

Donald D. Hoffman

Department of Cognitive Sciences

University of California

Irvine, CA 92697 USA

949-824-6795

ddhoff@uci.edu

Word Count: 2797

The Interface Theory of Perception

Donald D. Hoffman

Department of Cognitive Sciences

University of California

Irvine, CA 92697 USA

Word Count: 2797

References: 23

Abstract

Our perceptual systems are products of evolution and have been shaped, in part, by natural selection. It is widely assumed that natural selection favors veridical perceptions, viz., perceptions that accurately describe aspects of the objective world relevant to fitness. This assumption has been tested using the mathematics of evolutionary game theory. It is false. Monte Carlo simulations reveal that veridical perceptions are never more fit, and generically are less fit, than non-veridical perceptions of equal complexity that are tuned to fitness. Veridical perceptions go extinct, and their extinction rate increases as complexity increases. These results motivate a new theory of perceptual systems—as species-specific interfaces shaped by natural selection to hide objective reality and guide adaptive behavior. For *H. sapiens* space-time is the desktop of the interface and physical objects are icons on the desktop. The shapes and colors of physical objects no more resemble objective reality than the shapes and colors of desktop icons resemble files in a computer.

Keywords

evolution, natural selection, evolutionary game theory, perception, Bayesian inference, veridical perception, illusion

When you open your eyes and scan your environment, billions of neurons and trillions of synapses spring into action. About one third of your most advanced processing power is recruited for the apparently simple act of looking. This is the finding of modern visual neuroscience (e.g., Werner & Chalupa, 2014).

This might be counterintuitive. If we think of looking as simply taking a picture, then it is indeed a puzzle why such processing power is necessary. After all, cameras successfully took pictures long before computers were even invented.

The standard explanation is that visual perception is not simply a passive process that takes a picture, but an active process that constructs all the depths, shapes, colors, motions and textures that we see. This process of construction is stunningly complex, a fact that becomes strikingly obvious as soon as one tries to build a device, such as a robotic vision system, that enacts the construction (e.g., Szeliski, 2010). The starting point for such a device is just a time-varying two-dimensional array of numbers, which chronicle the changing activations of photosensors. There are no objects, shapes, textures, motions or depths explicitly given in this array. All must be constructed from an apparently meaningless array of numbers. According to visual neuroscience, we need all that neural heft to power this complex construction.

Disagreements arise as to whether this construction should be viewed as a form of information processing, whether large portions of the visual world are constructed

all at once or just a bit at a time on a need-to-know basis, and whether action and embodiment are critical to the construction process (e.g., Chemero, 2009; Frisby & Stone, 2010; Hoffman, 2000; Marr, 1982).

But there is almost universal agreement that, in the normal case, our perceptual constructions are accurate *reconstructions* of the true state of affairs in the objective world. Our perceptions are normally *veridical*, in the sense that they accurately describe the state of the environment.

Perceptual Evolution

The standard argument for veridical perception is based on evolution: Those of our ancestors who saw more accurately had a competitive advantage vis-à-vis their contemporaries who saw less accurately, and thus were more likely to pass on their genes that coded for the more accurate perceptions. We are the fortunate offspring of those who, in each generation, saw more accurately, and so we can be confident that, under normal circumstances, our perceptions accurately describe those aspects of the objective environment that we need to know to survive and reproduce. As Palmer (1999, p. 6) says, “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...”

This argument sounds plausible. But is it in fact correct? We don't have to guess. Evolution by natural selection has precise mathematical formulations, such as evolutionary game theory, evolutionary graph theory, and genetic algorithms (e.g., Mitchell, 1998; Nowak, 2006). We can precisely define an exhaustive classification of perceptual strategies, including veridical strategies and various non-veridical strategies, and have them compete in evolutionary games across a variety of simulated worlds and subjected to a variety of different fitness functions.

This has been done, and the answer is clear: Veridical perceptual strategies are *never* more fit than equally complex non-veridical strategies that are tuned to the relevant fitness functions (Hoffman, Singh & Prakash, 2015a; 2015b; Mark, Marion & Hoffman, 2010). When they compete, veridical strategies are routinely driven to extinction. The probability that they might be fit enough to avoid extinction goes to zero as the complexity of the strategies increases.

The problem is not that veridical perceptions are necessarily counter-adaptive, but rather that veridicality is irrelevant to adaptation, meaning that veridicality per se contributes nothing when reward value is varied orthogonally to it. One can, of course, construct payoff matrices in which reward and veridicality are correlated. But such correlations are not generic, in the mathematical sense that the unbiased probability of their occurrence is near zero.

Thus natural selection does not, in generic cases, favor veridical perceptions. To the contrary, if a veridical perception happens to appear due to some mutation, then natural selection will, generically, work to remove it from the population.

This might be counterintuitive. How can perceptions be useful if they are not veridical, if they are not accurate descriptions of the true state of the objective world?

The Interface Metaphor

A helpful metaphor is the desktop interface of a computer. Suppose you're editing a PowerPoint presentation for an upcoming talk, and the icon for the presentation is red, rectangular and in the center of the desktop. Does that mean that the PowerPoint presentation itself inside the computer is red, rectangular and in the center of the computer? Certainly not. Anyone who thought so would simply be misunderstanding the function of the desktop interface. It's not there to accurately depict the objective reality inside the computer. To the contrary, it's there to *hide* that reality. If you had to know the details of the transistors, voltages, magnetic fields, and megabytes of system and application software, you would never finish your presentation in time for your talk. The interface provides you with simplified symbols intended to help you to interact with the computer successfully, while remaining blissfully ignorant of the complex reality of that computer.

The perceptual systems with which we have been endowed by natural selection are a species-specific interface that allows us to interact adaptively and successfully with objective reality, while remaining blissfully ignorant of the complexity of that objective reality. Space-time is the desktop of our perceptual interface, and physical objects are icons on that desktop. To ask whether the red color and round shape that I perceive of that apple on the table is the veridical color and shape of something in objective reality is the same category mistake as asking if the red color and rectangular shape of the icon for the PowerPoint presentation is the veridical color and shape of something in the computer.

Some Natural Objections

But, one might ask, isn't our perception of, say, a rattlesnake more than just an icon of our interface? After all, if the snake strikes you it could kill you. So surely it is more than just an icon.

Indeed, one would be well advised to stay clear of the snake. But for the same reason that one would be well advised not to carelessly drag the red PowerPoint icon to the trashcan. Not because one should take that icon literally. The PowerPoint is not literally red and rectangular. But one should take it *seriously*. Dragging the icon to the trashcan might cause the loss of weeks of work.

And that is the point. Evolution has shaped us with perceptual symbols to help us survive and reproduce. We had better take them seriously. If you see a snake don't

touch it, if you encounter a precipice don't step over it, if you see a lion don't try to mate with it. Those who don't take their perceptions seriously tend to exit life early and leave no genes behind.

But from the fact that we must take our perception seriously, it does not follow that we must take them literally, i.e., as true descriptions of a reality independent of the observer. To assert otherwise is a logical fallacy. But it is a fallacy to which *H. sapiens* appears particularly prone—perhaps because there were no selection pressures that favored those who took perceptions seriously but not literally over those who took perceptions seriously and literally. As long as one takes perceptions seriously, it doesn't much matter to evolution whether one also takes them literally or not. It only matters when one tries to step back and look at perception as a subject of scientific investigation. Only then does our natural proclivity to conflate seriously with literally become a genuine impediment to success.

But, one might argue, we can all look and agree that there is a rattlesnake writhing on the ground. Surely this means that the rattlesnake is more than just an icon of a perceptual interface, that it is in fact a true description of objective reality.

However, subjective agreement between perceptions does not logically entail the objective accuracy of those perceptions. We can all agree, for instance, that we see a 3D cube when we view a standard Necker cube display (see Figure 1). But we know that, despite our unanimous agreement, there is in fact no 3D cube. We agree

because we all construct our perceptual icons in the same, species-specific, manner—not because we see veridically.

Still, one might argue, there are systematic and predictable variations in our perceptions that occur when we act and that are the foundation for many of the invariances of our perceptions. These variations are in fact so lawful that we can write down mathematical expressions for them. I'm looking at that cardboard box from this angle, but if I move just a couple feet to the right I know with near certainty how the appearance of that box will transform, and I could even write down matrix equations and projection operators that will tightly match my experience as I move.

But the existence of these predictable variations of perception does not logically entail anything about the structure of objective reality. This is the surprising conclusion of an "Invention-of-Symmetry Theorem" proved by Chetan Prakash (Hoffman, Singh & Prakash, 2015a). The theorem only permits one to infer a lower bound on the cardinality of states of the objective world, but not to infer anything at all about its structure.

But, one might argue, if our perceptions are not veridical, then that implies that they are all illusory. But they are not all illusory. Thus there is something radically wrong with the theory that our perceptions are not veridical.

Indeed, the standard definition of a perceptual illusion is that it is a perception which most observers experience when presented with a specific stimulus, but which fails to be veridical.

However, natural selection does not, in generic cases, favor veridical perceptions. Thus the standard definition of illusion is wrong. But it is easily fixed: An illusion is a perception that most observers experience when presented with a specific stimulus, but which *fails to guide adaptive behavior*. The Necker cube, for instance, is illusory because its perceived 3D shape, if taken seriously, invites grasping behaviors that are certain to fail and are thus not adaptive.

New Prediction

A theory earns its keep, in part, by making new falsifiable predictions. The interface theory of perception makes a surprising, but falsifiable, prediction.

Consider this hypothesis:

H1: When it is not observed, an object has a definite value for each of its dynamical physical properties (e.g., position, momentum, spin, energy).

Most vision scientists accept this hypothesis, and assume that experiments do, or would, support it.

Now consider the opposite hypothesis, which is a clear prediction of the interface theory of perception, because it says that physical objects in spacetime are simply icons of a species-specific interface, not an insight into objective reality:

H2: When it is not observed, an object does *not* have a definite value for any of its dynamical physical properties.

Most vision scientists reject this hypothesis, and some even doubt that it could be tested experimentally. After all, they argue, how can one show that a physical property doesn't exist *when it's not observed*? You might as well speculate about how many angels can dance on the head of a pin.

But note that H1 is falsifiable if and only if H2 is falsifiable. And, as it happens, there are cases where both hypotheses are falsifiable. This conclusion is the brilliant work of John Stewart Bell (1964) based on the quantum phenomenon of entanglement. Bell's groundbreaking insight has motivated several careful experiments testing H1 and H2. The result in each case supports H2, viz., the prediction of the interface theory of perception: When it is not observed, a quantum object does *not* have a definite value for any of its dynamical physical properties (e.g., Giustina et al., 2013; Hensen et al., 2015; Mermin, 1985; 1990). This has led many physicists to reject *local realism*: the assumption that (1) quantum particles have objective and definite pre-existing values for all possible measurements before any measurement is made (realism), and (2) information about these values never propagates faster than the speed of light (locality).

One objection to the H2 prediction of the interface theory is that some physicists have tried to distinguish between the quantum regime and the “macroscopic” regime, and it has been posed as a bit of a puzzle why macroscopic objects don't exhibit such quantum effects as superposition and entanglement. However, it turns out they can. Physicists have recently created entanglement in a room-temperature wafer of silicon carbide about the volume of a red blood cell (Klimov et al., 2015). The interface theory of perception predicts that local realism is false not just for special macroscopic objects such as the wafer of silicon carbide, but for all macroscopic objects.

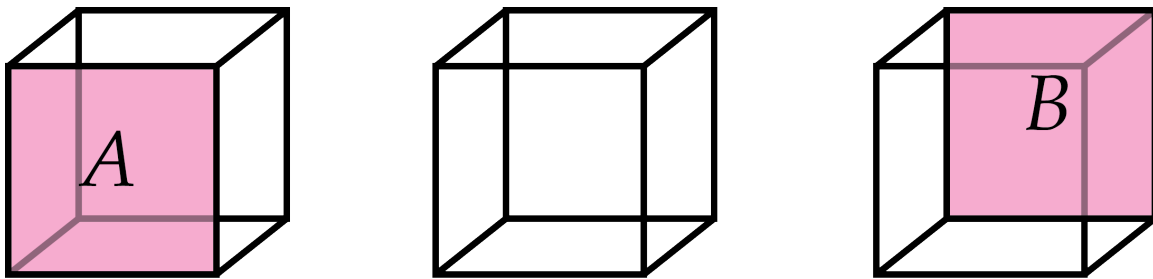


Figure 1. *The Necker cube. Sometimes face A is seen in front, other times face B is seen in front.*

To understand this prediction, it is helpful to consider the Necker cube in Figure 1. Sometimes you see the face labeled *A* in front; other times face *B*. When you don't

look, which face, *A* or *B*, is really in front? The answer, of course, is neither. There is no cube, and no front or back, unless you look and construct them.

Suppose you look away from the cube. Which face will be in front when you look back? You don't know. The best you can do is state probabilities that you will see either *A* in front or *B* in front. That is, when you don't look, your best guess about the cube that you'll see is a superposition of "*A* in front" and "*B* in front," with a probability attached to each element of the superposition. This is precisely the same as the "wavefunction" formalism used in quantum theory to describe physical systems between observations. That wavefunctions use complex amplitudes rather than probabilities is mere computational convenience; wavefunctions can equally well be written using standard probabilities, as is done in quantum Bayesianism (Fuchs, 2010).

If, when you look, you see face *A* in front then you know with probability one that *B* is behind; and vice versa. In other words, the states of the faces are entangled: Knowing the state of one face determines the state of the other. Thus, in the Necker cube, we have a model of superposition and entanglement in a macroscopic perception.

Concluding Thoughts

The idea that our perceptions might reflect poorly, or not at all, the true structure of objective reality has a long history going back at least to Plato's allegory of the cave,

according to which our perceptions are mere flickering shadows of reality cast on the wall of a cave by objects that remain unseen (for brief histories see, e.g., Koenderink, 2015; Mausfeld, 2015). The interface theory of perception contributes to this lineage of ideas by (1) using evolutionary game theory to demonstrate that veridical perceptions generically go extinct and (2) providing the metaphor of the user interface to help understand how non-veridical perceptions can be useful and fitness enhancing (Hoffman, 2000; 2009). Moreover, it provides a new formal framework for understanding perception that corrects and extends the current framework that treats perception as Bayesian decision making (Hoffman & Singh, 2012; Knill & Richards, 1996).

The interface theory of perception has radical implications for the notion of physical causation. If spacetime is a species-specific desktop and physical objects are species-specific icons in that desktop, then any apparent causal effects of physical objects in spacetime are a fiction.

The fiction is useful in everyday life. To compare, it's a useful fiction to think that when one drags a file icon to the trashcan it's the movement of the icon on the desktop, and its collision with the trashcan icon, that causes the file to be deleted. This is of course false. There is no feedback from the pixels of the icon to the computer. Similarly, for everyday life it's a useful fiction to think of physical objects as having true causal powers—to think that a bat hitting a ball can cause it to careen over a wall for a home run. But, if the interface theory of perception is correct, it's

nevertheless a fiction. Perception is not about seeing the true causal structure of reality, it's about having kids.

Recommended Reading

Hoffman, D. D., Singh, M., & Prakash, C. (2015a). The interface theory of perception. *Psychonomic Bulletin and Review*, 22, 1480-1506. A comprehensive presentation of the interface theory of perception.

Hoffman, D. D., Singh, M., & Prakash, C. (2015b). Probing the interface theory of perception: Replies to commentaries. *Psychonomic Bulletin and Review*, 22, 1551-1576. Replies to ten commentaries that canvas a wide range of natural objections to the interface theory of perception.

Mermin, N. D. (1985). Is the moon there when nobody looks? Reality and the quantum theory. *Physics Today*, 38(4), 38–47. A gentle introduction to research that investigates whether physical objects have definite values for dynamical physical properties when they are not observed.

Nowak, M. (2006). *Evolutionary dynamics: Exploring the equations of life*. Cambridge, MA: Belknap Press of Harvard University Press. The most accessible introduction to evolutionary game theory that I have found.

Declaration of Competing Interests

The author declares that he had no conflicts of interest with respect to his authorship or the publication of this article.

Funding

This work was supported by the Federico and Elvia Faggin Foundation. The content of this publication is solely the responsibility of the author and does not necessarily represent the views of the Federico and Elvia Faggin Foundation.

References

Bell, J. S. (1964). On the Einstein Podolsky Rosen Paradox. *Physics*, 1 (3): 195–200.

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

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. arXiv:1003.5209 [quant-ph]

Giustina, M., Mech, A., Ramelow, S., Wittmann, B., Kofler, J., Beyer, J., . . . Zeilinger, A. (2013). Bell violation using entangled photons without the fair-sampling assumption. *Nature*, 497, 227–230.

Hensen, B., et al. (2015). Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometers. *Nature*, 526, 682–686.

Hoffman, D. D. (2000). *Visual intelligence: How we create what we see*. New York: W. W. Norton.

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

Hoffman, D. D., Singh, M., & Prakash, C. (2015a). The interface theory of perception. *Psychonomic Bulletin and Review*, 22, 1480-1506.

Hoffman, D. D., Singh, M., & Prakash, C. (2015b). Probing the interface theory of perception: Replies to commentaries. *Psychonomic Bulletin and Review*, 22, 1551-1576.

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

Klimov, P.V., Falk, A.L., Christle, D.J., Dobrovitski, V.V., & Awschalom, D.D. (2015). Quantum entanglement at ambient conditions in a macroscopic solid-state spin ensemble. *Science Advances*, doi: 10.1126/sciadv.1501015.

Koenderink, J.J. (2015). Esse est percipi & verum factum est. *Psychonomic Bulletin and Review*, 22, 1530-1534.

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

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

Mausfeld, R. (2015). Notions such as “truth” or “correspondence to the objective world” play no role in explanatory accounts of perception. *Psychonomic Bulletin and Review*, 22, 1535-1540.

Mermin, N. D. (1985). Is the moon there when nobody looks? Reality and the quantum theory. *Physics Today*, 38(4), 38–47.

Mermin, N. D. (1990). What’s wrong with these elements of reality? *Physics Today*, 43(6), 9–11.

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

Szeliski, R. (2010). *Computer vision: Algorithms and applications*. New York: Springer.

Werner, J.S. & Chalupa, L.M. (2014). *The new visual neurosciences*. Cambridge, MA: MIT Press.