hesitancy to explore processes that are both innately determined and then themselves continue to develop via learning. Some of this tension is only sociological, in that nativist research programs are seen to somehow squelch the study of development: “Calling some skill or behavior innate tends to stop analysis of how it develops” (Fischer & Stewart, 1999, p. 150). In many contexts, however, it seems popular to assume that there is a deeper and more inherent conflict between nativism and development (e.g., Elman et al., 1996; Fischer & Bidell, 1991; Karmiloff-Smith, 1992; Quartz & Sejnowski, 1997; Thelen & Smith, 1994; for discussions of neuroscientific versions of such arguments, see Marcus, chapter 2 here). According to these views, innateness is seen as inherently “anti-developmental” (Gopnik, 1996, p. 174), and at root nativist research programs are seen as attempts at “minimizing change during development” (Quartz & Sejnowski, 1997, p. 537). Recent examples of this conflict are readily seen in two of the most active areas of cognitive developmental psychology: “initial knowledge” in infancy and “theory of mind.”

### 2.1 *Examples from “Object Cognition” in Infancy*

One area of active nativist debate in cognitive development is the study of what young infants know about the physical world. Using looking-time measures to study the infant’s object concept, developmental psychologists have demonstrated that infants even a few months old have a substantial amount of “initial knowledge” about objects, in domains such as physics and arithmetic (for reviews see Baillargeon, 2002; Spelke et al., 1995a; Wynn, 1998). This research has shown, for example, that infants have some appreciation of the fact that objects must trace spatiotemporally continuous paths through space (Spelke et al., 1995a); that objects will fall if unsupported (Needham & Baillargeon, 1993); that one plus one yields two, and other simple arithmetical facts (Feigenson et al., 2002; Wynn, 1992); that objects cannot pass through one another (Baillargeon et al., 1985; Spelke et al., 1992); that the mechanical interactions of objects will obey certain causal laws (Leslie & Keeble, 1987); and so on.

Some of the theorizing that has accompanied this research has had a strong nativist flavor, often phrased in terms of “core knowledge” (e.g., Spelke, 2000, 2003; Shusterman & Spelke, chapter 6 here). Under these proposals, the appreciation of physical and numerical laws revealed in looking-time experiments reflects the operation of innate principles of some form—albeit principles built into highly

task- and domain-specific processes that are largely independent and encapsulated from each other and from other aspects of the mind.<sup>2</sup> These principles form the core of our knowledge of the world and provide the critical representations needed for bootstrapping by later learning. Many types of evidence have been adduced to support nativist theories of such abilities, including familiar arguments based on precocity, learnability, universality, and continuity with other branches of the phylogenetic tree (e.g., Spelke, 1988, 1998).

These nativist proposals have proven highly controversial. Indeed, debates that center on nativism have been featured prominently every few years at major infant cognition conferences and have resulted in a number of recent vigorous exchanges in the literature (e.g., Cohen & Marks, 2002, v. Wynn, 2002; Haith, 1998, v. Spelke, 1998; Smith, 1999, v. Baillargeon, 1999). The proponents of empiricist viewpoints have been just as lively as their nativist counterparts—suggesting for example that “claims that types of knowledge are innate [constitute] misdemeanors, if not outright psychological felonies” (Haith, 1998, p. 168)—and the arguments used to fuel such views have often made strong appeals to the perceived inability of nativist theories to accommodate later development (e.g., Bogartz et al., 1997; Cohen & Cashon, 2003; Fischer & Biddell, 1991; Johnson, 2003; Smith, 1999). Nativist theories are seen to “shut down attempts at process explanations and developmental analyses” and to be inherently “nondevelopmental” (Haith, 1998, pp. 176, 172). The biggest problem for nativist views is alleged to be simply that “development *happens*” (Johnson, 2003, p. 103). Theories of causality, for example, that rely on “an innate module” are taken to imply that “there is no room . . . for causality perception to develop in stages. . . . Causality, according to Leslie and other nativists . . . is not something to study developmentally: it is a nonsequitur to study the development of something that doesn’t change” (Cohen et al., 1998, pp. 172–3). And some writers have even argued explicitly that our eventual ability to overrule the dictates of innate principles (e.g., to understand Star Trek transporter beams, which violate laws of spatiotemporal continuity) falsify nativist claims based on “core knowledge” (Gopnik & Meltzoff, 1997).

## 2.2 *Examples from “Theory of Mind”*

Strikingly similar arguments are found in other areas of developmental research, such as “theory of mind.” Many human behaviors are the result of internal mental states such as *beliefs* and *desires*. Such representations not only cause our behaviors but also form much of the currency of our mental lives: even young children perceive, interpret, predict, and explain the behavior of others in terms of their underlying mental states. The acquisition of such abilities—collectively referred to as a “theory of mind” (ToM)—is early, universal (except in certain clinical populations), seemingly effortless, and largely dissociable from more general intellectual

2. Even among theorists who are sympathetic to aspects of nativism, there is still much debate as to the exact format of these principles, and the degree to which they represent cognitive as opposed to perceptual (or attentional?) processing (e.g., Carey & Xu, 2001; Scholl & Leslie, 1999a).

development. In adults, the exercise of such abilities is often irresistible and seemingly instantaneous. Such facts—along with worries about learnability for concepts such as *belief*, whose referents cannot be directly seen, heard, or felt—have helped to inspire theories that take the core of ToM (though obviously not its mature competence) to be innate (e.g., Baron-Cohen, 1995; Leslie, 1994; Scholl & Leslie, 1999b, 2001)—“part of our genetic endowment triggered by appropriate environmental factors, much as, say, puberty is” (Scholl & Leslie, 2001, p. 697).

Again, nearly all of the arguments against such views tend to turn on the role of development (see Scholl & Leslie, 1999b, for extensive discussion). Several authors have said explicitly that certain nativist views are “nondevelopmental” or “anti-developmental” and that only developmental evidence can decide whether certain nativist views of ToM are correct (e.g., Gopnik & Meltzoff, 1997; Gopnik & Wellman, 1994). Others argue that the mere fact of observed development in ToM—paradigmatically, the fact that children become able to pass the “false belief task” around the age of four—argues against theories based on “modularity nativism” (Wellman et al., 2001).<sup>3</sup> Theories of ToM based on an innate modular core are seen as incapable of accounting for change, without appeal to “the maturation of another innate structure, a later module coming on line . . . . It is . . . difficult, however, to see why evolution would have designed a sequence of incorrect modules, each maturing only to be replaced by another” (Gopnik & Meltzoff, 1997, pp. 54–5). This type of ridicule seems rooted in the idea that an encapsulated process with an innate basis is somehow unable to develop via contact with the environment. Even authors who are sympathetic to nativist and modular views of ToM seem uncomfortable with the idea that the innate processes can themselves change and develop. Rather, such theorists appeal to notions, such as parameterization, derived from studies of linguistic nativism (e.g., Segal, 1996; Stich & Nichols, 1998). On such views, development is explained not by robust learning via interaction with the environment but by an executive “switching” mechanism that simply chooses at various times among preset options, all of which are still innately determined. Thus even nativist theorists are drawn to views that seem implausible (see Scholl & Leslie, 1999b, for discussion), on the basis of developmentally motivated concerns.

3. As with ‘core knowledge’ views of infants’ object-cognition abilities, nativist theories of ToM have also often claimed that the innate processes are highly domain specific, informationally encapsulated, and possibly embodied in cognitive modules (e.g., Scholl & Leslie, 1999b, 2001). In general, of course, innateness and modularity are entirely dissociable concepts: “The claim of innateness is obviously not required of the modularity view” (Scholl & Leslie, 1999b, p. 134). However, we have sometimes treated innateness as a critical part of how modularity applies to the case of ToM, as have others who refer to “modularity nativism.” This has sometimes lead to confusion (see Nichols & Stich, 2003), but it seems clear that most developmentally motivated arguments against the ‘modularity nativism’ view of ToM are in fact aimed at a certain conception of nativism, rather than anything intrinsic to modularity. For example, Gopnik and Meltzoff (1997) take themselves to be arguing against the following: “According to modularity theories, representations of the world are not constructed from evidence in the course of development. Instead, representations are produced by innate structures, modules, or constraints that have been constructed in the course of evolution” (p. 50).

### 2.3 *Nativism and Flexibility*

In the foregoing examples, arguments against innateness have been fueled by concerns that nativist origins preclude developmental flexibility. This view is not entirely unmotivated, of course. Such views are suggested by characterizations of the innate endowment as a set of *constraints* on development (e.g., Elman et al., 1996; Keil, 1991), a view that even some resolute nativist researchers take to be unfortunate: such constraint-based characterizations “imply that innate knowledge prevents people from learning” (Spelke, 1998, p. 194). Perhaps the strongest motivation for such views, however, is the simple fact that nativist theories rarely contain specific proposals for *how* later development would work, beyond peripheral ideas of triggering and maturation. Theories of core knowledge in infant object-cognition, for example, have been vague about just how later bootstrapping processes would harness the innate core representations, or just how it is possible for “hardwired” principles to later be unwired or rewired. Rather, such theories have appealed to “overriding” to explain radical change (see Stich & Nichols, 1998), wherein entirely new systems (perhaps with little or no innate basis) can eventually come online and overrule core knowledge. This issue also arises in the theory-of-mind literature, in suggestions that the innate contribution affects only “early ToM” (by analogy to “early vision”; e.g., Scholl & Leslie, 2001)—the implication being that development results from higher level processes that are not claimed to be innate. In sum, all of these views suggest that the actual innate processes themselves do not develop; rather, development occurs in a context beyond the innate content, which is still seen as static and unchanging.

## 3 Bayesian Visual Perception

The previous section sketched only a few examples a wider perceived conflict between innateness and development. But is such a conflict really necessary? Some theories of visual perception, at least, suggest a negative answer: it may be possible for the *very same process* to be both innately determined and yet to later change (even radically) in response to interaction with the environment. This section sketches out one general framework, based on Bayesian inference, in which this is possible. After this brief introduction to Bayesian perception, section 4 explores how both Bayesian models and models based on spontaneous endogenous stimulation can incorporate both nativism and learning in the same processes.

### 3.1 *The Impossibility of Visual Perception*

Visual perception is the process of recovering useful information about the structure of the world, based the shifting patterns of light that enter the eyes. Perhaps the most fundamental fact about visual perception is that this task is, strictly speaking, impossible. That is, the shifting patterns of light that enter the eyes are insufficient by themselves to fix the structure of the external world from which that light was reflected or emitted (Marr, 1982) because there are always a multitude of possible structures in the world that could have given rise to those same patterns of

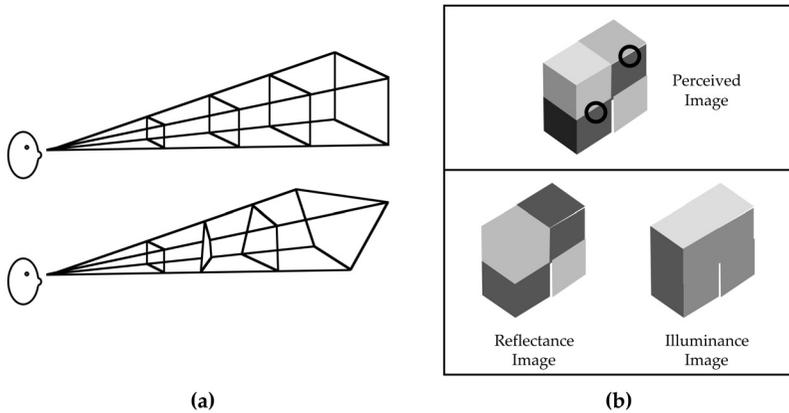


FIGURE 3.1 Two examples of underdetermination problems in visual perception. (a) A given 2D retinal shape could be caused by a real 3D object in the world of almost any size and shape (adapted from Feldman, 1999). (b) The light reaching the eyes is a product of the illumination from a light source and the reflectance of the ‘paint’, and it isn’t strictly possible to recover only one of these operands (the reflectance, which we would like to see as constant regardless of the lighting) without ‘unmultiplying’. Note that the two circled patches contain the same objective luminance edge, despite the fact that we see one caused by lighting and one caused by ‘paint’ (adapted from Adelson, 2000).

light. In this sense the visual system must solve an “inverse problem,” which is technically not possible via deductive inference.

This underdetermination is most commonly appreciated in the case of depth and three-dimensional shape. A given patch of retinal stimulation, for example, could correspond to an object in the world of almost any size (since a small nearby object will create the same retinal image as a larger object further away) and almost any shape (see fig. 3.1a). Furthermore, objects in the world could in principle be changing constantly among these possibilities yet always continue to project the same retinal stimulation. Such dilemmas of underdetermination are in no way specific to depth perception but hold for almost every aspect of visual processing. Another example is the perception of surface lightness, where in everyday experience we commonly perceive the reflectance of a surface as constant in the face of changing illumination, despite the fact that the same luminance edge can be produced by either a change in lighting or a change in reflectance (fig. 3.1b). This is also technically impossible to achieve, since the actual number of photons that reach the eyes after reflecting off a surface will always be the product of the illumination from the light source (which we want to discount) and the reflectance of the surface (the “paint,” which we want to know about). Separating these sources is not strictly possible, since it would essentially require the visual system to “unmultiply,” for example, solving for  $R$  in the equation  $R \times L = 12$  (see Adelson, 2000).

In each of these cases (and many others), accurate perception based only on the incoming visual information is impossible, due to an underdetermination problem. As a result, successful perception is possible only via the application of internal

processing constraints, and the focus of much research in perception has been to determine the nature of these constraints (Marr, 1982; Rock, 1983). So what is their nature? At first blush, it seems unlikely that there would be any global answer to this question, simply because perception is generally thought to consist of many specialized and independent subprocesses (Marr, 1982; Palmer, 1999). For example, the processes that compute depth from disparity information are thought to be functionally (and sometimes neuroanatomically) distinct from those that determine phenomenal colors based on distributions of different cone outputs, those that compute the correspondences between items undergoing apparent motion, or those responsible for recognizing faces. The outputs of such processes may in some cases serve as inputs to others, and they may all eventually feed into more central mechanisms, but in general the processes that subservise each such individual task are likely to be specialized and functionally independent—making it unlikely on the face of it that there would be any single “nature” of perception. Despite this degree of specialization in the visual system, however, researchers have identified a few general principles that seem to run through all manners of visual processing. Perhaps the most powerful such principle is what might be called “coincidence avoidance.”

### 3.2 *Coincidence Avoidance and Perception as Unconscious Inference*

The visual system, it appears, abhors a coincidence. One of the earliest and most forceful proponents of this view was Helmholtz, who popularized the idea in his principle of “unconscious inference” (1867/1925):<sup>4</sup>

#### Helmholtz’s Principle of “Unconscious Inference”

What is perceived are essentially those objects and events that under normal conditions *would be most likely* to produce the received sensory stimulation.

The idea isn’t that the visual system makes actual inferences by reasoning just as we do but rather that it operates *in accord with* such a principle, since doing so has proven selectively advantageous in the course of the visual system’s phylogenetic development. This type of principle proves extremely adept at accounting for a wide variety of perceptual phenomena.

One can readily appreciate the operation of such a principle by considering some of the simple shapes in figure 3.2. You see figure 3.2a as a straight line rather than a line that curves in the plane parallel to your line of sight (which would project the same retinal stimulation). Why not perceive this latter possibility? The visual system discounts the curved-line interpretation because, while possible, it would require an “accidental” viewpoint: of all possible viewpoints of a

4. Recent scholarship in the history of vision science has uncovered examples of this principle—and of an astounding number of other supposedly recent insights—in the work of the Islamic scholar Al-hazen (965–1039). For fascinating discussion, see Howard (1996) and Sabra (1978).

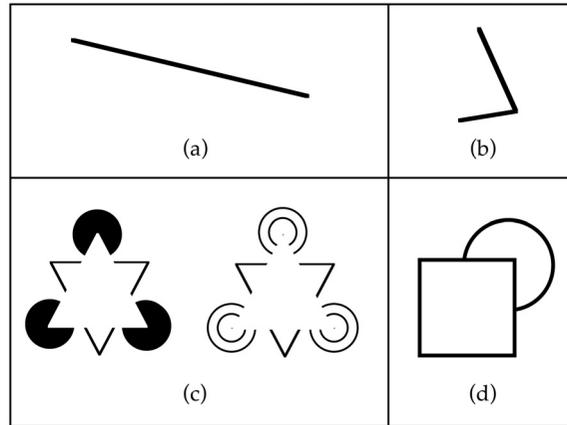


FIGURE 3.2 Four simple examples of the principle of unconscious inference in action. See text details.

curved three-dimensional wire, only a small number would produce a straight-line two-dimensional percept. (This is where the notion of likelihood enters Helmholtz's principle: if you chose a viewpoint randomly, it would be extremely *unlikely* to produce a single linear two-dimensional percept of a curved three-dimensional wire.) As such, the visual system assumes that a straight contour in an image corresponds to a straight contour in the world. Similarly, you see an "L" shape in figure 3.2b (rather than two independent segments which lie at different depths) because it would be unlikely for two independent contours to coterminate in the image if they weren't really connected somehow (e.g., if you randomly threw down matchsticks onto the floor, very few would end up aligned in this way). Slightly more complicated cases of illusory contours (fig. 3.2c) and amodal completion (fig. 3.2d) can also be interpreted in this way: you see illusory triangles in figure 3.2c because the other obvious possibility (involving a lucky perfect configuration of independent black shapes and contours) would be a coincidence and is judged to be unlikely; you see a partly occluded circle in figure 3.2d because the other obvious possibility (a pac-man perfectly abutting a square) would also require a coincidental and unlikely arrangement. Beyond these simple examples, similar judgments of probability and coincidence can explain many (or even most) other well-known aspects of visual processing (for extended discussion, see Hoffman, 1998; for specific discussions of "generic viewpoint" assumptions in vision, see Biederman, 1987; Freeman, 1996).

### 3.3 Visual Perception as Bayesian Inference

The idea of vision as unconscious inference has been adopted in recent years by more formal probabilistic theories of visual perception (Rao et al., 2002), especially models based on Bayesian inference. Bayesian inference is a method of optimal reasoning under uncertainty, and specifies how to choose rationally from among a set of mutually exclusive hypotheses (Hs) based on a given pattern of data (D).

Bayes' theorem requires you to first specify the "likelihood function," which models the probability of obtaining the observed data if you assume each hypothesis to be true—for each hypothesis  $H$ ,  $p(D|H)$ , "the probability of the data given the hypothesis." This distribution (one resulting value for each  $H$ ) is then convolved with a second distribution that models the probability of each hypothesis independent of the current data— $p(H)$ , "the probability of the hypothesis." This is called the Bayesian "prior" and models the degree of "prior belief" in the conclusion. The product of the prior and the likelihood function is then divided by the probability of the data irrespective of the hypothesis, and this quotient constitutes the "posterior distribution," which specifies the relative degree of resulting belief for each hypothesis:

$$p(H|D) = \frac{p(D|H)p(H)}{p(D)}$$

In the context of Bayesian perception, however, the probability distribution of the data can be safely ignored as a normalizing constant, leaving us with:

$$p(H|D) = p(D|H)p(H)$$

Bayes' theorem simply tells you to choose the hypothesis that maximizes this value.<sup>5</sup>

Bayes' theorem constitutes optimal reasoning under uncertainty, but people do not always find such reasoning to be intuitive. Thus a common result in the psychology of decision-making has been to find that people ignore the prior distribution—even in situations such as medical diagnosis, where such errors can be disastrous (e.g., Kahneman & Tversky, 1972; for a review see Koehler, 1996). As a result, a common applied project in this literature has been to develop tools that teach laypeople to reason according to Bayes' theorem (e.g., Sedlmeier & Gigerenzer, 2001). However, despite the fact that people do not find Bayesian reasoning to be intuitive in conscious decision-making, a wealth of recent evidence suggests that the visual system does engage in "unconscious inference" in accord with Bayes' theorem (for reviews see Kersten et al., 2004; Kersten & Yuille, 2003; Knill & Richards, 1996; Mamassian et al., 2002).<sup>6</sup>

In the context of visual perception, the data consist of the visual image that arrives at the retina, and the hypotheses under consideration are the possible scenes

5. Or, if the cost of making certain kinds of errors is higher than others, the resulting posterior distribution might be convolved again with a cost function, and the maximum value in the resulting distribution may not always correspond to the maximum of the posterior distribution itself. More generally, one can also analyze the shape of the posterior distribution to analyze the reliability of the information: sharply peaked distributions indicate high reliability, while broad gradual distributions reflect relatively lower reliability.

6. In fact, some aspects of higher level cognition also seem to implicate implicit Bayesian processing—including categorization (e.g., Tenenbaum, 1999), causal reasoning (e.g., Tenenbaum & Griffiths, 2001), and word-learning (e.g., Tenenbaum & Xu, 2000)—but such ideas have been most successful and popular in vision science, to the extent that it is perhaps not hyperbole to claim that "Bayesian concepts are transforming perception research" (Geisler & Kersten, 2002, p. 508).

in the world that may have given rise to the image. Because, as noted above, this problem is underdetermined in many ways, the visual system must engage in probabilistic decision-making to choose which scene to assign as the conscious percept corresponding to the incoming image—and in many cases it does this by maximizing the posterior distribution in accord with Bayes’s theorem:

$$p(\text{Scene} | \text{Image}) = p(\text{Image} | \text{Scene}) p(\text{Scene})$$

In this context, ideas about “unconscious inference” in visual perception can be rigorously modeled and psychophysically tested. The likelihood function— $p(\text{Image} | \text{Scene})$ —models aspects of optics and projection that have been increasingly well understood (especially in the context of computer graphics), and the prior— $p(\text{Scene})$ —models the prior assumptions (sometimes called “natural constraints”) that the visual system has about the structure of the world, and that are necessary in order to cope with underdetermination. When such models are constructed and tested against psychophysical data, the fit is often extremely close—suggesting (in the absence of other models that would yield similar predictions) that the visual system is actually reasoning in accord with Bayesian inference.<sup>7</sup>

One example of an assumption about the world that has been modeled in this way is the assumption that there is a single light source that comes from overhead (Rittenhouse, 1786). This is a particularly useful case, given that this “natural constraint” has typically been phrased as a vague proposition (just like many principles in infant cognition and theory of mind), yet it turns out to be explicable in Bayesian terms—and moreover, it turns out to be plausibly innate. Because of underdetermination problems, the visual system must use several heuristic cues to three-dimensional shape, and one such cue consists of luminance gradients on surfaces. In certain contexts, the visual system assumes that top-to-bottom lightness gradients in an image that are lighter at the top and darker at the bottom signal the presence of a convex surface—a “bump.” In contrast, a top-to-bottom image gradient that is darker at the top and lighter near the bottom is seen to signal the presence of a concave surface—a “dent.” As with most such notions, the results are easier to see than to read, and as such most observers will readily see the lower-middle disc in the “egg carton” in figure 3.3 as a concave dent, and the others as convex bumps (Ramachandran, 1988). In contrast, if you turn the page upside down, you’ll see the opposite the pattern. Note that this phenomenon does not depend on artificial images: you can observe the same phenomenon in the real world, and in actual photographs such as the moon craters in figure 3.4. This phenomenon makes some

7. The assumed absence of other models that would yield similar good fits is critical, though a detailed discussion of this issue is beyond the scope of this chapter. In general, most cognitive modeling efforts have suffered not from an inability to fit data but from an ability (driven by an overabundance of free parameters) to fit *any* possible pattern of data. In such situations, the good fit of a model confers no support whatsoever for the psychological reality of that model, since other models with very different assumptions could fit the data just as well. For discussion of this critical issue see Roberts and Pashler (2000).

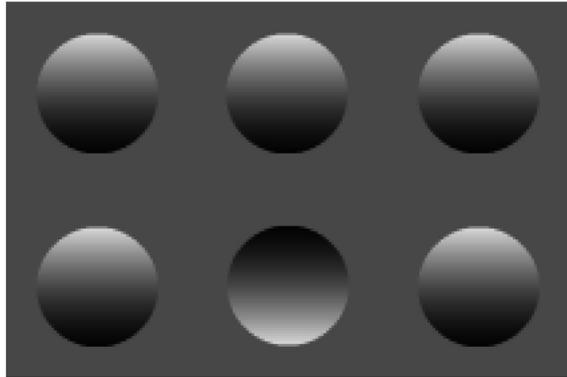


FIGURE 3.3 An example of the perceptual results of the visual system's assumption that light comes from overhead. The lower middle disc in the 'egg carton' can be readily seen as a concave hole, while the other discs appear as convex bumps. Turning the picture upside down reverses this pattern. See text for discussion (adapted from Hoffman, 1998, and Ramachandran, 1988).

intuitive sense: when facing a surface that is itself facing the sun at an oblique angle, bumps on the surface will in fact be lighter toward their tops, while dents will be lighter toward their bottoms—due simply to the differential shadows produced in each case.

The critical thing to note about this phenomenon, however, is that the inference from shading to shape is only reliable given two assumptions. The first assumption is that there is only a single light source (since all of the discs in fig. 3.3 could be bumps if you assume that the lower-middle disc is simply lit by a separate light source—an assumption your visual system is not willing to make). Second, this inference is only valid if you assume that the light source is overhead (e.g., Berbaum et al., 1983; Rittenhouse, 1786)! This appears to be another of the assumptions that the visual system uses as a heuristic cue when computing shape from shading, in an attempt to combat underdetermination.<sup>8</sup> Where does this assumption come from—is it innate or is it learned from experience with the sun overhead? At least in some species, there is good evidence to suggest that this assumption is innately determined. Hershberger (1970), for example, raised chickens from egghood in an

8. Here are two other interesting facts about this bias: first, note that the assumption is of a bias from *overhead* rather than from *above*. The reason for this is that the light-source assumption appears to operate in a head-centered reference frame rather than a world-centered reference frame (Howard et al., 1990)—a fact that you can see for yourself by studying figures 3.3 and 3.4 while standing on your head. Second, more contemporary studies have actually demonstrated that most observers also assume that the single light source comes from a bit to the left of center, in addition to being overhead (Mamassian & Goutcher, 2001; Sun & Perona, 1998)! While good explanations of this are hard to come by, it may be related to the fact that we consistently orient our bodies relative to light sources when manipulating objects with our hands, combined with the fact that most people are righthanded (Sun & Perona, 1998).

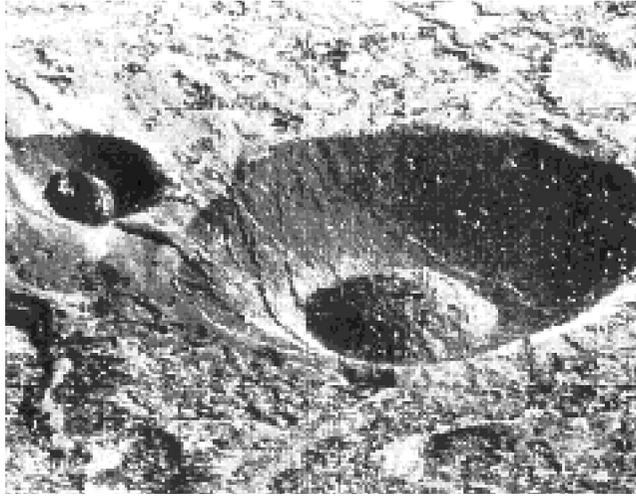


FIGURE 3.4 An example of the “overhead illumination” assumption in a real photograph of moon craters. Turning the picture upside-down reverses convex and concave.

environment that was always lit from below, and trained them to discriminate bumps from dents. When the chickens were later tested (in isotropic lighting conditions) using stimuli such as the discs in figure 3.3, their behavior indicated that their visual systems still assumed that the light source came from overhead—directly contradicting all of their visual experience! As a result, he concluded that “there appears to be an innate perceptual parameter corresponding to an ‘overhead source of illumination’” (p. 407).

Recent Bayesian modeling work has successfully cast this assumption as a Bayesian prior (Mamassian & Goutcher, 2001; Mamassian et al., 2002). Though the details of this model are beyond the scope of this chapter, it is worth noting that such a modeling effort has several advantages over simply talking in a vague but interesting way of an “assumption” about overhead illumination that is wired into the visual system. First, a Bayesian model shows how such a “principle” can exist in a more general visual architecture that also incorporates various other assumptions and priors. Second, such a model allows for a detailed study of how the “overhead illumination” prior interacts with other specific priors (such as the assumption of an overhead *viewpoint*; Mamassian & Landy, 1998), and how priors can interact and compete.

Third, Bayesian models embody optimal reasoning under uncertainty, and as such they allow for an objective assessment of how good human performance is relative to an “ideal observer.” Fourth, and perhaps most important, the model allows for a detailed investigation of the reliability of the prior: while human performance may always assume that the illumination comes from a particular angle, the model’s illumination-location prior can be set to multiple values, and the resulting impact on discrimination can be rigorously studied. In this way the Bayesian model serves as both a *tool* (for really understanding what an

overhead-illumination “assumption” could mean, and why and how it might have an impact on perception) and a *theory* (of how the human visual system actually employs the assumption). In this way, Bayesian models have allowed for “natural constraints” to be translated from vague statements into rigorous testable theories—“a psychophysics of constraints” (Knill et al., 1996).

Of course, this is just one example, but a Bayesian approach has also been successfully employed in similar ways to model phenomena in many other domains of visual perception—including the perception of color (Brainerd & Freeman, 1997), motion (Weiss et al., 2002), shape (Feldman, 2000), surfaces (Nakayama & Shimojo, 1992), and higher level aspects of perception such as object recognition (Liu et al., 1995) and perceptual grouping (Feldman, 2001). Of course there are many limitations to Bayesian modeling in perception (see note 7): for example, the Bayesian approach tends to work well for modeling specific functions, but non-Bayesian theories must account for the modular nature of perception itself (Knill et al., 1996) and are in many cases responsible for uncovering particular priors. Still, the Bayesian approach fits human performance exceedingly well in several contexts, and it allows for a detailed scientific understanding of how “principles” may be wired into the visual system. Moreover, because the Bayesian approach is in some sense optimal (a fact that motivates the increasingly common study of “ideal observers” in perception research), it is perhaps the most successful contemporary realization of David Marr’s famous advice that the mind should be studied first from the computational level, when possible, so that we understand the nature of the information-processing problems themselves that the mind faces.

#### 4 Reconciling Nativism and Development: Clues from Vision Science

This section aims to make a very simple point: there is no mystery in many theories of visual perception about how nature and nurture can interact. More specifically, both theories of Bayesian perception (sec. 4.1) and theories that appeal to spontaneous endogenous input (sec. 4.2) make it entirely unmysterious how the very same process can both be innately determined and yet later develop robustly via interaction with the environment.

##### *4.1 Nature and Nurture Are Easily Combined in Bayesian Perception*

As discussed in section 2, much of the controversy about nativist theories of higher level cognition has focused on the perceived conflict between innateness and development. Often nativism is simply written off as antidevelopmental, but even careful researchers tend to relegate development to processes outside the actual nativist components of their theories—to maturation, triggering, overriding, or simply other non-innate aspects of the mind. In contrast, Bayesian theories of visual perception constitute a case study of how such a separation is unnecessary. This simple point falls directly out of the general structure of the Bayesian framework. Innate assumptions and principles, in this framework, are realized as priors. And the underlying structure

of these priors are simply distributions of values of variables—“ $p(\text{Hypothesis})$ ” for each hypothesis (scene) under consideration. Thus an innate principle—for example, the assumption of a single overhead light source discussed in section 3.3—is architecturally realized as the default value or “factory setting” of the relevant variable. But in no sense is that principle then written in stone, since its value can later be updated and tuned via interaction with the environment. This point has never really been stressed by Bayesian theorists in vision science, perhaps because they tend not to traffic in debates about nativism in their day-to-day research.

Many priors may be innate. This has been empirically demonstrated in some cases (e.g., the overhead-illumination principle; Hershberger, 1970) and is widely assumed in many others—including the visual system’s prior assumptions that objects are rigid (Ullman, 1979), that objects are convex (e.g., Hoffman & Richards, 1984), and that motion is relatively slow (Weiss et al., 2002). In general theorizing beyond the scope of particular priors, moreover, most Bayesian theorists are happy to accept the possibility that “the priors are in the genes” (Kersten et al., 2004, p. 285). Other priors are probably formed by combining innate constraints with learning, however: “undoubtedly, the prior probability and likelihood distributions incorporated implicitly into the visual system arise through a combination of evolution and perceptual learning” (Geisler & Kersten, 2002, p. 509). Some specific priors, such as the bias for assuming that certain ambiguous image angles correspond to right angles in the actual scenes (e.g., Halper, 1997), may even have little or no innate component. In general, recent studies have shown that human perceptual systems—even those of young infants—are surprisingly good at picking out even subtle static and temporal statistical regularities in the environment (Chun, 2000; Fiser & Aslin, 2001, 2002a, b). Other theories of “opportunistic” learning suggest that some priors might be learned at only particular times when the visual system recognizes that relatively rare conditions of low ambiguity obtain (see Brady & Kersten, 2003; Kersten et al., 2004).

More generally, the Bayesian framework makes plain how priors can be *both* innate and later tuned via interaction with the environment: it is as simple as initializing a variable to a default value, and later updating the value of that variable (where each variable is really a distribution of values).<sup>9</sup> Indeed, it seems likely that this is the most common situation, since perceptual systems should take advantage of both nature and nurture as much as possible: “The Bayesian approach allows one to understand precisely how the reliability of different sources of information, including prior knowledge, should be combined by a perceptual system. Different sources of information do not always keep the same relative reliability, and hence a rational perceptual system should adjust the weights that it assigns to different information sources contingent upon their current relative reliabilities” (Geisler & Kersten, 2002, p. 509).

This updating could take place in several ways: by updating the actual distribution of a prior, by changing the relative weightings of various priors, by

9. Though many neuroscientists do not often talk about variables being stored in neural tissue, we know that they must exist, even in much simpler creatures such as ants (see Gallistel, 2000, for discussion).

adjusting the likelihood function, or even by adding (or removing) priors altogether. The critical point here is that it isn't the least bit mysterious how any of these forms of environmental tuning is possible, and several types of learning would occur not outside of the nativist framework but to *the very information that is itself thought to be innate*. Moreover, this type of updating could encompass bona fide *learning* (Pylyshyn, 1985). Unlike triggering, maturation, and strong parameterization, this form of development is not limited to highly constrained environmental cues but can in principle encompass entirely new and unexpected forms of information. Thus it needn't be true in a nativist theory that "the constraints remain unchanged through life" (Gopnik, 2003, p. 239). Rather, the priors in such architectures may simply characterize the *initial* state of the system, and may come to be updated or supplanted later. Moreover, some of the nativist content in a Bayesian framework may even be particular innate strategies for later learning: "it is largely an open question of how the human visual system learns the appropriate statistical priors, but some priors *as well as strategies for learning priors* are likely to be rooted in our genes" (Kersten & Yuille, 2003, p. 151, emphasis added).<sup>10</sup> In all of these ways, Bayesian theories of perception constitute a case study of an architecture in which innateness and development do not inherently conflict, and can richly interact.

#### 4.2 *A Second Example: Reconciling Nativism and Plasticity in Vision via Spontaneous Endogenous Activity*

In fact, several types of theories of visual perception—not only those based on statistical inference—invoke both innateness and development, and make clear how these processes can comfortably coexist in the very same visual mechanisms. In this section I will mention one other general strategy for reconciling nativism and plasticity in vision science, which invokes spontaneous endogenous activity in perceptual systems, and has its roots in visual neurophysiology.

When most cognitive psychologists think about the role of environmental influences on the design of the visual system, they are likely to recall the textbook cases of the effects of deprivation on the development of ocular dominance columns. These famous studies by Hubel and Wiesel (1965, 1970; Wiesel & Hubel, 1965a, b) involved raising cats in conditions of either complete monocular deprivation or severely attenuated binocular input (via imposed strabismus) and showing that these early experiences severely impaired the development of normal structure in visual cortex and also the resulting binocular visual function. Similar work has shown that exposing animals to contours of only certain orientations during rearing leads to a predominance of cells that select for those orientations,

10. Detailed theories of how the visual system should go about learning priors and tuning innate priors have yet to be developed with the same rigor that characterizes most Bayesian theories, but landmarks for this project do exist. In particular, detailed models have recently been proposed for how natural selection itself can be modeled in Bayesian terms (Geisler & Diehl, 2002), and it seems likely that such models will apply to both phylogenetic and ontogenetic updating.

and a relative paucity of selective cells for those orientations the animals have never seen (e.g., Blakemore & Cooper, 1970; Sengpiel et al., 1999).

This well-known research tradition, however, is not the full story. More contemporary work in neurophysiology has demonstrated that the development of ocular dominance columns, for instance, occurs in two stages: the initial formation of the structures and subsequent environmental tuning. Effects of deprivation tend to affect the second stage, whereas the basic structure of cortical circuits forms even without any visual experience (for a review see Katz et al., 2000). Moreover, we have some idea of the mechanism by which this is possible: spontaneous internally generated cortical signals (e.g., O'Donovan, 1999; Shatz, 1996). That is, innate structure is formed via the very same mechanism that drives later environmental tuning; the difference concerns only whether the signals are generated internally or externally. For the development of ocular dominance columns, this strategy is particularly effective: since the two eyes are unconnected, spontaneous early visual activity in the retina creates statistically independent oscillating signals from each eye, which can be used in the developing cortex to determine that the activity arises from two separate loci (see Katz et al., 2000). In the context of orientation tuning a similar story holds, and the disruption of this spontaneous activity correspondingly impairs orientation selectivity (Weliky & Katz, 1997).

This type of situation—wherein nature and nurture operate on the very same mechanisms via internal versus external signals—may apply much more generally to the development of perceptual abilities, including higher level perceptual skills. For example, face recognition is an area of perception research that has also seen the type of conflict between innateness and development that is characteristic of debates about higher level cognition (see Pascalis & Slater, 2003), and many popular models of face perception assume that innate factors and subsequent learning occur in completely separate brain systems (e.g., CONLERN v. CONSPEC, in Morton & Johnson, 1991). Again, however, this separation between nature and nurture may not be necessary. More recent models of face perception, for instance, suggest that the very same processes may be innately structured via spontaneous endogenous activity, and then later develop via external input (e.g., Bednar, 2003). Only the initial endogenous pattern generators are assumed to be innate. (Note that such generators may involve bona fide innate structure, not just innate signals, as in the case of uncorrelated input from the two retinas.) The rest of the machinery involved in face perception simply acts on that input in the very same way that it acts on environmental input, and such processing needn't even “know” whether its input consists of innately driven spontaneous signals or faces in the external world. Moreover, as with Bayesian theories, detailed computational models of this process exist, demonstrating that there is nothing mysterious about this clever marriage of innateness and development (Bednar & Miikkulainen, 2003).

#### 4.3 *Conclusions: From Principles to Priors?*

The same ultimate lesson can be drawn from work on both Bayesian models of perception and models of spontaneous endogenous activity in early visual development: the widespread perception of an inherent conflict between innateness

and development is illusory. Innateness and development can act together in several ways, and can even act on the very same underlying processes. Innately specified structure can *itself* develop, and there is nothing mysterious about this process.

A goal for future work will be to adapt these types of architectures to the cases of higher level cognition in which innateness and development have not fit so well together. In both of the domains highlighted in section 2—infant object-cognition and theory of mind—nativist theories have tended to be phrased in terms of static-sounding principles and concepts, and such characterizations have tended to fuel the perception that such innate structure is incapable of significant change. But note that the priors of Bayesian perception sound equally static and inflexible when phrased in terms of abstract principles and assumptions (“Illumination comes from overhead”). Rigorous models have demonstrated how such abstract principles can be implemented in ways that are malleable, however, and it will be worth exploring such options (along with the possibility of spontaneous endogenous signals) in cases of higher level cognition. Some of the development in these domains may in fact involve more extrinsic or peripheral developmental processes such as triggering, maturation, overriding, parameterization, and so on, but these needn’t be the whole story. Innate structure can in principle develop richly in response to environmental stimulation, and this development can bring about radical changes even in the initially innate structure itself.

The fact that nativist theories in higher level cognition also often appeal to modularity is no reason to doubt that such strategies can work. In practice, aspects of visual perception—including the foci of most low-level Bayesian modeling—are thought to be far more modular than higher level cognition (though see Carruthers, chapter 5 here; Sperber, chapter 4 here), and this has in no way frustrated the development of successful models that blend nature and nurture. In addition, there is no inherent conflict in principle between modularity and development (Scholl & Leslie, 1999b). Modules are often characterized primarily in terms of informational encapsulation, such that they can only act on a restricted range of possible inputs (Fodor, 1983, 2000). But this in no way stops modules from developing internally—and even altering their own input restrictions—on the basis of the information they do receive. Innate modules can also develop, in straightforward ways. Constructing theories such as those based on Bayesian inference or spontaneous endogenous input may be more challenging in the context of higher level cognition, but the work from visual perception can serve as a useful guide for such a project and demonstrates in rigorous case studies that there is at least no inherent conflict between innateness and development.

Even these future directions, however, are only a small piece of a much larger project. Researchers who are exploring foundational issues such as innateness, modularity, computation, evolution, and representation—all foci for intense controversy in the study of higher level cognition—might do well to explore how such notions are treated not only in simpler organisms (Gallistel, 2000) but also in simpler, less controversial, and more rigorously understood cognitive processes, such as visual perception. This is one specific context in which we see the hope of cognitive science: that scholars from superficially different research areas may still find common connections and useful insights that apply broadly to the organization of the mind.