ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/282045360

The Interface Theory of Perception

ARTICLE *in* PSYCHONOMIC BULLETIN & REVIEW · SEPTEMBER 2015

Impact Factor: 2.99 · DOI: 10.3758/s13423-015-0890-8

CITATIONS
7

READS
95

3 AUTHORS:



Donald Hoffman
University of California, Irvine
97 PUBLICATIONS **3,032** CITATIONS
SEE PROFILE



Manish Singh
Rutgers, The State University of New Jersey
108 PUBLICATIONS **2,008** CITATIONS
SEE PROFILE



Chetan Prakash
California State University, San Bernardino
10 PUBLICATIONS **172** CITATIONS
SEE PROFILE

Available from: Donald Hoffman
Retrieved on: 13 March 2016

Psychon Bull Rev
DOI 10.3758/s13423-015-0890-8

THEORETICAL REVIEW

The Interface Theory of Perception

Donald D. Hoffman[1] · **Manish Singh**[2] · **Chetan Prakash**[3]

© Psychonomic Society, Inc. 2015

Abstract Perception is a product of evolution. Our perceptual systems, like our limbs and livers, have been shaped by natural selection. The effects of selection on perception can be studied using evolutionary games and genetic algorithms. To this end, we define and classify *perceptual strategies* and allow them to compete in evolutionary games in a variety of worlds with a variety of fitness functions. We find that veridical perceptions—strategies tuned to the true structure of the world—are routinely dominated by nonveridical strategies tuned to fitness. Veridical perceptions escape extinction only if fitness varies monotonically with truth. Thus, a perceptual strategy favored by selection is best thought of not as a window on truth but as akin to a windows *interface* of a PC. Just as the color and shape of an icon for a text file do not entail that the text file itself has a color or shape, so also our perceptions of space-time and objects do not entail (by the *Invention of Space-Time Theorem*) that objective reality has the structure of space-time and objects. An interface serves to guide useful actions, not to resemble truth. Indeed, an interface *hides* the truth; for someone editing a paper or photo, seeing transistors and firmware is an irrelevant hindrance. For the perceptions of *H. sapiens*, space-time is the desktop and physical objects are the icons. Our perceptions of space-time and objects have been shaped by natural selection to hide the truth and guide adaptive behaviors. Perception is an adaptive interface.

Keywords Bayesian inference and parameter estimation · Perceptual learning · Perceptual categorization and identification · Spatial cognition · Evolution

✉ Donald D. Hoffman
ddhoff@uci.edu

[1] Department of Cognitive Sciences, University of California, Irvine, CA 92697, USA

[2] Department of Psychology and Center for Cognitive Science, Rutgers University, New Brunswick, NJ, USA

[3] Department of Mathematics, California State University, San Bernardino, CA, USA

Introduction

What is the relationship between perception and reality? This question motivated Fechner in 1860 to launch the field of psychophysics—and of experimental psychology more generally. Today, with the benefit of advances in experimental psychology and evolutionary biology, a broad consensus has emerged among perceptual scientists: *Natural selection has shaped our perceptions to be, in the typical case, accurate depictions of reality, especially of those aspects of reality that are critical for our survival.*

This consensus is spelled out in a standard textbook on vision: "Evolutionarily speaking, visual perception is useful only if it is reasonably accurate.… Indeed, vision is useful precisely because it is so accurate. By and large, *what you see is what you get.* When this is true, we have what is called veridical perception…perception that is consistent with the actual state of affairs in the environment. This is almost always the case with vision…" (Palmer 1999, emphasis his).

Marr (1982, p. 340) agrees: "We…very definitely do compute explicit properties of the real visible surfaces out there, and one interesting aspect of the evolution of visual systems is the gradual movement toward the difficult task of representing progressively more objective aspects of the visual world."

Pizlo and his collaborators (2014, p. 227) agree: "We close by restating the essence of our argument, namely, veridicality is an essential characteristic of perception and cognition. It is

Published online: 18 September 2015

absolutely essential. *Perception and cognition without veridicality would be like physics without the conservation laws.*" (emphasis theirs.)

The evolutionary theorist Trivers (2011) also agrees: "… Our sense organs have evolved to give us a marvelously detailed and accurate view of the outside world—we see the world in color and 3-D, in motion, texture, nonrandomness, embedded patterns, and a great variety of other features. Likewise for hearing and smell. Together our sensory systems are organized to give us a detailed and accurate view of reality, exactly as we would expect if truth about the outside world helps us to navigate it more effectively."

In mathematical models of perception, this theory often is couched in the language of Bayesian estimation, the idea being that evolution shaped our perceptual systems to estimate accurately, on the basis of sensory information, the true state of the environment. Yuille and Bülthoff (1996) for instance: "We define vision as perceptual inference, the estimation of scene properties from an image or sequence of images…there is insufficient information in the image to determine uniquely the scene. The brain, or any artificial vision system, must make assumptions about the real world. These assumptions must be sufficiently powerful to ensure that vision is well-posed for those properties in the scene that the visual system needs to estimate."

Geisler and Diehl (2003, p. 397) agree: "In general, it is true that much of human perception is veridical under natural conditions. However, this is generally the result of combining many probabilistic sources of information…Bayesian ideal observer theory specifies how, in principle, to combine the different sources of information in an optimal manner in order to achieve an effectively deterministic outcome."

Why should evolution favor veridical perceptions? The intuition is that those who see more truly outcompete those who see less truly and thus are more likely to pass on their genes that code for truer perceptions. Thousands of generations of this process have spread the genes for veridical perceptions throughout our species. We are thus the offspring of those who, in each generation, saw a bit more truly, and we can be confident that we too, in most situations, have veridical perceptions.

Although this is considered a good argument for veridical perception in humans, it is not considered so for simpler organisms, such as insects and amphibians. Marr, for instance, argued that fly vision, unlike human vision, is nonveridical: "Visual systems like the fly's serve adequately and with speed and precision the needs of their owners, but they are not very complicated; very little objective information about the world is obtained. *The information is all very subjective*…" [emphasis added] and "…it is extremely unlikely that the fly has any explicit representation of the visual world around him—no true conception of a surface, for example, but just a few triggers and some specifically fly-centered parameters.…" (1982, p. 34). In this quote, Marr explicitly states his view that for the fly "the information is all very subjective." Marr's whole point in discussing the fly is to show that fly vision can successfully control the flight of the fly without needing to compute objective descriptions of the world.

Similarly, Marr thought that frog vision was nonveridical: "In a true sense, for example, the frog does not detect flies—it detects small, moving, black spots of about the right size. Similarly, the housefly does not really represent the visual world about it—it merely computes a couple of parameters…which it inserts into a fast torque generator and which cause it to chase its mate with sufficiently frequent success" (1982, p. 340). Marr explained why evolution might shape nonveridical perceptions: "One reason for this simplicity must be that these facts provide the fly with sufficient information for it to survive. Of course, the information is not optimal and from time to time the fly will fritter away its energy chasing a falling leaf or an elephant a long way away.…" (1982, p. 34).

Marr's point is well taken. Natural selection is a search procedure that yields *satisficing* solutions, not optimal solutions. This is evident, for instance, in the backwards structure of the vertebrate eye, which forces light to pass through bipolar and ganglion cells before being caught by photopigments and which consequently requires a blind spot—a hole in the retinal mosaic—to allow the optic nerve to exit the eye; cephalopod eyes, which evolved separately, do not suffer these problems (Land and Nilsson 2012). Thus, perceptual systems that evolve by natural selection need not be optimal in structure and need not deliver optimal information, just information sufficient for survival and reproduction.

There are many examples of satisficing perception in nature. Dragonflies, for instance, have aquatic larvae and must find water to lay their eggs. Dragonfly vision has a simple trick to find water: Find horizontally polarized light reflections (Horvath et al 1998, 2007). Water strongly reflects horizontally polarized light, so this trick often guides successful oviposition. Unfortunately for the dragonfly, oil slicks and shiny tombstones also reflect such light, sometimes more strongly than water. Dragonflies are fooled by such slicks and tombstones to lay eggs where they cannot survive. In the niche where dragonflies evolved, their perceptual strategy normally works, but where that niche has been disturbed by *H. sapiens* with oil slicks and tombstones, the same strategy can be fatal.

Male jewel beetles fly about looking for the glossy, dimpled, and brown wing-casings of females. When males of *H. sapiens* began tossing out empty beer bottles that were glossy, dimpled, and just the right shade of brown, the male beetles swarmed the bottles and ignored the females, nearly causing the extinction of the species (Gwynne and Rentz 1983). The beetles' perceptions relied not on veridical information but rather on heuristics that worked in the niche where they evolved.

Thus, natural selection has shaped the perceptual systems of many organisms to rely on fallible heuristics. Yet there is consensus among perceptual scientists that natural selection has shaped the perceptions of *H. sapiens* to be, in the normal case, veridical. This raises obvious questions: What precisely are the conditions in which natural selection favors veridical perceptions? Are we correct in assuming that *H. sapiens*, unlike flies, frogs, and beetles, has been shaped to have veridical perceptions? Is there really such a discontinuity between *H. sapiens* and other animal species?

Fortunately, we aren't forced to speculate about the answers to these questions. Evolution by natural selection can be studied using mathematical tools, such as evolutionary game theory, evolutionary graph theory, and genetic algorithms (Hofbauer and Sigmund 1998; Mitchell 1998; Lieberman et al. 2005; Nowak 2006; Samuelson 1997; Sandholm 2007). One can study competitions between perceptual strategies and compute probabilities for strategies to emerge, go extinct, coexist, or dominate.

The first step is to define *perceptual strategies* and classify them by how informative they are about the objective world. This yields an understanding of possible relationships between perception and reality that is more nuanced than a simple dichotomy between veridical or not. We can use evolutionary games and genetic algorithms to study the relative fitness of these perceptual strategies.

The collection of strategies considered must be exhaustive; otherwise we might miss a winning strategy. In particular, we must include strategies that see none of the true facts, some of the true facts, and all of the true facts. Even if we suppose that human perception is veridical today, we must consider all possible strategies, veridical or not, in order to explore the plausible hypothesis that we evolved from species whose perceptions were not veridical. And we must entertain the hypothesis that any one of these strategies, veridical or not, might have evolved by natural selection, even for human perception.

We will see that there are really two separate questions to be answered. First, is the *vocabulary* of our perceptions isomorphic to aspects of objective reality so that our language of perceptions could, in principle, describe the objective truth? Second, if so, do our perceptual systems, using that vocabulary, in fact *succeed* in describing the true state of affairs in the world?

With this background, we now define the strategies that we will study.

Perceptual strategies

We need a definition of *perceptual strategy* that's broad enough to include all relevant strategies—otherwise our evolutionary games and genetic algorithms might inadvertently overlook viable strategies. Some models of color perception, for instance, posit a *metric* on color experiences and represent perceptual differences among colors by distances in the metric (Koenderink 2010; Mausfeld and Heyer 2003). This works well for color, but might not for other perceptions. Perhaps, for instance, no metric adequately models taste: Is a marinara sauce closer in taste to apples or to blueberries? This question might have no answer. Thus, it seems too restrictive to require that all perceptual strategies have associated metrics.

But it always seems necessary to model the *probabilities* of various perceptions and how these probabilities vary with states of the world. If, for instance, one assumes that the world contains surfaces and light sources, one must model how the probability of perceiving different colors is related to the reflectance functions of surfaces and the spectral distributions of the incident light sources. Or, if one assumes that the world contains various odorant molecules, one must model how the probability of perceiving different smells is related to the distribution of odorant molecules. Thus, it seems necessary to use so-called *measurable spaces* (i.e., probability spaces whose probability measure is not yet specified) to model possible perceptions and possible states of the world. And it seems necessary to require that probabilities of perceptual events can be systematically related to probabilities of events in the world, i.e., that the mapping from the world to perceptions is a so-called *measurable* mapping. (This is simply a generalization of the familiar notion of a random variable. For precise definitions of events, measurable spaces, measurable maps, and probability measures, see Appendix 1.)

These considerations lead us to define a perceptual strategy as follows. We represent the possible *perceptual experiences* of an organism by a measurable space $(X, \mathcal{X})$, where X is a set of possible experiences and $\mathcal{X}$ is a collection of subsets of X called *events*; in this case the events are perceptual events. The elements of X denote, we emphasize, *perceptual experiences* themselves and not, e.g., some kind of *objects* of perceptual experiences, such as so-called sense data. We represent the *world* by a measurable space $(W, \mathcal{W})$, where W is a set (of world states) and $\mathcal{W}$ is again a collection of subsets of W called events; in this case the events are in the world, but the notion of event used here differs from the notion of event in special relativity (i.e., a space-time point) and in particle physics (i.e., the results just after a fundamental interaction occurs between subatomic particles). Then the definition of a perceptual strategy is straightforward if there is no dispersion (such as noise), i.e., if each state $w \in W$ causes at most one perceptual experience $x \in X$.

Definition 1 A (dispersion-free) *perceptual strategy*, P, is a measurable function $P{:}W \to X$, where $(W, \mathcal{W})$ denotes a measurable space of states of the world and $(X, \mathcal{X})$ denotes a measurable space of perceptual experiences.

If there is dispersion (or noise), i.e., if there are states $w \in W$ associated to more than one perceptual experience $x \in X$, then the definition of a perceptual strategy requires a mapping that gives, for each state $w \in W$ of the world, the probabilities of the various perceptual experiences that the state w might cause. In the case where W and X are finite, this mapping can be written as a stochastic matrix, i.e., a matrix in which the values in each row sum to 1. In the more general case, this mapping can be written as a Markovian kernel P, which assigns to each $w \in W$ a probability measure, $P(w,\cdot)$, on $\mathcal{X}$ (Revuz 1984). In information theory, such Markovian kernels are *communication channels* (Cover and Thomas 2006). With this background the definition of a perceptual strategy when there is dispersion is as follows.

Definition 2 A *perceptual strategy with dispersion* is a Markovian kernel $P: W \times \mathcal{X} \rightarrow [0,1]$, where W denotes a measurable space of states of the world and $\mathcal{X}$ denotes the events for a measurable space X of perceptual experiences.

In this paper, we focus on dispersion-free strategies; the pattern of results we find also holds when there is dispersion. We begin with the strongest kind of veridical strategy, which we call the *omniscient realist* strategy, which accurately sees *all* aspects of the objective world and its structures. Thus, we define an *omniscient realist* strategy as follows.

Definition 3 An *omniscient realist* strategy is a perceptual strategy for which $X = W$ and P is an isomorphism, i.e., a one-to-one and onto map that preserves all structures on W (e.g., topologies, partial orders, groups).

Omniscient realism is, for good reason, not taken seriously by perceptual scientists or philosophers (except, perhaps, for the odd metaphysical solipsist). In the case of vision, for instance, it's widely agreed that we see just a small fraction of the electromagnetic spectrum, and only the front surfaces of opaque solid objects. Thus, we see, at best, a small part of the objective world, which contradicts the condition $X = W$ of omniscient realism. However, we include omniscient realism for sake of completeness.

Naïve realist strategies model a weaker version of veridical perception in which the perceiver accurately sees all aspects and structures of a *subset* of the objective world. In philosophy, naïve realism is roughly the view that what you see is really there even in the absence of any perceiver, that "objects of awareness are actually the mind independent objects that inhabit the world" (Fish 2010). Versions of naïve realism have long been debated and are defended to this day (e.g., Brewer 2011; Campbell and Cassam 2014; Chemero 2009; Fish 2009; Gibson 1986; McDowell 1996; Noë 2012; Searle 2015; Travis 2013). Searle (2015), however, prefers the term "direct realism" rather than naïve realism, to distinguish his view from disjunctivist accounts.

Gibson (1986), for instance, says "The environment consists of the earth and the sky with objects on the earth and in the sky, of mountains and clouds, fires and sunsets, pebbles and stars. Not all of these are segregated objects, and some of them are nested within one another, and some move, and some are animate. But the environment is all these various things—places, surfaces, layouts, motions, events, animals, people, and artifacts that structure the light at points of observation." According to Gibson's ecological theory of vision, we directly and truly see those aspects of the world that are *affordances*, i.e., those aspects relevant to "what it *offers* the animal, what it *provides* or *furnishes*, either for good or ill" [emphasis his]. Gibson's theory of vision also is motivated by evolution but, as we will see, his conclusions about how evolution shapes perception differ dramatically from ours.

We mathematically define a *naïve realist* strategy as follows.

Definition 4 A *naïve realist* strategy is a perceptual strategy for which $X \subset W$ and P is an isomorphism on this subset that preserves all structures on W.

One argument against naïve realism cites the phenomenon of metamers, in which illuminants with different spectra look the same color, or surfaces with different reflectance functions look the same color. This is sometimes taken to show that color is not part of the objective world. Hardin (2008, p. 143), for instance, says "Perceived colors are therefore two removes from the occurrent bases of the dispositions to see them. Many different mechanisms can produce the same SPD [spectral power distribution], and many different SPDs can cause us to see the same color. It is also important to note that animals with different receptoral sensitivities are unlikely to experience the same colors that we do under the same circumstances. It is little wonder that color categories have been described as 'gerrymandered' and 'anthropocentric.'"

Hardin goes on to note that "a color realist's appeal to 'normal' or 'standard' conditions to determine the 'true' or 'actual' colors of objects is mere hand-waving unless there is some clear reason for preferring one set of illumination or background conditions to another. So far, nobody who has held a realist position has been prepared to propose and defend such a set of conditions. What is to be said about the other half of the equation, the 'normal' observer to whom philosophers so casually refer?"

Not everyone agrees, however; some philosophers claim that color is indeed part of the objective world (e.g., Byrne and Hilbert 2003).

Examples, such as metamers, suggest the need to consider *critical realist* strategies, in which the perceptions need not be a subset of the objective world, but in which relations among perceptions nevertheless preserve relations between states in the objective world. Thus, we define a *critical realist* strategy as follows.

Definition 5 A *critical realist* is a perceptual strategy for which X need not be a subset of W, but P is nevertheless a homomorphism that preserves all structures on W.

Many scientists and philosophers today are critical realists but of a special type that we will call *hybrid realists*. They claim that some of our perceptual experiences, such as color and taste, are not part of the objective world, but that other perceptions, such as object shapes and motions, are part of the objective world (Pizlo 2010; Pizlo et al. 2014). Among philosophers this view is sometimes expressed as a variant of Locke's distinction between primary and secondary qualities (Locke 1690). This gets a bit tricky because, according to Locke, secondary properties, such as color, are strictly speaking, *dispositions* of mind-independent objects, i.e., dispositions to trigger perceptual experiences in us that we describe with terms such as colors; but these perceptual experiences, according to Locke, do not *resemble* any objective properties of the mind-independent objects. Thus, one must be careful when interpreting the writing of hybrid realists to determine when they are discussing properties, such as dispositions, of mind-independent objects and when they are discussing perceptual experiences that are the consequences of such properties. Exegesis on this point can be controversial.

Hybrid realism dates back at least to the early years of science. Galileo, for instance, said "I think that tastes, odors, colors, and so on are no more than mere names so far as the object in which we locate them are concerned, and that they reside in consciousness. Hence if the living creature were removed, all these qualities would be wiped away and annihilated" (1623/1957, p. 274).

We define hybrid realism as follows.

Definition 6 A *hybrid realist* strategy is a critical realist strategy that requires that there exists a strict subset $\hat{X} \subset X$ that satisfies $\hat{X} \subset W$ and requires that P is an isomorphism on this subset that preserves all structures.

Most vision researchers who use Bayesian models of perception assume hybrid realism (Knill and Richards 1996). They assume, for instance, that our perceptions of object *shapes* are normally veridical and that Bayesian techniques illuminate how we estimate true shapes from images. They typically assume, however, that *color* is not an objective property of the world but that Bayesian methods can model the relationship between perceived colors and, say, equivalence classes of surface reflectances.

However, there is a class of perceptual strategies even more general than critical realist and hybrid realist. This class of strategies, which we call *interface strategies*, does not require any perceptions to be veridical or to reflect any structures of objective reality, such as orders or metrics. Therefore, we define interface strategies as follows.

Definition 7 An *interface perceptual* strategy is a perceptual strategy that does not require X to be a subset of W and for which the mapping P has no restrictions other than being measurable (so that the probabilities of perceptions are systematically related to probabilities of events in W).

Thus, an interface strategy is simply a dispersion-free perceptual strategy, with no additional constraints. However, the new name is useful for understanding these strategies. Consider a *strict* interface strategy, i.e., an interface strategy that is *not* a critical realist strategy. For such a strategy, *no* perceptions are veridical ($X \not\subset W$) and *no* structure of W is preserved other than measurable structure (P is not a homomorphism). It is natural to ask for such a strategy how it could possibly be useful to an organism. If none of its perceptions are veridical, and none of its perceptions reflect the structure of the world, then aren't its perceptions completely useless?

It turns out that they can, in fact, still be quite useful, and a familiar metaphor helps to see this. Consider the desktop of the windows interface on your laptop computer. Suppose that there is a blue rectangular icon in the upper right corner of the desktop for a text file that you are editing. Does this mean that the text file itself is blue, rectangular, or in the upper right corner of the laptop? Of course not. Anyone who thinks so misunderstands the purpose of the desktop interface. No features of the icon are identifiable with any features of the file in the computer. Moreover, one would be hard pressed to find a natural sense in which the icon is a veridical representation of the file. However, the icon is intended to guide useful behaviors. If, for instance, you drag the blue icon to the trash you can delete the text file; if you drag it to the icon for an external drive, you can copy the file.

So if our perceptions are in fact *strict* interface perceptions, then none of our perceptions are veridical and none of our perceptions reflect the structure of the world. This would mean that our perceptions of physical objects and even of space-time itself are not veridical. Instead, space-time would be our desktop and physical objects would be the icons on the desktop. If natural selection has appropriately shaped our perceptions of space-time and physical objects, then they could still be useful guides to behavior even though they are not veridical. It is for this reason that the most general perceptual strategies are called interface strategies.

The interface metaphor is offered merely as an aid to intuition. Like most metaphors, it suffers weaknesses. One might for instance argue that—contrary to the "simplify and hide" hallmark of interfaces that we have touted—the nesting of folder icons on a desktop is an accurate guide to the nesting of real folder hierarchies in the computer. This critique is well taken. However, it is the interface strategy as precisely defined, not as metaphor, which lives or dies by the sword in our evolutionary games. Moreover, the strength of the metaphor is what it highlights: The simplicity of a desktop, which hides the complexity of the computer, and the nonveridicality

of a desktop, which allows it to be tailored instead to the needs of the user, are in fact huge advantages that promote efficient interactions with the computer.

The relationships among strategies are shown in Fig. 1. One sees from the diagram, for instance, that all hybrid realists and naive realists are critical realists but that no hybrid realist is a naive realist; some critical realists are neither hybrid realists nor naive realists.

Evolutionary games

Does natural selection favor veridical perceptions? The five classes of perceptual strategies defined in the previous section allow us to ask this question with greater precision: Which of the five perceptual strategies are favored by natural selection, and under what conditions are they favored?

Fortunately, we need not speculate about the answer. We can devise evolutionary games and genetic algorithms to obtain precise answers in precise contexts, and from these we can extrapolate general principles.

Using evolutionary games, we compel perceptual strategies to compete in a variety of environments and under a variety of selection pressures, and discover which will coexist, disappear, and dominate.

Evolutionary games have the power to model frequency-dependent selection, in which the fitness of strategies is not fixed, but instead varies with the proportion of individuals in the population that employ each strategy (Allen and Clarke 1984; Hofbauer and Sigmund 1998; Nowak 2006; Samuelson 1997; Sandholm 2007).

For instance, frequency-dependent selection governs the strategies of hunter-gatherers who share their daily haul. Some are "producers" and work hard to hunt and gather, whereas others are lazy "scroungers" and simply eat what others provide (Barnard and Sibly 1981). If most are producers, then scroungers do well; but as the proportion of scroungers increases, the fitness of their strategy declines until, in the limit where everyone scrounges, everyone starves.

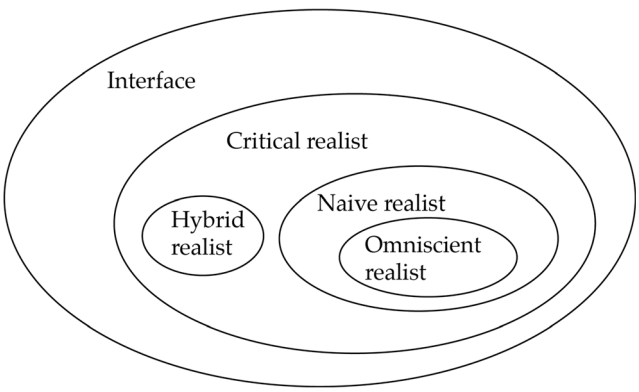


Fig. 1 Venn diagram of the relationships among the different perceptual strategies

A perceptual example is Batesian mimicry, in which a benign species avoids predation by resembling a dangerous species. In regions where the dangerous species is frequent, even poor mimics avoid predation; but where the dangerous species is infrequent, only good mimics enjoy this benefit (Harper and Pfennig 2007).

Evolutionary games assume infinite populations of players, each having a fixed strategy. Players are chosen at random to interact in games—a situation called *complete mixing*, because all interactions are equally likely. Each player receives a payoff from each of its interactions and, critically, this payoff is interpreted as fitness, and thus as reproductive success. This leads to natural selection: strategies that excel in games reproduce more quickly and thus outcompete other strategies.

Formally, natural selection is modeled by a differential equation called the *replicator equation* (Bomze 1983; Taylor and Jonker 1978). If n strategies interact, we let a_{ij} denote the payoff to strategy i when interacting with strategy j; we let $[a_{ij}]$ denote the $n \times n$ "payoff matrix" for all such interactions; and we let x_i denote the frequency of strategy i. Then, the expected payoff for strategy i is $f_i = \sum_{j=1}^{n} x_j a_{ij}$ and the average payoff is $\phi = \sum_{i=1}^{n} x_i f_i$. The replicator equation follows by equating payoff with fitness: $\dot{x}_i = x_i(f_i - \phi)$, where $i = 1, \ldots, n$ and $\dot{x}_i$ denotes the time derivative of the frequency of strategy i.

In the case of two strategies one finds the following: Strategy 1 dominates if $a_{11} > a_{21}$ and $a_{12} > a_{22}$; Strategy 2 dominates if $a_{11} < a_{21}$ and $a_{12} < a_{22}$; Strategies 1 and 2 are bistable if $a_{11} > a_{21}$ and $a_{12} < a_{22}$; Strategies 1 and 2 coexist if $a_{11} < a_{21}$ and $a_{12} > a_{22}$; Strategies 1 and 2 are neutral if $a_{11} = a_{21}$ and $a_{12} = a_{22}$ (Nowak 2006).

In the case of three strategies, there also can be cyclic domination among the strategies, much as in the children's game of Rock-Paper-Scissors. In the case of four or more strategies, there can be limit cycles and chaotic attractors (Nowak 2006).

The evolution of perceptual strategies has been studied in games that force players to compete for resources that are distributed over a set of territories (Marion 2013; Mark et al. 2010; Mark 2013). On each trial, quantities of resources are distributed at random (e.g., uniformly) to each territory. For each quantity of resources in a territory there is an associated payoff specified by a fixed payoff function. In each interaction, each player looks at each territory and decides which territory to seize. Each player receives the payoff for the resources in the territory it nabs.

These games have many variations, including the number of territories, the number and distributions of resources per territory, the payoff function, the order in which players choose, the perceptual strategies of the players, the number of distinct perceptions each

player can have, and the costs for computation and storage of information.

To see how an interface strategy differs from a critical realist strategy, consider the case where each territory has one resource that varies in quantity from 0 to 100, and where the perceptions of each player are limited to just four colors—e.g., red, yellow, green, and blue. Let the order on colors be the "energy order": red < yellow < green < blue. Let the payoff function be a (roughly) Gaussian function of the resource quantity: the greatest payoffs are associated with quantities near 50, and fall off for quantities greater or smaller than 50. Such a nonlinear payoff function is quite common: Not enough water and one dies of thirst; too much and one drowns; somewhere in between is just right. Similarly for salt and a variety of other resources. Indeed, for organisms that must maintain homeostasis of a wide variety of variables, one can expect many nonmonotonic payoff functions.

In this case, a critical realist whose perceptions veridically represent the quantity of resources is illustrated in Fig. 2. On the horizontal axis is the resource quantity, varying from 0 to 100. The translucent Gaussian depicts the payoff function (which is equated with fitness), having a maximum around 50. The colored rectangles indicate how resource quantities map to perceived colors. For instance, resource quantities between 0 and 25 map to red. This perceptual strategy is a critical realist because (1) the perceived colors are not a subset of the resources and (2) the mapping from resources to colors is an order-preserving homorphism, i.e., every resource quantity that maps to red is less than every resource quantity that maps to yellow, and so on.

If the resources are uniformly distributed then this critical realist, given its perceived color, can optimally estimate the true quantity of resources. If, for instance, it sees green, then it knows that the resource quantity lies between 50 and 75 with an expected value around 62.5. However, it cannot optimally estimate the payoffs. If it sees green, then the payoff could range from nearly 100 to less than 25; if it sees yellow, it's exactly the same—green and yellow are redundant. Thus this strategy is an efficient communication channel for information about the true value of the resource quantity but a poor channel for payoffs.

An interface strategy tuned to payoffs is illustrated in Fig. 3. It's a *strict* interface strategy, because the mapping from resource quantities to colors is not a homomorphism. For instance, some resource quantities that map to green are smaller than all resource quantities that map to blue (green bar on the left), but other resource quantities that map to green are greater than all resource quantities that map to blue (green bar on the right).

However, although this strategy is not a homomorphism for information about resources, it is for payoffs: All resource quantities that map to blue have higher payoffs than all resource quantities that map to green, and so on. This strategy is an efficient communication channel for information about payoffs but a poor channel for truth.

In evolution by natural selection, whenever payoffs and truth differ it is payoffs, not truth, that confer (indeed, are identified with) fitness and reproductive success.

From Monte Carlo simulations of many versions of this game the pattern is clear: *strict interface strategies that are tuned to fitness routinely drive naïve realist and critical realist strategies to extinction* (Marion 2013; Mark et al. 2010; Mark 2013). Adding more complexity to the environment, either by greatly increasing the number of territories or the number of resources per territory, doesn't help the realist strategies, in part because increasing complexity just saddles realist strategies with the burden of representing a greater quantity of irrelevant information; this extra burden reduces their fitness relative to the interface strategies, and thus pushes them to swifter extinction. Increasing costs for information and computation, or adding dispersion to the perceptual maps, generally makes matters worse for naïve realists and critical realists. The only situation in which realists have a chance against interface strategies is when payoff varies monotonically with resource quantity, i.e., when truths and payoffs are roughly the same thing.

The key insight from these evolutionary games is this: *Natural selection tunes perception to payoffs, not to truth*. Payoffs and truth are different, unless payoff functions happen to vary monotonically with truth. But we cannot expect, in general, that payoff functions vary monotonically with truth,

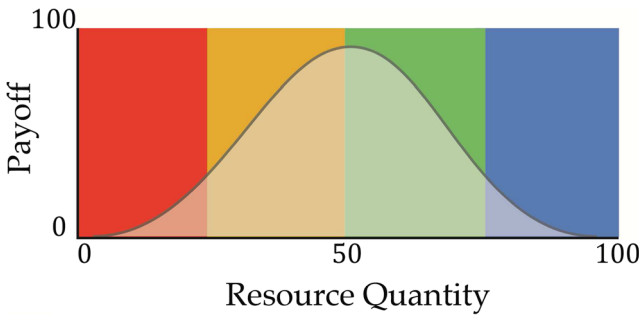


Fig. 2 A critical realist. The payoff function is approximately Gaussian. Any resource quantity between 75 and 100 maps to blue

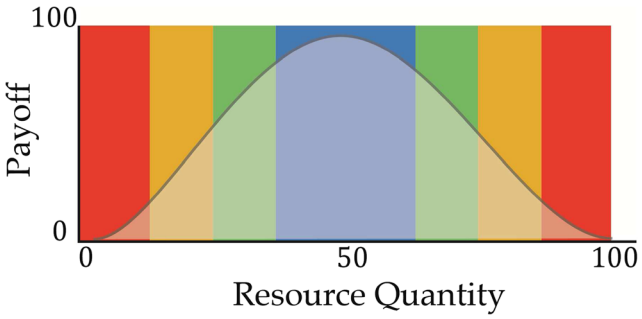


Fig. 3 An interface strategy. The resource quantities with the highest payoffs map to blue, and those with the lowest payoffs to red

because (1) monotonic functions are a (unbiased) measure zero subset of the possible payoff functions, and (2) even if they weren't, the ubiquitous biological need for homeostasis militates against them. Thus, we cannot expect, in general, that natural selection has tuned our perceptions to truth, i.e., we cannot expect our perceptions to be veridical. This is perhaps shocking news to perceptual scientists who assume that "much of human perception is veridical under natural conditions" and that "veridicality is an essential characteristic of perception and cognition."

Is it possible that this result—viz., that natural selection generically drives veridical perceptions to extinction—is an artifact of unrealistic assumptions in evolutionary game theory itself? In particular, is it an artifact of the assumptions of infinite populations and complete mixing?

Possibly, but unlikely, because the core reason that interface strategies dominate realist strategies is that, when payoffs are not monotonic with truths, interface strategies can be tuned entirely to the right information whereas realist strategies are necessarily tuned to the wrong information. Such mistuning will continue to cripple realist strategies even if the populations are finite and even if complete mixing is replaced with plausible networks of interactions. This claim should of course be checked using, e.g., simulations based on evolutionary graph theory.

What might prove interesting are spatial games, in which players interact with only nearest neighbors on a 2D grid. In this case it might be possible for groups of individuals having realist strategies to survive together in small enclaves.

Can this result be dismissed as an artifact of the overly simplistic and high-level example of perception that was studied? Could more realistic examples, say of shape perception or color perception, favor realist strategies? Again, it's possible, but unlikely. Monte Carlo simulations indicate that greater complexity does not, in general, favor realist strategies. Instead the mistuning of realists exacts a greater toll as the complexity of the situation increases. Indeed, with increasing complexity the need for simplification and abstraction is only accentuated.

Genetic algorithms

With evolutionary games we find that veridical perceptions fare poorly against interface perceptions when both are on the same playing field. But there is a prior question to be asked: Will veridical perceptions even get on the playing field? Or are they so unfit that evolution is likely to pass them over completely?

To study this question we turn to genetic algorithms, which are search heuristics based on features of natural evolution in sexually reproducing species, features such as mutation, inheritance, selection, and crossover (Hoffman et al. 2013; Mark 2013; Mitchell 1998; Poli et al. 2008).

The genetic algorithms we explore are variants of one introduced by Mitchell (1998) that evolves, over many generations, a robot named Robby who can efficiently gather soda cans that are randomly distributed on a 10×10 grid of squares. Surrounding the grid is a wall, which we can model as a perimeter of squares. Thus the world, call it W, that Robby inhabits can be represented as a 12 × 12 grid of squares. We denote the state of square (i, j) by $W(i, j)$ and stipulate that its value is 0 if the square has no cans, 1 if it has one can, and 2 if it is a wall. Because the wall is fixed, and the state of each square of the inner 10 × 10 grid is either 0 or 1, the possible states of W are $2^{10\times10} = 2^{100}$.

Robby can only see the state of the square he occupies and of the four immediately adjacent states. For instance, if Robby is at location (i, j) then he sees the world states $(W(i,j), W(i,j+1), W(i,j-1), W(i+1,j), W(i-1,j))$. Because there are at most three states at each of these five locations, the space of Robby's possible perceptions, call it X, is no larger than $3^5 = 243$; in fact it's a little smaller because, for instance, if the square to Robby's right is a wall, then the square to his left is not. Robby does not know which square (i, j) he is in, or even that he is in a 12 × 12 grid; he only sees the states of the squares, but not the locations or structure of the squares.

The goal of the genetic algorithm is to evolve a version of Robby that efficiently gathers soda cans, despite his ignorance of the structure of the grid. To this end, Robby has a set, call it G, of seven primitive actions he can take: stay where he is, pick up a can, step north, step south, step east, step west, or step randomly. What must be learned phylogenetically (i.e., over many generations of the genetic algorithm) is a foraging strategy that specifies which of the seven actions in G to take for each possible perception of the roughly 240 in X that Robby can have. The set of possible foraging strategies is thus approximately of cardinality $7^{243} \approx 2.3 \times 10^{205}$, a large search space in which to evolve good strategies.

The payoff function that provides the selection pressures for Robby's evolution is as follows: Robby gets 10 points for each can he picks up, but loses 1 point each time he tries to pick up a can where there is none, and loses 5 points each time he tries to walk into a wall.

There are roughly 240 "genes" that the genetic algorithm evolves, each having seven possible values, corresponding to the seven actions that can be taken in response to each potential perceptual state. Mitchell starts the genetic algorithm with an initial generation of 200 robots, each having randomly chosen values for each gene. Each robot is forced to forage through 100 randomly chosen worlds, taking 200 actions in each world. The fitness of a robot is the average number of points it collects over its 100 foraging runs. The fitter robots are preferentially chosen to be parents for the next generation. The genes for two parents are randomly split into two parts,

and the parts swapped to create two new genomes. A small amount of mutation is applied. In this way a new generation of 200 robots is created, and their fitness again measured by their success at foraging. This process is repeated for 1000 generations.

The first generation is comically stupid, bumping into walls, grabbing for cans in empty squares, perseverating in obvious mistakes. But the last generation sports expert foragers, racking up impressive point totals with surprising cleverness and methodical efficiency.

In this genetic algorithm, only the foraging strategies evolve, while the perceptual strategy remains fixed. All robots are naïve realists, seeing the true state of the world in their immediate vicinity. To take the next step, to study the coevolution of foraging and perceptual strategies, Mark (2013) modifies Mitchell's genetic algorithm in several ways. He allows each square to have up to 10 cans, and stipulates the following payoff function: (0,1,3,6,9,10,9,6,3,1,0). For instance, a robot gets 6 points for grabbing the cans in a square having 3 or 7 cans, and 0 points for a square having 0 or 10 cans. However, each robot cannot see the exact number of cans in each square, but instead sees just two colors, *red* and *green*. Each robot thus has a perceptual strategy, namely a mapping that assigns the percept *red* or *green* to each of the 11 possible numbers of cans. Perhaps, for instance, it sees *red* if a square has 0 cans and *green* otherwise. There are 2^{11}=2048 possible perceptual strategies. To allow for perceptual strategies to coevolve with foraging strategies, each robot has 11 more genes in its genome, which code the color that the robot sees for each quantity of cans. In the first generation the assignment of colors to the 11 genes is random.

In Mark's genetic algorithm, the first generation is again comically stupid. But after 500 generations there are again many skilled foragers, and all wield one of two perceptual strategies. In the first, squares are seen as *red* if they contain 0, 1, 9, or 10 cans, and as *green* otherwise. In the second, it is the reverse, with squares seen as *green* if they contain 0, 1, 9, or 10 cans, and as *red* otherwise.

These robots have evolved a strict interface strategy, tuned to payoffs rather than truths. A strategy tuned to truths would see squares having between 0 and 5 cans as, say, *red* and squares having between 6 and 10 cans as *green*, so that the perceived color would be as informative as possible about the true number of cans. But such a realist strategy would provide no information about payoffs (since *red* squares would have the same expected payoff as *green* squares) and would thus fail to guide effective foraging.

Instead, the robots wield a strategy that sees high payoff squares as *green*, and low payoff squares as *red*, or vice versa. This perceptual strategy provides the information required for effective foraging and is favored by the genetic algorithm. Given that Mark's simulation explored a space containing only 2048 perceptual strategies, it's likely that realist strategies were randomly tried and discarded during the 500 generations of evolution. But in a slightly more complex case, say where there are 30 genes and 10 possible colors, then the search space has 10^{30} possible perceptual strategies, and it's likely that a realist strategy, with no selection pressures in its favor, would never appear in any generation, because it could only appear by chance. Thus, it's likely that veridical strategies never enter the playing field. They're so unfit that they're not worth trying.

Interface theory of perception

Studies of perceptual evolution using evolutionary games and genetic algorithms render a clear verdict: Natural selection discards veridical perceptions and promotes interface strategies tuned to fitness. This motivates the interface theory of perception, which we now discuss (Fields 2014; Hoffman 1998; 2009; 2011; 2012; 2013; Hoffman and Prakash 2014; Hoffman and Singh 2012; Hoffman et al. 2013; Koenderink 2011; 2013; Mark et al. 2010; Mausfeld 2002; Singh and Hoffman 2013; see also von Uexküll (1909; 1926; 1957) for his related idea of an Umwelt).

Informally, the interface theory of perception says that the relationship between our perceptions and reality is analogous to the relationship between a desktop interface and a computer.

A desktop interface makes it easy to use the computer. To delete or copy files, for instance, one simply needs to drag icons around on the desktop.

But a desktop interface does not make it easy to know the true structure of a computer—its transistors, circuits, voltages, magnetic fields, firmware, and software. Indeed, it's in part by hiding this complex structure that the desktop makes it easier to use the computer. Why? Because if you were forced to be aware of the true facts about circuits, voltages, and magnetic fields, when your goal was simply to edit a photo or write a paper, you would be wasting time, memory, and energy on truths of no relevance to accomplishing your goal.

In similar fashion, says the interface theory of perception, our perceptions have been shaped by natural selection to make it easier for us to act effectively in the world, so that we can survive and reproduce (or, more accurately, so that our genes can survive and reproduce). Our perceptions have not been shaped to make it easy to know the true structure of the world but instead to hide its complexity.

Our perception of space-time is analogous to the desktop, and our perception of objects and their properties is analogous to the icons on the desktop. Just as the language of desktops and icons is the wrong language for describing the true structure of the computer, so also the language of space-time and physical objects is the wrong language for describing the true structure of the objective world.

A blue and rectangular icon on a desktop does not represent that something in the computer is blue and rectangular. Not because the icon is false or misleading or illusory, but because the icon is there to help you use the computer, not to distract you with irrelevant details about its innards.

This might seem odd. We're claiming that our normal perceptions are not veridical and yet not illusory. Isn't this self-refuting? After all, the standard definition of illusory perceptions is that they are perceptions that are not veridical.

The standard definition of illusory perceptions is, however, based on an incorrect understanding of perception and its evolution. It assumes that evolution has shaped our perceptions to be, in the normal case, veridical. But evolution has done no such thing. Instead, it has shaped our perceptions to be, in the normal case, adequate guides for adaptive behaviors. No perceptions are veridical. But it would be wrong to conclude that therefore all perceptions are illusory. They are not. They usually guide our behaviors quite well. It is only when we misunderstand the evolution of perception that we identify illusory perceptions with nonveridical perceptions.

For instance, when one sees a long, brown rattlesnake, this perception does not mean that something in the objective world is long and brown. Not because the perception is misleading or illusory but because the snake perception is there to adaptively guide your behavior, not to distract you with irrelevant details about the true structure of the world.

There is an obvious rejoinder: "If you think that snake is just an icon, why don't you pick it up? You'll soon learn that the snake is not just an icon, it's part of objective reality, and reality bites."

Of course, I won't pick up the snake. For the same reason I wouldn't carelessly drag a blue rectangular icon to the trash. Not because I take the file icon literally—the file isn't blue and rectangular. But I do take the icon seriously. If I drag the icon to the trash, I could lose many hours of work.

And that is the point. Natural selection has shaped our perceptions in ways that help us survive. We had better take our perceptions seriously. If you see a snake, don't grab it. If you see a cliff, avoid it. But taking our perceptions seriously doesn't entail that we must take them literally. To think otherwise, to think that "I must take my snake perception seriously" entails "I must take my snake perception to be literally true of the objective world," is an elementary error of logic but one that seems to enjoy a strong grip on the human mind, even the brightest of minds. Samuel Johnson, for instance, famously claimed to refute the idealism of George Berkeley by kicking a stone and exclaiming, "I refute it *thus*" (Boswell 1791). Kicking a stone can hurt; one must take the stone seriously, or risk injury. From this Johnson concludes, against Berkeley, that one must take the stone literally. Berkeleyian idealism may be false, but Johnson's argument against it is based on a logical fallacy.

We must take our perceptions seriously not because they reveal the true structure of the world, but because they are tuned, by natural selection, to fitness. The distinction between fitness and truth is elementary, and central to evolutionary theory. Fitness is a function of the objective world. However, a fitness function depends not just on the objective world but also on the organism, its state, and an action. For a hungry fly, a pile of dung conveys substantial fitness. For a hungry human, the same pile conveys no fitness.

Fitness is, in general, a complicated function of the objective world that depends on an organism, its state, and its action. There's no simple relation between fitness and truth, although many perceptual researchers assume otherwise. Geisler and Diehl (2002), for instance, assert "In general, (perceptual) estimates that are nearer the truth have greater utility than those that are wide of the mark." This would be convenient, but unfortunately it's not true. Fitness functions are more complex and versatile than that and rarely track truth.

Formally, the interface theory of perception proposes that the perceptual strategies of *H. sapiens* and, indeed, of all organisms are, generically, strict interface strategies. Recall that this means, in the dispersion-free case, that the perceptual function, $P{:}W{\rightarrow}X$, that maps states of the external world W onto perceptual experiences X, is not veridical in the following two senses. First, X is not a subset of W, so that none of our perceptual experiences are literally true of the world. Second, P is not a homomorphism of any structures intrinsic to W, other than the event structure required for probability, so that no structural relationships among our perceptions are literally true of the world.

The interface theory of perception certainly runs counter to our normal intuitions about the relationship between our perceptions and reality. It runs counter, for instance, to what Bertrand Russell (1912) took to be obvious: "If, as science and common sense assume, there is one public all-embracing physical space in which physical objects are, the relative positions of physical objects in physical space must more or less correspond to the relative positions of sense-data in our private spaces. There is no difficulty in supposing this to be the case."

The interface theory of perception is counterintuitive, but it can be seen as a natural next step along an interesting path of the intellectual history of *H. sapiens*.

The pre-Socratic Greeks, and other ancient cultures, believed that the world is flat, in large part because it looks that way. Pythagoras, Parmenides, and Aristotle, and soon many others, came to believe that our perceptions are misleading here and that the earth is in fact spherical. But they still believed that the earth is the center of the universe, because it certainly looks like the earth doesn't move and that the sun, moon, stars, and planets orbit around it. Kepler and Copernicus discovered that, once again, our perceptions have misled us, and the geocentric theory is false. This was difficult to accept. Galileo was forced to recant, and Giordano Bruno

was burned at the stake. Eventually we accepted the counterintuitive fact that, in this specific case, reality differs from our perceptions and the earth is not the center of the universe.

The interface theory of perception takes the next step. It says that reality differs from our perceptions not just in this or that specific case but in a far more fundamental way: our perception of physical objects in space-time no more reflects reality than does our perception of a flat and stationary earth. The space-time and physical objects of our perceptions are a species-specific adaptation, shaped by natural selection, which allow *H. sapiens* to survive long enough to reproduce. They are not an insight into the nature of objective reality. Quite simply, perception is about having kids, not seeing truth.

The argument for the interface theory is not an inference based on epistemological assumptions—that we can only be sure of our perceptions and so, for all we know, the world differs dramatically from our perceptions. Nor is it an argument for idealism—that to be is to be perceived, and that something not perceived by my mind exists only if perceived by another (Berkeley 1710/2012; 1713/1979).

Instead, the argument is that evolution by natural selection, one of the best-confirmed theories of contemporary science, applies not just to bodily traits but also to perceptual and cognitive traits. This entails that, for a perceptual strategy, the ticket to the next generation, indeed the only ticket other than dumb luck, is reproductive success. Reproductive success and veridicality are entirely distinct concepts. Whenever they diverge, reproductive success trumps veridicality. They diverge if the relevant payoff functions are nonmonotonic; indeed they diverge with unbiased probability one. Thus, it is almost certain that our perceptions have not been shaped to be veridical.

It is no surprise, then, that evolution has shaped beetles that are fooled by bottles, dragonflies that mistake gravestones for water, gull chicks that prefer red disks on cardboard to their real mothers, frogs that die of starvation when surrounded by mounds of unmoving edible flies, and birds that prefer brightly speckled rocks or the eggs of cowbirds to their own eggs. These are not shocking outliers but exactly what one expects from a careful understanding of evolutionary theory. The reason it seems counterintuitive that our own perceptions of space-time and objects are not veridical is that we are blind to our own blindness—we cannot step outside our perceptions and look back to make the shocking discovery that they are just a satisficing interface, not an insight into truth. For that discovery, for the realization that once again there is no fundamental divide between us and other animals, we need the aid of the mirror view provided by the theory of evolution. In retrospect, we should have expected all along what that mirror reveals.

Steven Pinker (1997) portrays well the view from the mirror: "We are organisms, not angels, and our minds are organs, not pipelines to the truth. Our minds evolved by natural selection to solve problems that were life-and-death matters to our ancestors, not to commune with correctness."

Robert Trivers (1976/2006, p. xx; also 2011) has peered into the mirror and seen the same view: "If deceit is fundamental to animal communication, then there must be strong selection to spot deception and this ought, in turn, to select for a degree of self-deception, rendering some facts and motives unconscious so as not to betray—by the subtle signs of self-knowledge—the deception being practiced. Thus, *the conventional view that natural selection favors nervous systems which produce ever more accurate images of the world must be a very naïve view of mental evolution*." (emphasis ours).

The standard Bayesian framework for vision

The standard contemporary framework for vision research takes vision to be fundamentally an inductive problem of inferring true properties of the objective world: any image on the retina is consistent with many different scene interpretations; that is to say, the same image could in principle have been generated by many (usually infinitely many) distinct 3D scenes. This raises the natural question of how the visual system converges upon a single interpretation, or small number of interpretations. The fundamental ambiguity inherent in perception can be resolved only by bringing to bear additional biases or constraints, e.g., concerning how probable different scene interpretations are *a priori*. The environment in which our species evolved is a highly structured place, containing many regularities. Light tends to come from overhead, there is a prevalence of symmetric structures, objects tend to be compact and composed of parts that are largely convex, and so on. Over the course of evolution, such regularities have been internalized by the visual system (Feldman, 2013; Geisler, 2008; Shepard, 1994). Thus, they help to define probabilistic biases that make some interpretations of an image much more probable than others.

Formally, given an image input y_0, the visual system must compute and compare the *posterior probabilities* $p(x|y_0)$ for candidate scene interpretations x. By Bayes' Theorem, this posterior probability is proportional to the product of the likelihood of the scene x, $p(y_0|x)$, and its prior probability, $p(x)$:

$$p\left(x\middle|y_0\right) \propto p\left(y_0\middle|x\right)p(x)$$

The likelihood of the scene x corresponds to the probability of obtaining the image y_0 from scene x; it is therefore a measure of the extent to which scene x is consistent with—or "can explain"—image y_0. In vision applications, the likelihood $p(y_0|x)$ often is defined in terms of a projective or rendering map from 3D scenes to projected images (possibly with noise). The prior probability captures the visual system's implicit knowledge, based on phylogenetic and ontogenetic

experience, that certain scene interpretations are more probable *a priori* than others. This knowledge is "prior" in the sense that it exists in the system prior to obtaining the current image input. Given the fundamental ambiguity of perception noted above, the likelihood is often equally high for many different scene interpretations (i.e., many different 3D interpretations can in principle explain the given image). However, these scenes are not equally probable *a priori*. The product of the likelihood and prior—the posterior distribution over scenes, given the image y_0—thus strongly favors some scene interpretations over others (Kersten, Mamassian & Yuille, 2004; Knill & Richards, 1996; Mamassian, Landy & Maloney, 2002).

The scene interpretation with the highest posterior probability often is taken to be the "best" scene interpretation, given the image. More generally, however, the selection of a single "best" interpretation based on the posterior distribution requires the application of a loss function. A loss function defines the consequences of making errors, i.e., of making interpretations that deviate from the "true," although unknown, value of the relevant variable to different extents. Technically, the *maximum-a-posteriori* (or MAP) decision rule noted above follows if the loss is equally "bad" for all non-zero errors, and 0 when the error is effectively 0. A quadratic loss function—where loss increases as the square of error magnitude—leads to a decision rule that picks the mean of the posterior distribution as the single best interpretation (Mamassian et al. 2002). Other decision rules used in models of vision include using maximum-local-mass loss (Brainard & Freeman, 1997) and sampling from the posterior (or probability matching; e.g., Wozny, Beierholm, & Shams, 2010).

Figure 4a summarizes pictorially the standard Bayesian approach to vision. In this framework, space X corresponds to states of the world (generally taken to be "3D scenes"), and Y to the set of projected images. The likelihood map L corresponds to the projective, or rendering, map from 3D scenes to 2D images—possibly with noise.[1] Given a particular image y_0 in Y, the Bayesian posterior B defines a probability distribution on scene interpretations in X. The choice of a loss function then allows one to pick a single best interpretation based on this full posterior distribution on X.

[1] In other words, in the general case, given a scene x in X, $L(x)$ defines a probability distribution on Y, i.e., $L(x|y) = p(y|x)$. Note that L is in fact a function of two variables, x and y; but in different contexts one or the other of these is assumed fixed. In the current discussion (the "forward mapping"), we are considering a situation where x is fixed and we compute L for different values of y. Whereas, in the discussion above (the "inverse map"), we considered the situation where y_0 was fixed and we were comparing the likelihood for different scenes, x.

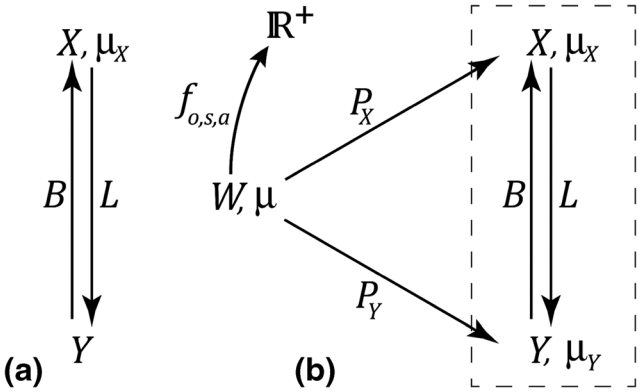


Fig. 4 **(a)** Standard Bayesian framework for vision. **(b)** Our CEP framework. In this framework, the interpretation space X in probabilistic inference is not assumed to be identical to the objective world W. Importantly, there is a fitness function f on W. The perceptual channels from W to the perceptual representational spaces X and Y are "tuned" to increase expected-fitness payout to the organism

Limitations of the standard Bayesian framework

Given the two probabilistic sources of information embodied in the likelihood and the prior, Bayes' Theorem provides a provably optimal way to combine them (Cox 1961; Jaynes, 2003). Hence, once a likelihood map and a prior distribution have been specified on a given space of possible interpretations (or scene *hypotheses*), there is principled justification for using Bayes to make perceptual inferences. However, as we clarify below, the standard Bayesian framework for vision makes certain key assumptions that make it much too limiting.

Note, in particular, that in the standard Bayesian framework summarized above, space X plays two distinct roles. First, X corresponds to the set of objective world states. Second, X corresponds to the space of interpretations (or hypotheses) from among which the visual system must select. In other words, in the standard Bayesian framework for vision, the observer's hypothesis space is implicitly assumed to be identical with the objective world. This dual role played by X is consistent with the conceptualization of vision as *inverse optics,* according to which the goal of vision is essentially to "undo" the effects of optical projection (Adelson & Pentland, 1996; Pizlo, 2001). It also is consistent with the historical roots of Bayesian methods as providing ways of estimating "inverse probability." Laplace (1774), for instance, considered the problem of estimating underlying causes C from an observed event E: What one would *like* to estimate is the probability $p(C|E)$ of a particular underlying cause C given observation E, but what one actually knows is the probability $p(E|C)$ of observing any particular event E given cause C. Bayes' Theorem, of course, provides a means of inverting these conditional probabilities.

The dual role played by space X clarifies the way in which the standard Bayesian framework for vision embodies the assumption that the human visual system (and perception more

generally) has evolved to perceive veridically. Clearly, it is *not* the case that a Bayesian observer always makes veridical inferences. Given the inherently inductive nature of the problem, that would be impossible. Specifically, because a Bayesian observer must rely on assumptions of statistical regularities in the world (e.g., light tends to come from overhead), it will necessarily make the *wrong* scene interpretation whenever it is placed in a context where its assumptions happen to be violated (say in a scene where light happens to come from below; e.g., Kleffner & Ramachandran, 1992).

There is a more fundamental sense, however, in which the standard Bayesian framework assumes veridicality: it assumes that the hypothesis space X—the observer's representational space, which contains the possible scene interpretations from which it must select—corresponds to objective (i.e., observer-independent) reality. In other words, it assumes that the observer's representational language of scene interpretations X is the correct language for describing objective reality. Even if the observer's estimate might happen to miss the "correct" interpretation in any given instance, the assumption is nevertheless that the representation space X contains somewhere within it a true description of the world. It is in this more fundamental sense that the standard Bayesian framework embodies the assumption that vision has evolved to perceive veridically.

When viewed in light of our earlier discussion on possible relationships between X and W (recall the section on Perceptual Strategies), it becomes clear that the standard Bayesian framework for vision essentially assumes that $X = W$ (or that X is isomorphic to W). This is a strong assumption—essentially a form of naïve realism—that makes it impossible to truly investigate the relation that holds between perception and the objective world. A genuine investigation must begin with minimal assumptions about the form of this relation. This is especially true if one's goal is to have a mathematical model of the *evolution* of perceptual systems. Clearly, as perceptual systems evolve, their representational spaces can change, as can the mapping from the world W to a given representational space X. Thus a framework that simply assumes that $X = W$, or X is isomorphic to W, will (by definition) be unable to capture this evolution.

Consideration of perceptual systems in simpler organisms makes the simplistic nature of this assumption especially clear. As mentioned in the *Introduction*, in discussing simpler visual systems such as those of the fly and the frog, Marr (1982) noted that they "…serve adequately and with speed and precision the needs of their owners, but they are not very complicated; very little objective information about the world is obtained. The information is all very subjective.…" He clarified what he meant by *subjective* by adding that "…it is extremely unlikely that the fly has any explicit representation of the visual world around him—no true conception of a surface, for example, but just a few triggers and some specifically fly-centered parameters…" (p. 34). Marr was acknowledging that visual systems that do not compute objective properties of the world can nevertheless serve their owners well enough for them to survive. This should not be surprising. Clearly, what matters in evolution is fitness, not objective truth; and even perceptual systems that compute only simple, "subjective" properties can confer sufficient fitness for an organism to survive—even thrive.

As also noted in the *Introduction*, when it comes to human vision, Marr held a different position (as do most modern vision scientists). He believed that the properties computed by human vision are, or correspond to, observer-independent properties of the objective world. Such a sharp dichotomy between the visual systems of "simpler" organisms on the one hand, and human vision on the other, seems implausible. After all, the evolution of *Homo sapiens* was governed by the same laws that govern the evolution of other species. Nor is it viable to assume that evolution is a "ladder of progress" that leads perceptual systems to compute incrementally more and more objective properties of the world. So what justification do we have to believe that the representational spaces employed by human perceptual systems correspond to objective reality?

As always, what matters in evolution is fitness, not objective truth. One must examine the role that fitness plays. As we noted in the *Interface Theory of Perception* section, the first thing to note about fitness is that it depends not only on the objective state of the world but also on the organism in question (e.g., frog vs. tiger), the state of the organism (e.g., starving vs. satiated), and the type of action in question (e.g., mating vs. eating).[2] Thus, one's formal framework must be broad enough to include the possibility that the representations computed by human vision also do not capture objective truth (in the more fundamental sense noted above—namely, that the interpretation space X does not contain anywhere within it a true description of the objective world). Moreover, if such a framework is to be sufficiently general to model the evolution of perceptual systems, it must clearly allow for different possible relations between X and W.

Computational Evolutionary Perception

We generalize the standard Bayesian approach to a new framework that we call Computational Evolutionary Perception or CEP (Hoffman & Singh, 2012; Singh & Hoffman, 2013). Given the intrinsically inductive nature of perception, CEP incorporates probabilistic inference in a

[2] For this reason, evolutionary fitness cannot be incorporated into the standard Bayesian framework for vision simply by identifying it with the *loss function* of Bayesian Decision Theory.

fundamental way. Importantly, it places the objective world W outside of the Bayesian inferential apparatus (Fig. 4b). In CEP, X and Y are simply two representational spaces—neither is assumed to correspond (or be isomorphic) to W. In any given context, Y may be a lower-level visual representational space (say a representation of some 2D image structure), and X may be a higher-level representation (say one that involves some 3D structure). The more complex representation X may, for example, have evolved subsequently to the simpler representation Y; however, there is no assumption in our framework that $X = W$, i.e., that X contains somewhere within it a true description of the state of the objective world. Nor do we assume that X is in any objective sense "closer" to W than is Y. X is simply a representational space that has evolved, presumably because it has some adaptive value for the organism within its ecological niche. We cannot assume that the properties of X correspond to properties of the objective world W. In other words, if we find some structure on X, it does not follow from this that W necessarily has that structure as well.[3] (For a proof of this, see the Measured World section, where we state and prove an *Invention of Symmetry Theorem*).

For each representational space, say X, there is a perceptual channel from W to that representational space, i.e., $P_X: W \rightarrow X$. We previously used the term *perceptual strategy* to refer to such channels (see Definitions 1 and 2). These channels define the correspondence between the objective world W and the representational spaces (X and Y). Recall that, in the general case (i.e., with dispersion), these perceptual channels are Markovian kernels (see Definition 2). That is, for each w in W, P_X specifies a probability measure on X, and P_Y specifies a probability measure on Y. Recall also, that we make no assumptions of any structure on W, except for probability structure—namely, that it is meaningful to talk of probabilities on W. Specifically, we assume there is a space of events $\mathcal{W}$ on W, and a probability measure μ on this space. This probability measure μ on W induces, via the channel P_X , a so-called "pushdown" measure μ_X on X, and similarly it induces, via P_Y, a probability measure μ_Y on Y. One implication of this is that the *prior probability distribution* on X, used in making Bayesian inferences from Y to X, is *not* the "world prior" in our framework, but rather its *pushdown* probability measure—via the perceptual channel P_X—onto the representational space X.

Fitness of course plays a fundamental role in CEP, with the high-level idea being that evolution "tunes" the perceptual channel P_X (including the representational space X itself) so as to increase the expected-fitness payout to the organism. In other words, fitness is the key signal that the perceptual channels are "tuned" to communicate. In order to bring fitness into the framework, we view organisms as gathering "fitness points" as they interact with the world. As noted earlier, fitness depends not only on the objective state of the world, but also on the organism in question, its state, and the type of action under consideration. We thus define a *global fitness function* $f: W \times O \times S \times A \rightarrow \mathbb{R}^+$ where O is the set of organisms, S their possible states, and A their possible action classes. Once we fix a particular organism o in O, its state s in S, and action class a in A, we have a *specific fitness function* $f_{o,\,s,\,a}: W \rightarrow \mathbb{R}^+$ that assigns fitness points (nonnegative real numbers) to each w in W (e.g., to a starving lion eating a gazelle).

The CEP framework thus differs from the standard Bayesian framework for vision in three key respects: (1) it separates the objective world W from the interpretation space X (used in the Bayesian inference from Y to X); (2) it introduces perceptual channels P_X and P_Y from W to the spaces X and Y, respectively; and (3) it introduces a fitness function on W (Fig. 4b). Fitness is in fact the key signal that the perceptual channels are tuned to communicate. Given a specific fitness function $f_{o,s,a}$, evolution shapes a source message about fitness and a channel to communicate that message, in such a way so as to hill-climb toward greater expected-fitness payout for the organism. Thus, the perceptual channel $P_X: W \rightarrow X$ can be expressed as the composition of two Markovian kernels: (1) a message *construction* kernel $P_{C_X} : W \rightarrow M$, where M is the set of messages, and (2) a message *transfer* kernel $P_{T_X} : M \rightarrow X$ that transmits the messages. The construction kernel is needed because the message to be transmitted depends not only on W and X, but also on the specific fitness function $f_{o,\,s,a}$. Thus, if we consider a different specific fitness function on W, the set of messages to be transmitted may be very different.

Consider again an interface game on a simple example of a "world" W involving a single variable that ranges from 0 to 100 (so that each value in this range is a particular "world state"; a similar example is discussed in the section on *Evolutionary Games*). Now consider a nonmonotonic fitness function on W with two peaks—a slight complication of the fitness function in Figs. 2 and 3—as shown in Fig. 5a. As is clear from this plot, world state values near 25 and 75 are associated with the most fitness, whereas values near 0, 50, and 100 are associated with the least fitness. Let's assume that we're given a representational space X, containing exactly 4 elements, $X = \{A, B, C, D\}$. If we want to construct an efficient perceptual channel for the above fitness function to this representational space, a natural way to proceed is to construct a message set, $M = \{B, G, Y, R\}$, and map values of W into M by clustering their fitness values into four classes. Specifically, world states in W with very high fitness values are mapped onto B ("blue"); those with somewhat high fitness values onto G ("green"), those with somewhat low values into Y ("yellow"), and those with very low values into R ("red"; Fig. 5b). In this case, the representation activated in X (based

[3] If X bears no simple relation to W, it is natural to ask: How is perception (and cognition) then grounded in the objective world? We address this question in the next section where we introduce the Perception-Decision-Action (or PDA) loop.

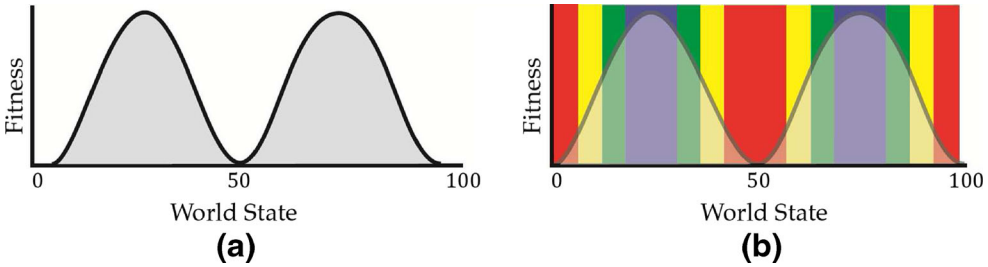


Fig. 5 **(a)** A nonmonotonic fitness function on a range of world states. **(b)** Constructing an efficient message for a representational space with four elements. World state values are mapped onto these four elements based on a clustering of their fitness values ("very high," "somewhat high," "somewhat low," and "very low"). The resulting channel is highly informative about expected-fitness payout but uninformative about objective world states

on the received message) will be highly informative about fitness. So if an organism has to choose between two world states based on the knowledge that one was *B* and the other *Y*, it would always be able to pick the world state with the higher fitness value. Note, by contrast, that this perceptual channel is poor at conveying the actual state of the world *W*. A message of *R,* for instance, could be indicative of a world state near 0, 50, or 100; there is no way to tell—similarly for the other possible received messages. For the same reason, this perceptual channel also will be poor at conveying information about a (different) fitness function that increased monotonically with world-state values.

The above example clarifies that the notion of "tuning" a perceptual channel depends critically on the specific fitness function. We propose the following general definitions:

Definition 8 Given a specific fitness function $f_{o,s,a}$, a *Darwinian ideal observer* consists of a representational space *X*, and a perceptual channel $P_X: W \rightarrow X$ that *maximizes* the expected-fitness payout to the organism.

We term such an observer *ideal*, because natural selection does not, in general, produce perceptual channels that *maximize* expected-fitness payout. It produces satisficing solutions, rather than optimizing solutions. This more typical satisficing-type of solution defines a *Darwinian observer*:

Definition 9 Given a specific fitness function $f_{o,s,a}$, a *Darwinian observer* consists of a representational space *X*, and a perceptual channel $P_X: W \rightarrow X$ that has been shaped by natural selection as a satisficing solution to the problem of increasing expected-fitness payout to the organism.[4]

[4] Of course, what counts as a satisficing solution, or "sufficient" increase in expected-fitness payout, depends on a number of factors, including the characteristics of the ecological niche of the organism, and of its competitors.

Evolution of perceptual channels and representations

While incorporating the role of probabilistic inference in a fundamental way, CEP generalizes the standard Bayesian framework for vision by: (1) allowing for different possible relationships between the world *W* and perceptual representations *X* (e.g., in evolving perceptual systems); (2) introducing fitness into the framework in a way that does not simply reduce it to the Bayesian loss function; and (iii) modeling the evolution of perceptual systems as hill-climbing towards greater expected-fitness payout for the organism. We next consider some different ways in which such hill-climbing can occur.

Evolution of perceptual channels

An obvious way to increase the expected-fitness payout is to "tweak" a perceptual channel P_X appropriately while keeping the representational space *X* fixed. A key component of such tweaking is the crafting of a set of messages *M*, and a message construction kernel $P_{C_X}: W \rightarrow M$ that is highly informative about the fitness function on *W*. As we saw in the example in Fig. 5 above, it is possible to have a perceptual channel (a composition of a message construction kernel and a transfer kernel) that is good at communicating information about fitness but bad at communicating information about truth, and vice versa.

Evolution of representational spaces

In the situation considered above, the representational space *X* remained fixed, only the channel P_X evolved. However, in biological evolution, it is clear that perceptual representations themselves evolve. If a representational space *X* has little relevant structure (say), even with the perceptual channel P_X tuned optimally (i.e., to maximize expected-fitness payout), the amount of information carried about expected fitness may still be quite limited. In such cases, there would be

evolutionary pressure to evolve the representational space X itself, rather than just the perceptual channel to the fixed representational space X. At one point in the course of its evolution, for example, an organism's visual system might represent only some rudimentary 2D image structure, whereas much later in its evolution, it may acquire representations that segment the perceptual world into objects, and represent some 3D structure. Note that this is a more dramatic change—one that alters the qualitative format of a representation—compared with a situation where a parameter value within a fixed representational space (such as the peak of a spectral sensitivity function) is tweaked by evolutionary pressures.

In the CEP diagram above (Fig. 4b), we considered representational spaces X and Y. These are of course just two of many such possible representations. In studying the evolution of representations, one must consider evolutionary sequences of perceptual representations, $\langle X_1, P_{X_1}\rangle \rightarrow \langle X_2, P_{X_2}\rangle \rightarrow \langle X_3, P_{X_3}\rangle \rightarrow \ldots$.[5] It is then natural to consider whether, and under what conditions, such a sequence might converge to the objective world structure. Given our arguments so far, and our review of results with evolutionary games (see the sections on Evolutionary Games and Genetic Algorithms), it seems unlikely that a sequence of perceptual representations resulting in monotonically increasing expected-fitness payout would generically result in monotonically increasing capacity to transmit the "truth" signal (i.e., information about objective world structure). The advantage of our formal framework is that it permits one to pose and address such questions in a mathematically precise manner.

Dedicated vs. general-purpose representations

Both possibilities considered above—evolving the perceptual channel P_X for a fixed X versus evolving X itself—assume a context where a specific fitness function is given, and the perceptual channel and/or representational space is tuned to increase the fitness-payout *for that specific fitness function*. Recall that a specific fitness function $f_{o,s,a}$ presupposes not only a particular organism o, but also a particular state s, and particular action class a. Because organisms engage, of course, in a wide variety of action classes, and each action class is associated with its own specific fitness function $f_{o,s,a}: W \rightarrow \mathbb{R}^+$, one must consider not just one but many such specific fitness functions. However—importantly—optimizing (in the ideal case) a perceptual representation and channel to maximize the expected-fitness payout for one specific fitness function does not guarantee that this representation and channel will be optimized for other fitness functions (associated with other action classes). This raises the problem of how best to tune the perceptions of an organism to a variety of different fitness functions.

There are, broadly speaking, two ways in which the above problem can be addressed—both of which evolution seems to have employed. The first is to evolve distinct perceptual representations that are dedicated to different types of tasks or actions. In this case, each dedicated representation/channel allows for high expected-fitness payout for the specific fitness function associated with a particular action class. When considering a different action class, a different representation/channel would be dedicated to communicating information about its expected-fitness signal. Although there is some evidence of such dedicated representations in the evolution of vision (e.g., dorsal versus ventral pathways in the primate cortex), adopting this strategy indiscriminately can lead to quickly proliferating representational spaces—which would quickly become untenable.

At the other end of the spectrum, one can imagine a single general-purpose representation/channel being "tuned" to increase simultaneously expected-fitness payout for a large number of specific fitness functions (associated with different action classes). In this case, it is unlikely that the perceptual channel can be tuned optimally for all of those specific fitness functions. However, if the specific fitness functions are sufficiently similar, it is certainly possible that the general-purpose channel can increase expected-fitness payout sufficiently to make this strategy feasible—especially because doing so avoids the "costs" associated with producing multiple representational spaces.

Although neither strategy is feasible in its extreme form (i.e., using a single strategy throughout), a compromise based on a mixture of the two strategies seems reasonable. Given a large number of specific fitness functions, group them into clusters based on their similarity. Now dedicate a different representational space and channel to each cluster. So, all specific functions within a cluster are subserved by a single representation/channel. This mixed-strategy allows the different representational spaces and channels to do a reasonable job of increasing expected-fitness payout for all specific fitness functions in a particular cluster, while keeping the total number of distinct representations relatively low.

Perception-Decision-Action (PDA) loop

The basic claim of Interface Theory is that our representational spaces need not be isomorphic or homomorphic to the objective world W (or to a subset of W). Hence when we observe

[5] We do not imply, of course, that there is single, linear evolutionary sequence of perceptual representations. Such sequences can branch out, with possibly multiple representations evolving more or less in parallel while emphasizing different aspects of the original representation (e.g., color, motion, and shape in the context of vision). We can nevertheless meaningfully consider ascending sequences within this larger lattice of evolving representations.

some structure in a representational space X (e.g., three dimensionality), we cannot simply infer from this observation that W must also have that same structure. However, this raises a natural question: In the absence of a homomorphic relation, how is it possible for perception to guide actions in the world?

Successful interaction with the world requires, at least, that an organism be able to predict how its perceptions will change when it acts. When we perceive the 3D shape of an object, we can predict—based on its various forms of perceived symmetries, for instance—what the object might look like if we were to pick it up and rotate it in a certain direction. Similarly, when we toss an object in a certain way, we can predict the trajectory, spin, and other behaviors we are likely to observe. Our success in interacting with the world in many different ways might suggest that our representations of the objective world W are veridical (Pizlo et al. 2014). In other words, it might suggest that our representations must include an accurate model of the objective world. How else could we account for such successful interactions?

Thus, a natural question that the Interface Theory must address is: If we can assume no simple (e.g., isomorphic or homomorphic) correspondence between our representations and the objective world, how can we explain our successful interactions with the world? We will flesh out an answer to this question in formal terms below, but in short the answer is that this is possible, because we do not simply passively view the world, but also act on it, and moreover we perceive the consequences of those actions. In other words, it is possible to interact with a fundamentally unknown world if (1) there are stable perceptual channels; (2) there is a regularity in the consequences of our actions in the objective world; and (3) these perceptions and actions are coherently linked. Although the role of action is emphasized in sensorimotor and enactive approaches to perception as well (Noe 2006; Chemero 2009), our position differs in a crucial respect. In our view, having a perceptual experience does not require motor movements. Our claim is rather that, over the course of evolution, perceptual-motor interactions have played an important role in shaping perceptual mechanisms.

To return to the metaphor of the desktop interface on a PC, even though visible characteristics of the file icons (their shape, color, etc.) do not reflect their objective properties (the computer files themselves are not inherently shaped or colored), the interface nevertheless allows us to interact successfully with the computer because of the coherence between the "perceptual" and "action" mappings. By its very design, the desktop interface allows us to interact successfully with the computer even if we are fundamentally ignorant of its objective nature. Similarly, the claim of Interface Theory is that perceptual properties of space-time and objects simply reflect characteristics of our perceptual interface; they do not correspond to objective truth. They are simply perceptual representations that have been shaped by natural selection to guide adaptive behavior.

Action plays a crucial role in the evolution of perceptual representations because fitness, to which perceptual channels are tuned, depends on the actions of an organism. Recall that specific fitness functions depend not just on the organism and its state, but also on the action class under consideration. Thus different action classes correspond to different expected fitnesses. Because perceptual channels are tuned to efficiently communicate information about expected fitness, one can expect a coupling between the evolution of perceptual channels/representations and the actions they inform.

Recall that, in the CEP framework, we have representational spaces (X, Y) and perceptual channels (P_X, P_Y) from the world W to these representational spaces. Let us focus on one of these representational spaces, say X. (We can therefore drop the subscript X from the perceptual channel for the remainder of this section.) To introduce action into the framework, we add a space G of possible actions, as shown in Fig. 6. (G may have, as a subset, a group that acts on the world W, but even so, the action of this group on the world may not technically be a *group* action. See Appendix.)

Given a perception x in X, the perceptual system must decide which action g to take (including the possibility of taking no action). Once an action g has been selected, the observer must then act on the world W: if the action g is deterministic (as in, e.g., a group action), then the previous state w of the world is moved to a new state w' denoted $g.w$. However, in general we want to allow the possibility that the action on the world is stochastic, so we think of g acting via a Markovian kernel A, called the *action kernel*, as follows. For each $g \in G$ and $w \in W$, $A(g,\mathrm{d}w)$ defines a probability distribution on states of W.

As a result, we have three Markovian kernels—for perception, decision, and action respectively. P is a kernel from W to X, D is a kernel from X to G, and, given the previous state w of the world, A is a kernel from G back to W (strictly speaking, A is a kernel from $G \times W$ to W; similarly, mutatis mutandis, for the other kernels). These three kernels therefore form a loop that we call the PDA loop. Because, in our framework, the

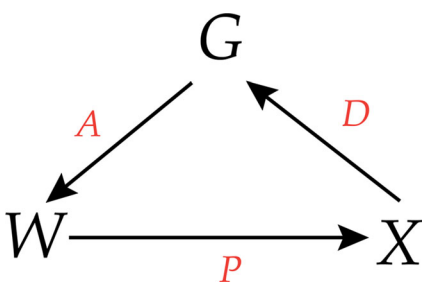


Fig. 6 Perception-Decision-Action, or PDA, loop. Note that we have added a space of possible actions G to the CEP framework. This now yields three Markovian kernels: the perception channel P from W to X, the decision kernel D from X to G, and the action kernel A from G back to W

observer does not know *W*, it cannot know the perceptual channel *P* (from *W* to *X*), nor the action kernel *A* (from *G* to *W*). In other words, just as the observer does not know the true source of its perceptions in the objective world, similarly it does not know the true effects that its actions are having in the objective world. Importantly, however, the observer does know the perceptual consequences of those effects, i.e., the results those effects have, via perceptual channel *P*, back in its perceptual representation *X*.

In other words, even though the observer cannot know kernels *P* and *A* individually, it *can* know the composition kernel *AP* from actions in *G* to perceptual representations in *X*. Similarly, it *can* know the composition kernel *DAP* from perceptual representations *X* back to *X*. This is what allows the observer to interact with *W*, even though it is in a fundamental sense ignorant of it. By trying various actions, and observing their perceptual consequences, it can tweak its decision kernel (the one that picks actions) so that the resulting perceptual consequences of its actions more consistently enhance fitness; note that this logic applies both phylogenetically and ontogenetically.

We should note that the PDA formalism just described applies not just to humans, but also to all organisms. Moreover, a given organism can have many PDA loops, and its PDA loops can be nested and networked in an endless variety of ways. Thus, the PDA formalism provides a powerful abstract framework for cognitive modeling (Hoffman and Prakash 2014; Singh and Hoffman 2013).

Measured world

Despite the evidence from evolutionary games and genetic algorithms that militates against veridical perceptions, a hard-nosed critique might still be unfazed: "Look, it's still the case that what you see is what you get. If it looks to me that a rock is round and 5 feet away, I can verify this with rulers, laser rangefinders and a host of other instruments, and then confirm it with other observers endowed with similar instruments. So my perceptions are in fact veridical."

This argument is *prima facie* plausible and has two key parts. The first part, the *measured world* argument, claims that our perceptions of the world are veridical because they generally agree with our careful measurements of the world. The second part, the *consensus* argument, claims that our perceptions are veridical because human observers normally agree with each other about their perceptions and the results of their measurements.

Both arguments fail. One problem with the measured world argument is that there are obvious cases where our perceptions radically disagree with our careful measurements. The sun, moon and stars, for instance, all look far away, but they all look about equally far away. Nothing in our perceptions prepares us to expect that the sun is almost 400 times further away than the moon, or that the nearest star, Proxima Centauri, is more than 250,000 times further away than the sun. Even at close distances our perceptions differ from our careful measurements (Kappers 1999; Cuijpers et al. 2003; Koenderink et al. 2010; Pont et al. 2012), leading Koenderink (2014) to conclude "The very notion of veridicality itself, so often invoked in vision studies, is void" and "It is a major obstacle on the road to the understanding of perception."

A second problem with the measured world argument arises even if the results of measurements *agree* with our perceptions. We express our measurements in terms of predicates that our perceptual representations use. For example, we arrive at a notion of Euclidean space (in Newtonian physics) by extending our perceptual representations using symmetry assumptions, such as translation and rotation invariance. In this sense, our measured world is simply an extension of our perceptual representations. A measurement of depth, for instance, like our perception of depth, is described using spatial predicates (e.g., using centimeters or relative distances). It is these very predicates themselves that, according to the results of the evolutionary games, have no correspondence with objective reality. Natural selection instead favors predicates tuned to fitness functions.

The consensus argument fares no better, for the simple reason that agreement among observers does not entail the veridicality of their perceptions or measurements. Agreement can occur if, for instance, the perceptions and measurements of observers are all nonveridical in the same way. Indeed, if the interface theory of perception is correct, and natural selection has shaped *H. sapiens* to share a nonveridical interface, then that is precisely why we agree. But that entails nothing about reality. All flies agree: dung tastes great. We might beg to differ.

Another argument that our perceptions match the measured world is given by Bertrand Russell (1912): "If a regiment of men are marching along a road, the shape of the regiment will look different from different points of view, but the men will appear arranged in the same order from all points of view. Hence, we regard the order as true also in physical space, whereas the shape is only supposed to correspond to the physical space so far as is required for the preservation of the order." The idea is that certain aspects of our perceptions are invariant under changes in viewpoint, and this entails the veridicality of these aspects.

This argument also fails, but the reason is deeper. Our perceptions of space and time can be extended systematically using symmetry groups, e.g., Euclidean, Galilean, Lorentz, Poincare and supersymmetry groups (Cornwell 1997). The measured worlds that result share the same predicates of space and time as our perceptions but don't suffer the same myopia: the Euclidean extension, for instance, easily handles the huge difference in distance between the moon and the stars.

Changes in an observer's viewpoint or frame of reference can then be modeled by actions of these groups (e.g., translations and rotations) on the appropriately extended space-time. Russell claims that if a feature of our perceptions is invariant under these group actions, then it can be taken as veridical.

This claim is false. The following theorem shows that the world itself may not share any of the symmetries that the observer observes. The world need not have the structure the observer perceives, no matter how complex that structure is and no matter how predictably and systematically that structure transforms as the observer acts.

Invention of Symmetry Theorem. Let an observer have at its disposal a group G of actions on the world W, such that its own perceptual space X is a G-set. This means that G *acts* on X via the kernel $PA = P(A(g)) = \int P(w,dx)A(g,dw)$, i.e., the action of A followed by that of P; moreover G *acts* on X by a *transitive group action*, so that G is a *symmetry group* of X (see Appendix). Moreover, let G act on W in such a way that the observer's perceptual channel mediates this action: $P(g.w) = g.P(w)$, where the dot signifies the action of G on each set. Then, the perceptual experiences X of this observer will admit a structure with G as its group of symmetries.

Proof Let S_x be the *fiber* of P over $x \in X$. (The points, $w, w' \in W$, are in the same fiber if the probability measure $P(w, \cdot)$ on $(X, \mathcal{X})$ is the same as the probability measure $P(w', \cdot)$.) Then, we may view $W = \cup_{x \in X} S_x$ and think of each element of W as a pair (x,s) with $s \in S_x$. Because the function P is onto X, we can view P as a projection: $P(x,s) = x$.

When G acts on W, it will take each element (x,s), where $s \in S_x$, to an element $(g.x,s')$ with $s' \in S_{g.x}$. This preserves the fibers of P. Also, we see that when G acts on W via the group element g, it automatically acts on X by the same element, because $g.x = g.P(w) = P(g.w)$. □

Meaning An observer's perceptual experiences can have a rich structure, e.g., a 3D structure that is locally Euclidean, and that transforms predictably and systematically as the observer acts, but this entails absolutely nothing about the structure of the objective world. This is wildly counterintuitive. We naturally assume that the rich structure of our perceptual experiences, and their predictable transformations as we act, must surely be an insight into the true structure of the objective world. The Invention of Symmetry Theorem shows that our intuitions here are completely wrong.

Note that the action of G on the world need not be a group action: the coordinate s in the fiber could go to *any* s' at all in $S_{g.x}$. Also, there is no requirement on the nature of the different S_x: they could be anything at all. So, G need not be a symmetry group of the world: *the world need not have the structure the observer sees.* All that is required is that the observer's action on the world faithfully flows back to a *group* action on itself via its own perceptions: the observer's actions and its own perceived symmetries are compatible. That this is *mediated* in the world does not imply that the world shares the symmetry: that the world has this symmetry could be merely a conceit of the observer (see also Terekhov and O'Regan (2013) and Laflaquiere et al. (2013) for how Euclidean perceptions could be learned by interacting with a non-Euclidean world). An important special case of the *Invention of Symmetry Theorem* arises when the symmetry group is the Lorentz or Poincare group (or super-symmetries which include these). In this case, we have the corollary that an observer can successfully invent space-time even if the objective world has no space-time or has only local versions of space-time. We call this corollary the *Invention of Space-Time Theorem*.

Perhaps this theorem seems artificial: Why in the world would an observer's perceptions carve up W into such strange subsets S_x? Well, one good reason would be a fitness function on W that happened to be constant, or roughly constant, within each subset S_x but differed between subsets. Then selection pressures would tend to shape precisely this strange carving of the world. In this case, we see the world as Euclidean not because this perception is veridical, but because it suitably represents what matters in evolution: fitness. For example, our perception of space might simply be a representation of the fitness costs that we would incur for locomoting and similar actions.

Taking this a speculative step further, because the observer itself is part of the world W which is the domain of the fitness function, it follows that, as the structure of the observer evolves, the fitness function itself is likely to change. In this sense, the observer and its fitness functions coevolve. If, as seems plausible, observers that are less costly in their requirements of information and computation are, *ceteris paribus,* fitter, then we might find it to be a theorem that the coevolution of observer and fitness function leads inexorably to group structures and actions. If so, this result would show how the groups that appear in physical theories might in fact arise from evolutionary constraints.

Illusion and hallucination

Perceptual illusions have been subjects of interest for millennia (see Wade 2014, for a review). The modern textbook account of perceptual illusions treats them as rare cases in which perception fails to be veridical. The textbook *Vision Science*, for instance, says "…veridical perception of the environment often requires heuristic processes based on assumptions that are usually, but not always, true. When they are true, all is well, and we see more or less what is actually there. When these assumptions are false, however, we perceive a situation that differs systematically from reality: that is, an illusion" (Palmer 1999, p 313).

Gregory (1997) agrees with the textbook account, but admits "It is extraordinarily hard to give a satisfactory definition of an 'illusion.' It may be the departure from reality or from truth; but how are these to be defined? As science's accounts of reality get ever more different from appearances, to say that this separation is 'illusion' would have the absurd consequence of implying that almost all perceptions are illusory. It seems better to limit 'illusion' to systematic visual and other sensed discrepancies from simple measurements with rulers, photometers, clocks, and so on."

The interface theory of perception claims, on evolutionary grounds, that we should expect *none* of our perceptions to be veridical. This entails that the textbook theory of illusions as departures from truth can't be right. If we concede that there is a vital divide between perceptions deemed normal and illusory (which, e.g., Rogers 2014 does not), then we must find new grounds for that divide and construct a new theory of illusions.

The obvious place to seek new grounds is the theory of evolution. The basic mistake of the textbook theory is its claim that selection shapes perceptions to be true. This forces illusions to be departures from truth. The correct claim is that selection shapes perceptions to guide adaptive behavior. This forces the interface theory of illusions to identify illusions as perceptions that fail to guide adaptive behavior (Hoffman 2011).

Is this plausible? Let's check a couple of cases. Consider the Necker cube in Fig. 7. The textbook theory says that what we see is illusory, because it's untrue: we see a 3D cube when in truth it's flat, and we see it flip in depth when in truth nothing changes. The interface theory says that what we see is illusory, because it fails to guide adaptive behavior: we see a 3D shape that we normally could grasp (or avoid, etc.) but here cannot, and we see flips in depth that normally require a change in grasp but here do not. In other words, our perception is illusory because it invites us to initiate behaviors or make categorizations that don't work.

Of course, we're not fooled by the figure or tempted to grab in vain at thin air. The textbook theory explains this by claiming that some of our perceptions of this figure are veridical: Stereovision reports the truth that the page is flat, and our hands confirm this. The interface theory explains that stereovision invites behaviors at odds with those appropriate for a cube. This mismatch in behavioral advice, and our confidence in, e.g., the advice of stereovision, keeps us from being fooled.

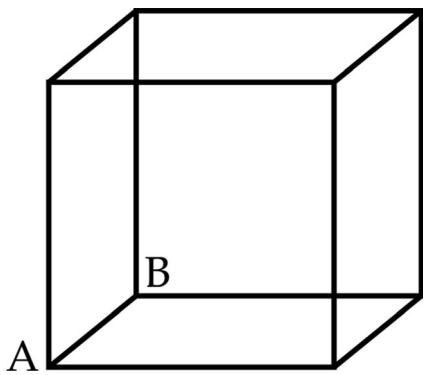


Fig. 7 Necker cube

But doesn't the textbook theory also say that normal perceptions guide adaptive behavior whereas illusory perceptions do not? So what's the difference, and what's new about the interface theory? Indeed, the textbook theory does say this and even points to evolution as the reason. The difference is that the textbook theory, but not the interface theory, claims that perceptions guide adaptive behaviors because, and only if, they are veridical. This claim is stronger than that of the interface theory and is in fact false. It gets evolution wrong.

Changing modalities from vision to taste, a striking gustatory illusion can be induced by miraculin—a protein found in the red berries of *Richadella dulcifica* (Koizumi et al. 2011). For more than an hour after eating these berries, sour substances taste sweet. The textbook theory of illusions would say that the sweet taste is illusory, because it's not veridical. But this sounds odd. What can we possibly mean by the *veridical* taste of a molecule? What objective standard tells us its true taste? Couldn't taste vary across species? One might hope, for instance that dung tastes different to coprophagic creatures, such as pigs, rodents, and rabbits, than it does to us (Hübner et al. 2013).

The interface theory of illusions does not require implausible claims about the true taste of a molecule. It simply says that the sweet taste induced by miraculin is illusory, because it does not guide adaptive behaviors. An animal with low blood sugar, for instance, that needed quick carbs, would eat the wrong foods. Thus, according to the interface theory, illusory perception cannot be defined in terms of nonveridicality: indeed all perceptions are fundamentally non-veridical, but only some of them are illusory.

Discrepancies between perception and the measured world may provide a distinct way of defining illusions (cf. Gregory, 1997). We consider this a weaker form of "illusion"; it is more a lack of consistency between the results of two different measurement procedures.

The textbook account of hallucinations claims that they are nonveridical perceptions. They differ from illusions in a key respect: whereas most normal people report seeing an illusion if placed in the right context, hallucinations are idiosyncratic perceptions seen by just one, or perhaps a few, individuals, and need not depend on the context. The interface theory of hallucinations simply modifies the textbook account in one respect: it replaces the claim that hallucinations are nonveridical perceptions with the claim that hallucinations are perceptions that do not guide adaptive behavior. The interface theory still says that hallucinations differ from illusions in

that an illusion is seen by most normal people if placed in the right context, but a hallucination is an idiosyncratic perception seen by just one, or perhaps a few, individuals, and need not depend on the context.

Conclusion: objections and replies

Numerous objections have been raised against the interface theory of perception. We conclude by canvassing some objections and offering replies.

Objection 1 What's new here? Of course perception is adaptive. We can go back to Gibson and see the same point. But how could it be anything else?

Reply Indeed, Gibson and others recognized that perception is adaptive. But Gibson's theory differs from the interface theory on three key points. First, Gibson got evolution wrong: He claimed that evolution shapes *veridical* perceptions of those aspects of the world that have adaptive significance for us. Thus Gibson proposed naïve realism, not the interface theory. Koenderink (2014) takes Gibson to task for this, noting that he "…holds that a stone of the right size has the affordance of being throwable, even in the absence of any observer. His affordance is like a property of the stone, much like its weight, or shape. This is quite unlike von Uexküll, who holds that a stone can indeed appear throwable—namely, to a person looking for something to throw. Here, the affordance is not a property of the stone but of an observer in a certain state. Gibson's notion derives from his reliance on the All Seeing Eye delusion.…"

Second, Gibson denied that perception involves information processing. The interface theory does not. Evidence for information processing is now overwhelming.

Third, in place of information processing Gibson proposed direct perception: We directly perceive, for instance, that something is edible; we do not use information processing to infer from visual and tactile cues that it is edible. But this raises a problem for Gibson: Are illusions direct misperceptions? What could one possibly mean by direct misperception? How could a theory of direct perception explain illusions? Gibson never solved this problem (Fodor and Pylyshyn 1981). Instead, as Gregory (1997) notes: "To maintain that perception is direct, without need of inference or knowledge, Gibson generally denied the phenomena of illusion." The interface theory does not deny the phenomena of illusion. Instead, one of its strengths is that it offers a new theory of illusions that seems far more plausible than the textbook account.

Objection 2 The interface theory of perception makes science impossible. If our perceptions are not veridical, then we can never have reliable data to build our theories.

Reply The interface theory poses no problem for science. It claims that our perceptions are not veridical reports of reality. If this claim is correct, then we can discard a particularly simple theory of perception. But that is not to discard the methodology of science. We can continue in the normal fashion to propose scientific theories and make falsifiable predictions about what we will observe. If our theory attributes some structure to the world W, and posits some functional relation $P: W \rightarrow X$ between the world and our perceptions that is not veridical, we can still deduce from W and P what measurement results we should expect to find in X. The methodology of science is not so fragile that it fails entirely if P happens not to be some simple function, such as an isomorphism.

Objection 3 You use the theory of evolution to show that our cognitive faculties are not reliable guides to the true nature of objective reality. But if our faculties are not reliable, then the theories we create are not reliable, including the theory of evolution. Thus, you are caught in a paradox.

Reply We use evolutionary games to show that natural selection does not favor veridical perceptions. This does not entail that *all* cognitive faculties are not reliable. Each faculty must be examined on its own to determine how it might be shaped by natural selection.

Perhaps, for instance, selection pressures favor accurate math; one who accurately predicts that the payoff for eating an apple today when hungry, combined with the payoff for eating an apple yesterday when equally hungry, is roughly twice the payoff obtained on either day, might have a selective advantage over his math challenged neighbor. Perhaps selection favors accurate logic; one who combines estimates of payoff in accord with probabilistic logic might avoid having nature and competitors make fitness Dutch books against him.[6] This is not to predict that natural selection should make us all math whizzes for whom statistical inference is quick and intuitive. To the contrary, there is ample evidence that we have systematic weaknesses and rely on fallible heuristics and biases (Kahneman 2011). Whereas in perception the selection pressures are almost *uniformly* away from veridicality, perhaps in math and logic the pressures are not so univocal, and partial accuracy is allowed. The point is that we don't know

[6] Recall that a Dutch book in gambling is a set of odds and bets that guarantee one person a profit regardless of the outcome of the gamble. One can show that if one's degrees of belief don't satisfy the axioms of probability and Bayes rule, then one can be the victim of a Dutch book (e.g., Jeffrey 2004).

until we study the implications of natural selection for these specific mental faculties.

Objection 4 You say in the abstract, and elsewhere, that our perceptions have been shaped to hide the truth. This is a fallacy. Adaptation doesn't work that way. Our perceptions have been shaped to improve fitness wherever possible—where "fitness" could be nonveridical or veridical—so there is no "hiding" which implies an intentionality that evolution can't and doesn't have.

Reply Yes, of course. We use "hide," because it powerfully and succinctly makes an important point, and we're not terribly worried that readers might be taken in by any connotations of intentionality.

Objection 5 Isn't the interface theory of perception just the utilitarian theory of perception proposed earlier by Braunstein (1983) and Ramachandran (1985; 1990)?

Reply Not at all. The utilitarian theory of perception claims that evolution has shaped perception to employ a set of heuristics or "bag of tricks", rather than sophisticated general principles. It claims that these tricks are employed to *recover* useful information about an objective physical world (a claim which the interface theory explicitly denies). Accordingly, when these tricks are sufficiently successful (which sometimes they're not), our perceptions are thus *veridical* about useful aspects of reality. The utilitarian theory is a naïve realist theory, not an interface theory.

Objection 6 The interface theory says that our perceptions of objects in space-time are not veridical, but are just species-specific icons. Doesn't it follow that (1) no object has a position, or any other physical property, when it is not perceived, and (2) no object has any causal powers? If so, isn't this a *reductio* of the interface theory? It entails, for instance, that neurons, which are objects in space-time, have no causal powers and thus cause none of our behaviors.

Reply The interface theory indeed makes both predictions. If either proves false, then the interface theory is false. No one can claim that the interface theory makes no falsifiable predictions.

But neither prediction has yet proven false. Moreover, both predictions are made by the standard "Copenhagen" interpretation of quantum theory and by more recent interpretations, such as quantum Bayesianism (Allday 2009; Fuchs 2010). According to these interpretations an electron, for instance, has no position when it is not observed and the state of the electron does not, in general, allow one to predict the specific position one will find when making a position measurement, i.e., no causal account can be given for the precise measurement obtained. Thus, both predictions of the interface theory are compatible with current physical theory and experimental data.

Both predictions are, of course, deeply counterintuitive. Our intuitions here are the result of evolutionary pressures to interact successfully with the world, e.g., tracking objects behind occluders and predicting where they're likely to reappear. Hence, they are innate as far as an individual child is considered. Belief in "object permanence," the belief, e.g., that a doll still exists and has a position even when it's hidden behind a pillow, begins as early as 3 months postpartum and is well-ensconced by age 18 months (Bower 1974; Baillargeon and DeVos 1991; Piaget 1954). Rich causal interpretations of physical objects are evident in children by age 6 months (Carey 2009; Keil 2011). We have been shaped by evolution to believe early on that objects exist unperceived and have causal powers.

The interface theory predicts that these beliefs are adaptive fictions.

Objection 7 The interface theory is nothing but the old sense-datum theory of perception—which claims that we see curious objects called sense data and do not see the world itself—that philosophers rightly discarded long ago.

Reply The short reply is: No, the interface theory is not a sense-datum theory and does not entail the existence of the sense data, or sensibilia, posited by such theories.

The longer reply is: "Sense-datum theory" covers a diverse set of philosophical ideas about perception. Precursors to these ideas can be found in the notion of sensory impressions or ideas proposed by the British empiricists, Locke, Berkeley and Hume. The origin of the modern conceptions of sense data can be traced to the writings of Moore (1903) and Russell (1912; 1918).

According to the act-object theory of sense data originated by Moore, each sense datumis a real concrete object with which an observer has a primitive relation in an act of perceptual awareness, but which nevertheless is distinct from that act of awareness. The act of perceptual awareness is a kind of knowing, and the sense datum thus known has exactly the properties it appears to have. Moreover, some philosophers propose that sense data have exact and discernible properties (if a sense datum is speckled, the sense datum has a precise number of speckles), and that they are objects that are private to each subject and distinct from physical objects.

The sense datum theory has been criticized by philosophers for, *inter alia*, conflating nonconceptual phenomenal consciousness with the physical events that are perceived (Coates 2007; Sellars 1956), for getting wrong the phenomenology of ordinary perceptual experience (Austin 1962; Firth 1949; Merleau-Ponty 1945), for requiring determinate phenomenal properties (Barnes 1944), and for breeding epistemological issues, such as skepticism or idealism. Logical

positivists and logical empiricists conscripted sense data into service as the incorrigible foundation for a verificationist program of knowledge, and when this program was discredited, e.g., by Quine's (1951) attack on the analytic/synthetic division and Hanson's (1958) attack on the theory neutrality of observation data, the theory of sense data suffered similar decline.

Sense data also run afoul of current theory and empirical data in vision science. The shapes, lightness, colors, and textures of sense data were claimed to be seen directly and without intervening inferences. It is now clear that these visual properties are the end products of computations of such sophistication that they are still not fully understood (Frisby and Stone 2010; Knill and Richards 1996; Marr 1982; Palmer 1999; Pizlo et al. 2014).

The interface theory does not entail that perception is an act whose objects are sense data or that sense data are an incorrigible foundation for an edifice of verified knowledge. The interface theory is metaphysically neutral, in that it does not posit anything about the world W other than measurability (in the probability-theory sense, rather than the scientific measurement sense). In particular, in addition to not entailing the existence of sense data, the interface theory does not entail idealism. However, it can be embedded in a mathematically rigorous theory of idealism (Hoffman 2008; Hoffman and Prakash 2014).

The interface theory is a general, but mathematically precise, theory of perception and action. It says that in a world represented by the probability space $(W, \mathcal{W}, \mu)$, a perceiving agent, $\mathcal{A}$, is a six-tuple $\mathcal{A} = (X, G, P, D, A, N)$, where X and G are measurable spaces, $P : W \times \mathcal{X} \rightarrow [0, 1]$, $D : X \times \mathcal{G} \rightarrow [0, 1]$ and $A : G \times \mathcal{W} \rightarrow [0, 1]$ are Markovian kernels, and N is an integer. X denotes the agent's possible perceptions, G its possible actions, P its perceptual mapping, D its decision process, A its action on the world, and N its counter of perceptions (as described more fully in the section on the PDA Loop). Perceiving agents can be combined, in several mathematically precise ways, to create new perceiving agents that are not reducible to the original agents (Hoffman and Prakash 2014).

When the evolution of a perceiving agent is shaped by a (suitably normalized) fitness function $f: W \rightarrow \mathbb{R}^+$, then that agent is shaped *towards* an X and P that maximize the mutual information $I(\mu f; \mu f P)$ and not the mutual information $I(\mu; \mu P)$; this is the formal way to state that perception is tuned to fitness rather than to veridicality.

This, in a nutshell, is the mathematical structure of the interface theory. The proper philosophical interpretation of this structure is a separate and interesting question. In response to this question we, as the authors of the theory, can opine but are not final authorities. When Schrodinger, for instance, first proposed his famous equation, he mistakenly interpreted its wave functions as waves of matter; Born later corrected that interpretation to waves of probability amplitudes.

With this proviso, we interpret X as the possible phenomenal states of the observer, and we interpret a specific $x \in X$ not in terms of an act-object relation as proposed by the sense-datum theory, but as a specific phenomenal aspect or constituent of the observer's mind; ours is a one-place account rather than the two-place account of the sense datum theory. In this regard, our interpretation is much like the critical realist interpretation of Coates (2007). Also like Coates, we take phenomenal qualities to carry information about the environment that normally triggers them. However, whereas Coates takes this information to be about mind-independent physical objects, we take it to be information about fitness and the fitness consequences of possible actions; there is a mind-independent world, but it almost surely does not consist of physical objects in space-time that are the targets of intentional content proposed by Coates.

Objection 8 The interface theory entails that there are no public physical objects. But this is absurd. Even our legal system knows this is absurd. My car is a public object, and if you steal it you break the law.

Reply The interface theory denies that there are public physical objects, but it does not deny that there is an objective reality that exists even if not perceived by a specific observer. When you and I both look at your car, the car I experience is not numerically identical to the car you experience. We both interact with the same objective reality, and we both represent our interaction with a species-specific set of experiences that we refer to as a car. But the objective reality is not a car and doesn't remotely resemble a car; moreover, the car of your experience is distinct from the car of my experience.

This might seem puzzling or logic chopping, but it's quite straightforward. Consider, for instance, the Necker cube of Fig. 4. Sometimes you see a cube with corner A pointing forward (call it "cube A"), and other times a cube with corner B pointing forward ("cube B"). Your cube A experience is not numerically identical to your cube B experience. If you and a friend are both looking at Fig. 4, and she experiences cube A while you experience cube B, then clearly your cube experiences are not numerically identical. Even if you both see cube A at the same time, your cube A experiences are not numerically identical. And yet we have no problem talking about "the cube," because we both assume that the experience of the other, although numerically distinct from our own experience, is nevertheless similar enough to permit communication. In the same way, we can discuss our migraine headaches, even though there are no public headaches; we assume that the headaches of others are similar enough to our own to make communication possible.

When I see your car, I interact with an objective reality, but my experience of that reality as a car is not an insight into that reality, but is merely a species-specific description shaped by

natural selection to guide adaptive behaviors. The adaptive behaviors might include complimenting you on your car or offering to wash it, but not stealing it. If I do steal it, I've changed objective reality in a way that injures you and rightly puts the law on your side, but the reality that I've changed doesn't resemble a car.

Similarly, if I'm in California and you're in New York and we're competing in an online video game trying to steal cars, I might find "the Porsche" before you do and steal it. But the Porsche on my screen is not numerically identical to the Porsche on your screen. What is behind my screen that triggers it to display a Porsche is a complex tangle of code and transistors that does not resemble a Porsche. I assume that the Porsche on my screen is similar to the Porsche on yours, so that we can discuss genuinely and compete for the Porsche. But there is no public Porsche.

We understand that our denial of public physical objects—our claim that physical objects are simply icons of one's perceptual interface—appears, to almost everyone, as not just counterintuitive but *prima facie* false. To many it's not worth dignifying with a response. That's how deeply *H. sapiens* assumes the existence of public physical objects. This assumption is an adaptive fiction shaped by natural selection, because it's helpful in the practical endeavors required to survive and reproduce. This fiction becomes an impediment when we turn to scientific endeavors, such as solving the mind-body problem. Here the assumption that neurons are public physical objects that exist unperceived and have causal powers is the starting point for almost all theories, and is, we propose, the reason for the (widely acknowledged) failure of all such theories to solve the mind-body problem.

Objection 9 You say that evolution drives veridicality to extinction only when it conflicts with fitness. In general, truth is useful and indeed optimal within the everyday human scheme: e.g., my chances of rendezvousing with you are better if I know the truth about where you are.

Reply Yes, my chances of rendezvousing with you are better if I know the truth about where you are, just as my chances of deleting a text file are better if I know the truth about where the icon of the text file is on my desktop interface. However, a truth about the state of the interface is not *ipso facto* a truth about objective reality. Knowing that the icon is in the center of the desktop does not entail that the file itself is in the center of the computer. Similarly, knowing where in space-time to rendezvous with you does not entail any knowledge of objective reality; indeed it does not even entail that space-time itself is an aspect of objective reality (as we proved above in the *Invention of Space-Time Theorem*). An interface can be an accurate guide to behavior without being an accurate guide to the nature of objective reality.

This research was supported in part by grants from the National Science Foundation, Procter & Gamble, VF Corporation, and the Federico and Elvia Faggin Foundation.

Appendix 1: measure theory and group actions

Measure theory

The minimal commonly accepted structure for probability spaces and their appropriate mappings is as follows. A *measurable space* is a set X together with a distinguished collection of subsets (the collection of events) $\mathcal{X}$ In order to be able to assign probabilities to events, this collection needs to satisfy the following three properties: (1) $X \in \mathcal{X}$, (2) $\mathcal{X}$ is closed under countable union, and (3) $\mathcal{X}$ is closed under complement. We say that $\mathcal{X}$ is a σ-*algebra* on X, and then any function μ assigning values between 0 and 1 to the events in $\mathcal{X}$ is a *probability measure* if it satisfies: (1) $\mu(\phi) = 0$, where ϕ is the empty set, (2) μ is σ-*additive on* $\mathcal{X}$: If $\{E_i\}_{i=1}^{N}$ is a disjoint collection of events in $\mathcal{X}$ (i.e., $E_i \cap E_j = \phi, i \neq j$), then $\mu\left(\cup_{i=1}^{N}\{E_i\}\right) = \sum_{i=1}^{N} \mu(E_i)$, for N either finite or ∞ Once this is done, all the usual rules of probability obtain.

Finally, the collection of "appropriate" mappings $f{:}X \to Y$ between measurable spaces $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$ are called, again, the *measurable* mappings: f is measurable if all its values in an event F in $\mathcal{Y}$ came precisely from an event E in $\mathcal{X}$. Technically, the measurability of the function f can be written as: $\forall F \in \mathcal{Y}, f^{-1}(F) \in \mathcal{X}$ (Revuz 1984). In statistics, measurable mappings on X are the familiar *random variables* on X, taking values in Y.

Group actions

Definition A group G *acts* on a set X if there is a mapping $G \times X \to X$, which we will denote by $(g, x) \to g, x$ such that, if e is the identity element in G, $e, x = x, \forall x \in X$ and $\forall g, h \in G, g(h.x) = (gh).x, \forall x \in X$.

The group action is *transitive* if, given any pair of points x, $y \in X$ there is an element g $\in G$ with $g.x = y$. A set X acted on transitively by the group G is said to be a G-*set.*

In terms of geometry, suppose the G-set X is a differentiable manifold and G is a Lie group with a smooth action on X. In this instance we say that G is the *symmetry group* of the geometry on the *homogeneous space X.* We do not address this situation in this paper.

Examples The action of G in the PDA loop on W, relevant to the IOS Theorem

Example 1 Suppose that all fibres S_x (see the proof of the Invention of Symmetry Theorem) are the same set, S say. Then $W=X\times S$: the world is a Cartesian product. If X is a G-set, let the action of $g\in G$ on (x,s) be given by $g.(x,s)=(g.x,s)$. This *is* a group action on W, but it is trivial on the fibers and is certainly not transitive on the whole of W, which is thus *not* a G-set.□

Example 2 Here is a toy example of an agent's G acting on W in such a way that there is a *bona-fide* group action of G on X and we always have $g.s\in S_{g.x}$ whenever $s\in S_x$, yet there is *no* group action on W as a whole: W is not a G-set.

Let $X=\{x,x'\}$ be a set with two elements and let $G=\{e,a\}$, with $a^2=e$; this group is denoted $\mathbb{Z}_2$. Let $S=\{y,y'\}$ and $W=X\times S$. Let G act on W as follows: $e.w=w,\ \forall w\in W$ and $a.(x,y)=(x',y')$; $a.(x,y')=(x',y')$; $a.(x',y)=(x,y)$; $a.(x',y')=(x,y')$. Then G acts transitively on X, but its effect on W as a whole is *not* a group action. We have $a^2=e$ but $a.(a.(x,y))=a.(x',y')=(x,y')$ and this does not equal (x,y), as would have been required for a group action.

The reader can see how easy it is to generate further examples of this behavior, because we have not specified any structure on the action mappings A, other than that they are functions. The world is much richer than the agent, so it is clear that no A can be *onto* the world. In this example, A is not 1:1, but it is not at all hard to generate examples where A*is* 1:1, as the next example shows.

Example 3 Let $X=\{x,x'\}$ and $S=\mathbb{N}$, the natural numbers. Again, let the group be $\mathbb{Z}_2$. Let $a.(x,s)=(x',s+3)$ and $a.(x',s)=(x,s)$. Then $a.a.(x,s)=(x,s+3)\neq e.(x,s)=(x,s)$.

References

Adelson, E. H., & Pentland, A. (1996). The perception of shading and reflectance. In D. C. Knill & W. Richards (Eds.), *Perception as Bayesian inference* (pp. 409–423). Cambridge: Cambridge University Press.

Allday, J. (2009). *Quantum reality: theory and philosophy*. Boca Raton, FL: CRC Press.

Allen, J. A., & Clarke, B. C. (1984). Frequency-dependent selection—homage to Poulton, E.B. *Biological Journal of the Linnean Society, 23,* 15–18.

Austin, J. L. (1962). *How to do things with words*. Oxford: Clarendon Press.

Baillargeon, R., & DeVos, J. (1991). Object permanence in young infants: Further evidence. *Child Development, 62,* 1227–1246. doi:10.2307/1130803

Barnard, C. J., & Sibly, R. M. (1981). Producers and scroungers: A general model and its application to captive flocks of house sparrows. *Animal Behavior, 29,* 543–550.

Barnes, W. F., (1944). The myth of sense-data. *Proceedings of the Aristotelian Society, 45*

Berkeley, G. (1710/2012). *A treatise concerning the principles of human knowledge*. New York: Dover.

Berkeley, G. (1713/1979). *Three dialogues between Hylas and Philonous*. Indianapolis: Hackett.

Bomze, I. M. (1983). Lotka-Volterra equations and replicator dynamics: A two dimensional classification. *Biological Cybernetics, 48,* 201–211.

Boswell, J. (1791/1993). *The life of Samuel Johnson*. New York: Everyman's Library.

Bower, T. G. R. (1974). *Development in infancy*. San Francisco: Freeman Press.

Brainard, D. H., & Freeman, W. T. (1997). Bayesian color constancy. *Journal of the Optical Society of America A, 14,* 1393–1411.

Byrne, A., & Hilbert, D. R. (2003). Color realism and color science. *Behavioral and Brain Sciences, 26,* 3–26.

Braunstein, M. L. (1983). Contrasts between human and machine vision: should technology recapitulate phylogeny. In J. Beck, B. Hope, & A. Rosenfeld (Eds.), *Human and machine vision*. New York: Academic Press.

Brewer, W. (2011). *Perception and its objects*. New York: Oxford, UK: Oxford University Press.

Campbell, J., & Cassam, Q. (2014). *Berkeley's puzzle: what does experience teach us?* Oxford, UK: Oxford University Press.

Carey, S. (2009). *The origin of concepts*. Oxford, UK: Oxford University Press.

Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, MA: MIT Press.

Coates, P. (2007). *The metaphysics of perception*. New York: Routledge.

Cornwell, J. F. (1997). *Group theory in physics*. San Diego: Academic Press.

Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory*. New York: Wiley.

Cox, R. T. (1961). *The algebra of probable inference*. London: Oxford University Press.

Cuipers, R. H., Kappers, A. M. L., & Koenderink, J. J. (2003). The metrics of visual and haptic space based on parallelity judgements. *Journal of Mathematical Psychology, 47,* 278–291.

Fechner, G. (1860/1966). Elements of psychophysics. (H.E. Adler translator). New York: Holt, Rinehart and Winston.

Feldman, J. (2013). Tuning your priors to the world. *Topics in Cognitive Science, 5,* 13–34.

Fields, C. (2014). This boundary-less world. In D. Chopra (Ed.), *Brain, mind, cosmos*. La Costa, CA: Chopra. **Chapter 13**.

Firth, R., (1949). Sense-data and the percept theory. *Mind, 58 & 59*.

Fish, W. (2009). *Perception, hallucination, and illusion*. New York: Oxford University Press.

Fish, W. (2010). *Philosophy of perception: A contemporary introduction*. New York: Routledge.

Fodor, J. A., & Pylyshyn, Z. W. (1981). How direct is visual perception?: Some reflections on Gibson's Ecological Approach. *Cognition, 9,* 139–196.

Frisby, J. P., & Stone, J. V. (2010). *Seeing: The computational approach to biological vision*. Cambridge, MA: MIT Press.

Fuchs, C. (2010). *QBism, the Perimeter of Quantum Bayesianism*. Available online at: arXiv:1003.5209v51.

Galilei, G. (1623/1957). *The assayer*. In S. Drake (translator) *Discoveries and opinions of Galileo*. New York: Random House.

Geisler, W. S. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology, 59,* 167–192.

Geisler, W. W., & Diehl, R. L. (2002). Bayesian natural selection and the evolution of perceptual systems. *Philosophical Transactions of the Royal Society of London B, 357,* 419–448.

Geisler, W. S., & Diehl, R. L. (2003). A Bayesian approach to the evolution of perceptual and cognitive systems. *Cognitive Science, 27,* 379–402.

Gibson, J. J. (1986). *The ecological approach to visual perception*. Hillsdale, New Jersey: Lawrence Erlbaum.

Gregory, R. L. (1997). Knowledge in perception and illusion. *Philosophical Transactions of the Royal Society of London B, 352,* 1121–1128.

Gwynne, D. T., & Rentz, D. C. F. (1983). Beetles on the bottle: Male buprestids make stubbies for females. *Journal of Australian Entomological Society, 22*(79), 80.

Hanson, N. R. (1958). *Patterns of discovery.* Cambridge: Cambridge University Press.

Hardin, C. L. (2008). Color qualities and the physical world. In E. Wright (Ed.), *The case for qualia* (pp. 143–154). Cambridge, MA: MIT Press.

Harper, G. R., & Pfennig, D. W. (2007). Mimicry on the edge: Why do mimics vary in resemblance to their model in different parts of their geographical range? *Proceedings of the Royal Society B, 274,* 1955–1961.

Hofbauer, J., & Sigmund, K. (1998). *Evolutionary games and population dynamics*. New York: Cambridge University Press.

Hoffman, D. D. (1998). *Visual intelligence: How we create what we see*. New York: W. W. Norton.

Hoffman, D. D. (2008). Conscious realism and the mind-body problem. *Mind & Matter, 6,* 87–121.

Hoffman, D. D. (2009). The interface theory of perception. In S. Dickinson, M. Tarr, A. Leonardis, & B. Schiele (Eds.), *Object categorization: Computer and human vision perspectives* (pp. 148–165). New York: Cambridge University Press.

Hoffman, D. D. (2011). The construction of visual reality. In J. Blom & I. Sommer (Eds.), *Hallucinations: Theory and practice* (pp. 7–15). New York: Springer.

Hoffman, D. D. (2012). The sensory desktop. In J. Brockman (Ed.), *This will make you smarter: New scientific concepts to improve your thinking* (pp. 135–138). New York: Harper Perennial.

Hoffman, D. D. (2013). Public objects and private qualia: the scope and limits of psychophysics. In L. Albertazzi (Ed.), *The Wiley-Blackwell handbook of experimental phenomenology* (pp. 71–89). New York: Wiley-Blackwell.

Hoffman, D. D., & Prakash, C. (2014). Objects of consciousness. *Frontiers of Psychology, 5,* 577. doi:10.3389/fpsyg.2014.00577

Hoffman, D. D., & Singh, M. (2012). Computational evolutionary perception. *Perception, 41,* 1073–1091.

Hoffman, D. D., Singh, M., & Mark, J. T. (2013). Does natural selection favor true perceptions? *Proceedings of the SPIE 8651, Human Vision and Electronic Imaging, XVIII,* 865104. doi:10.1117/12.2011609

Horváth, G., Bernáth, B., & Molnár, G. (1998). Dragonflies find crude oil visually more attractive than water: Multiple-choice experiments on dragonfly polarotaxis. *Naturwissenschaften, 85,* 292–297.

Horváth, G., Malik, P., Kriska, G., & Wildermuth, H. (2007). Ecological traps for dragonflies in a cemetery: The attraction of sympetrum species (odonata: Libellulidae) by horizontally polarizing black gravestones. *Freshwater Biology, 52,* 1700–1709.

Hübner, S., Töle, J., & Meyerhof, W. (2013). Taste and nutrition. *Ernaehrungs Umschau international, 60*(12), 222–227. doi:10.4455/eu.2013.042

Jaynes, E. T. (2003). *Probability theory: The logic of science*. Cambridge: Cambridge University Press.

Jeffrey, R. (2004). *Subjective probability: The real thing*. Princeton, NJ: Princeton University Press.

Kahneman, D. (2011). *Thinking fast and slow.* New York: Farrar, Strous and Giroux.

Kappers, A. M. L. (1999). Large systematic deviations in the haptic perception of parallelity. *Perception, 28,* 1001–1012.

Keil, F. C. (2011). Science starts early. *Science, 33,* 1022–1023.

Kersten, D., Mamassian, P., & Yuille, A. L. (2004). Object perception as Bayesian inference. *Annual Review of Psychology, 555,* 271–304.

Kleffner, D., & Ramachandran, V. (1992). On the perception of shape from shading. *Perception and Psychophysics, 52,* 18–36.

Knill, D., & Richards, W. A. (Eds.). (1996). *Perception as Bayesian inference*. New York: Cambridge University Press.

Koenderink, J. J. (2010). *Color for the sciences*. Cambridge, MA: MIT Press.

Koenderink, J.J. (2011). Vision as a user interface. *Human Vision and Electronic Imaging XVI, SPIE Vol. 7865*. doi: 10.1117/12.881671

Koenderink, J.J. (2013). World, environment, umwelt, and inner-world: a biological perspective on visual awareness, *Human Vision and Electronic Imaging XVIII, SPIE Vol. 8651*. doi: 10.1117/12.2011874

Koenderink, J. J. (2014). The all seeing eye? *Perception, 43,* 1–6.

Koenderink, J. J., van Doorn, A. J., de Ridder, H., & Oomes, S. (2010). Visual rays are parallel. *Perception, 39,* 1163–1171.

Koizumi, A., Tsuchiya, A., Nakajima, K., Ito, K., Terada, T., Shimizu-Ibuka, A., ... Abe, K. (2011). Human sweet taste receptor mediates acid-induced sweetness of miraculin. *Proceedings of the National Academy of Sciences, 108*(40), 16819–16824.

Laflaquiere, A., Terekhov, A.V., Gas, B., & O'Regan, K.O. (2013). Learning an internal representation of the end-effector configuration space. *IEEE Conference on Intelligent Robots and Systems*, Tokyo, Nov. 3-7, 1230–1235.

Land, M. F., & Nilsson, D. (2012). *Animal eyes*. New York: Oxford University Press.

Laplace P S, 1774 "Mémoire sur la probabilité des causes par les évènemens (Memoir on the probability of the causes of events)", in Mémoires de mathématique et de physique présentés à l'Académie royale des sciences, par divers savans, & lûs dans ses assemblées 6 621–656 (English translation by S M Stigler published in *Statistical Science, 1,* 364–378, 1986)

Lieberman, E., Hauert, C., & Nowak, M. A. (2005). Evolutionary dynamics on graphs. *Nature, 433*(7023), 312–316.

Locke, J. (1690). *An essay concerning human understanding*. London.

Mamassian, P., Landy, M., & Maloney, L. T. (2002). Bayesian modeling of visual perception. In R. Rao, B. Olshausen, & M. Lewicki (Eds.), *Probabilistic models of the brain: perception and neural functioN* (pp. 13–36). Cambridge, MA: MIT Press.

Mark, J.T. (2013). Evolutionary pressures on perception: when does natural selection favor truth? Ph.D. Dissertation, University of California, Irvine.

Mark, J. T., Marion, B. B., & Hoffman, D. D. (2010). Natural selection and veridical perceptions. *Journal of Theoretical Biology, 266,* 504–515.

Marion, B.B. (2013). The impact of utility on the evolution of perceptions, Ph.D. Dissertation, University of California, Irvine.

Marr, D. (1982). *Vision*. San Francisco: Freeman.

Mausfeld, R. (2002). The physicalist trap in perception theory. In D. Heyer & R. Mausfeld (Eds.), *Perception and the physical world: Psychological and philosophical issues in perception* (pp. 75–112). New York: Wiley.

Mausfeld, R., & Heyer, D. (Eds.). (2003). *Colour perception: Mind and the physical world*. Oxford, UK: Oxford University Press.

McDowell, J. (1996). *Mind and world*. Cambridge, MA: Harvard Press.

Merleau-Ponty, M. (1945). *Phenomenology of perception*. (translated by Colin Smith) London: Routledge, 2005.

Mitchell, M. (1998). *An introduction to genetic algorithms*. Cambridge, MA: Bradford Books MIT Press.

Moore, G. E. (1903). The refutation of idealism. *Mind, 12,* 433–453.

Noe, A. (2006). *Action in perception*. Cambridge, MA: MIT Press.

Noë, A. (2012). *Varieties of presence*. Cambridge, MA: Harvard Press.

Nowak, M. (2006). *Evolutionary dynamics: Exploring the equations of life*. Cambridge, MA: Belknap Press of Harvard University Press.

Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press.

Pinker, S. (1997). *How the mind works*. New York: W.W. Norton.

Pizlo, Z. (2001). Perception viewed as an inverse problem. *Vision Research, 41,* 3145–3161.

Pizlo, Z. (2010). *3D shape: Its unique place in visual perception*. Cambridge: MIT Press.

Pizlo, Z., Li, Y., Sawada, T., & Steinman, R. M. (2014). *Making a machine that sees like us*. New York: Oxford University Press.

Piaget, J. (1954). *The construction of reality in the child*. New York: Basic.

Poli, R., Langdon, W.B., & McPhee, N.F. (2008). *A field guide to genetic programming*. Published via http://Lulu.com and freely available at http://gp-field-guide.org.uk (with contributions by Koza, J.R.).

Pont, S. C., Nefs, H. T., van Doorn, A. J., Wijntjes, M. W. A., te Pas, S. F., deRidder, H., & Koenderink, J. J. (2012). Depth in box spaces. *Seeing and Perceiving, 25,* 339–349.

Quine, W. V. O. (1951). Two dogmas of empiricism. *The Philosophical Review, 60,* 20–43.

Ramachandran, V. S. (1985). The neurobiology of perception. *Perception, 14,* 97–103.

Ramachandran, V. S. (1990). In C. Blakemore (Ed.), *Interactions between motion, depth, color and form: The utilitarian theory of perception, in Vision: coding and efficiency*. New York: Cambridge University Press.

Revuz, D. (1984). *Markov chains*. Amsterdam: North-Holland.

Rogers, B. (2014). Delusions about illusions. *Perception*. doi:10.1068/p7731

Russell, B. (1912). *The problems of philosophy*. New York: Oxford University Press.

Russell, B. (1918). *The philosophy of logical atomism*. Reprinted (2010), New York: Routledge.

Samuelson, L. (1997). *Evolutionary games and equilibrium selection*. Cambridge, MA: MIT Press.

Sandholm, W. H. (2007). *Population games and evolutionary dynamics*. Cambridge, MA: MIT Press.

Searle, J. (2015). *Seeing things as they are: A theory of perception*. New York: Oxford University Press.

Sellars, P. (1956). Empiricism and the philosophy of mind. In H. Feigl & M. Scriven (Eds.), *Minnesota studies in the philosophy of science* (The Foundations of Science and the Concepts of Psychology and Psychoanalysis, Vol. I, pp. 253–329). Minneapolis, MN: University of Minnesota Press.

Shepard, R. (1994). Perceptual–cognitive universals as reflections of the world. *Psychonomic Bulletin & Review, 1,* 2–28.

Singh, M., & Hoffman, D. D. (2013). Natural selection and shape perception: shape as an effective code for fitness. In S. Dickinson & Z. Pizlo (Eds.), *Shape perception in human and computer vision: An interdisciplinary perspective* (pp. 171–185). New York: Springer.

Taylor, P., & Jonker, L. (1978). Evolutionarily stable strategies and game dynamics. *Mathematical Biosciences, 40,* 145–156.

Terekhov, A., & O'Regan, K.O. (2013). Space as an invention of biological systems, arxiv.org/abs/1308.2124.

Travis, C. (2013). *Perception: Essays after Frege*. Oxford, UK: Oxford University Press.

Trivers, R.L. (1976/2006). Foreword, in R. Dawkins, *The selfish gene*. New York: Oxford University Press.

Trivers, R. L. (2011). *The folly of fools*. New York: Basic Books.

von Uexküll, J. (1909). *Umwelt und Innenwelt der Tiere*. Berlin: Springer-Verlag.

von Uexküll, J. (1926). *Theoretical biology*. New York: Harcourt, Brace & Co.

von Uexküll, J. (1957). *A stroll through the worlds of animals and men: A picture book of invisible worlds, reprinted in Instinctive behavior: Development of a modern concept, C.H. Schiller*. New York: Hallmark Press.

Wade, N. J. (2014). Geometrical optical illusionists. *Perception*. doi:10.1068/p7695

Wozny, D. R., Beierholm, U. R., & Shams, L. (2010). Probability matching as a computational strategy used in perception. *PLoS Computational Biology, 6*(8), e1000871.

Yuille, A., & Bülthoff, H. (1996). Bayesian decision theory and psychophysics. In D. Knill & W. Richards (Eds.), *Perception as Bayesian inference*. New York: Cambridge University Press.