

Does Evolution Favor Accurate Perception?¹

Adriano Angelucci, Vincenzo Fano,* Gabriele Ferretti,** Roberto Macrelli,* and Gino Tarozzi**

** University of Urbino Carlo Bo*

*** Ruhr University Bochum*

Abstract

The currently mainstream view is that, in normal conditions, our perceptual representations are largely accurate, as natural selection tends to favor epistemically reliable perceptual systems. This latter assumption has been questioned by Donald Hoffman and his collaborators by drawing on the formal tools of evolutionary game theory. According to their model, an organism whose visual system were tuned to objective reality would be driven to extinction. We argue that their model fails to take environmental modifications into due account, and we show that, once such changes are incorporated into the model, the latter will predict that an organism whose visual representations are at least partially accurate will in fact be more successful from an evolutionary point of view.

Keywords: Perception, Perceptual strategies, Evolutionary game theory.

1. Introduction

The currently mainstream view among scientists studying perception is that, in normal conditions, our perceptual representations are largely accurate—i.e., that, to some extent, they do a good job at tracking the objective structure of the external world.² The view in question usually rests on a specific evolutionary assumption—i.e., that natural selection will in the long run favor individuals whose perceptual systems are epistemically reliable. Within the relevant literature it is indeed typically argued that if our perceptual representations were not somehow tuned to the objective structure of reality, evolutionary pressures would long have driven our species to extinction.³ In a series of papers, Donald Hoffman and his

¹ In this article we bring out what we take to be the main philosophical consequences of the two models presented in Angelucci et al. 2021.

² Cf., e.g., Marr 1982: 340, Trivers 2011: 2, and Pizlo et al. 2014: 227.

³ Cf., e.g., Geisler & Diehl 2003, and Yuille & Bülthoff 1996.

collaborators (henceforth, H&C) made use of evolutionary game theory in order to question this widely held assumption.⁴ Evolutionary games, in their view, would conclusively establish that our visual systems are in fact tuned to *utility*, not to objective reality. As a consequence, H&C maintain, we would have little or no reason to believe that our visual representations are always, or even usually, accurate.

The far-reaching philosophical implications of this purported fact about human vision can hardly be overstated. Arguably, if H&C's conclusion were to prove correct, then large swaths of contemporary epistemology and philosophy of mind would have to be called into question, to say the least. Naturalistic approaches to knowledge and justification, for instance, are more or less explicitly premised on the assumption that, in normal conditions, our perceptual systems and processes are generally reliable,⁵ and the same seems to hold for naturalistically minded accounts of the semantic content of our mental states.⁶ Moreover, H&C's conclusion, if true, would arguably lend significant support to the skeptical—yet nonetheless popular in some intellectual milieus—idea according to which empirical science would not in the end be entitled to any justified claims about what the external world is like, independently of the way in which it happens to be perceived or thought of by sentient beings.

In what follows, we intend to argue that H&C's epistemically grim conclusion is still far from being the only one licensed by the formal tools of evolutionary game theory. Our main goal will be to show that, contrary to their view, the mere fact that the complex evolutionary dynamics responsible for shaping our perceptual systems will in the long run increase our fitness does not entail that our visual representations will therefore be generally inaccurate. What led H&C astray, in our view, is that their model fails to take the relevance of environmental modifications into due account. As we will try to show, however, a model that incorporates a dynamic, rather than static view of the organism's environment, will predict that—up to a certain point—the acquisition of apparently useless information about said environment will in fact increase fitness.⁷ In particular, we suggest that this will be the case even when the organism which detects such apparently useless information and the one which does not make use of the same number of bits. Our model then suggests that, in general, an organism whose visual representations were at least partially accurate would be more successful from an evolutionary point of view.

⁴ Cf., in particular, Mark, Marion, and Hoffman 2010, Hoffman and Manish 2012, Hoffman, Manish, and Mark 2013, Hoffman, Manish, and Prakash 2015.

⁵ Alvin Goldman, the father of *process reliabilism*, found it plausible to suppose that “many cognitive functions subserving the attainment of true beliefs [...] were selected for in evolution because of their biological consequences, that is their contribution to genetic fitness” (Goldman 1986: 98, quoted in Stich 1990: 161).

⁶ Consider, e.g., the following two passages from Ruth Millikan and Daniel Dennett respectively: “The mechanisms in us that produce beliefs [...] all have in common at least one proper function: helping to produce true beliefs” (Millikan 1984: 317, quoted in Stich 1990: 162); “natural selection guarantees that *most* of an organism's beliefs will be true” (Dennett 1981: 75, quoted in Stich 1990: 55).

⁷As we shall see, the information in question is here said to be ‘apparently’ (as opposed to ‘actually’) useless in the sense that, by gathering it, the organism will incur costs which—while increasing its fitness *in the long run*—are bound to have an *immediate* negative impact in terms of fitness. Thanks to an anonymous referee for inviting us to clarify this point.

The plan is as follow. In the next section we will introduce a basic formal framework which allows us to define three different perceptual strategies, dubbed *realist*, *critical realist* and *interface* strategy respectively. This framework will also provide us with the means to gauge the accuracy of each strategy. In section 3 we will then consider and assess H&C's argument for the inaccuracy of our visual representations, according to whose general conclusion an interface strategy will in the long run clearly outcompete a critical realist one. In section 4 we put forward an alternative model in order to show that, once a biologically more realistic view of the environment is incorporated into the model, a critical realist strategy will in the long run outcompete an interface one. In section 5 we will sum up our considerations and draw some conclusions.

2. Perceptual Strategies

In line with most contemporary philosophical theories of perception we will assume that perception is at bottom a representational process, i.e., that our perceptual systems represent reality by ascribing various features to individual objects as well as to the visual scene as a whole.⁸ As H&C focus on vision, our first task will consist in developing a plausible and empirically testable model of visual perception, accordingly conceived as a process whereby a given environmental stimulus causally interacts with our visual system, thereby giving rise to a more or less accurate representation of its source—i.e., a *visual representation*. So let us do just that.

Our model—just as any model—will inevitably involve a fair amount of idealization. So let us begin by thinking of an organism's environment as a given set E of features. Every subset of E can then be seen as a stimulus capable of causing in the organism a corresponding subset of a further set V of visual representations. Let us now call T_E and T_V the “best possible theories” of, respectively, E and V , and let us further conveniently suppose that these two theories are developed enough to possess their respective state-spaces S_{TE} and S_{TV} .⁹ By so doing, we can then let a *representation function* F stand for the organism's ability to visually represent its environment, and an inverse *causal function* Q stand for the environment's causal effects on the organism's visual system—whereas F will map S_{TV} regions onto S_{TE} ones, Q will map S_{TE} regions onto S_{TV} ones. A *perceptual strategy*, at this point, will be a composite function FQ that maps S_{TE} regions onto S_{TE} ones.

We can now provide an exact definition of three distinct perceptual strategies that, following H&C, we may call *realist*, *critical realist*, and *interface* respectively.

⁸ Cf., e.g., Nanay 2013, Siegel 2006, Brogaard 2014, and Ferretti & Zipoli Caiani 2019. In spite of various interesting attempts at developing nonrepresentational views of perception (cf., e.g., Noë 2004, Chemero 2009, and Hutto and Myin 2013), representationalism still remains the dominant view on the matter, and this is arguably mainly due to the undeniable explanatory advantages of the latter (cf. Pautz 2010, Nanay 2013), especially in case of the study of perceptual reality (Ferretti, forthcoming). It is however clear that, if perception were direct even in a weak sense, then H&C would be a fortiori wrong.

⁹ We hasten to add that, for the purposes of the present argument, there is no need to think of our ‘best possible theories’ as actual scientific theories— T_E and T_V are rather intended as merely useful fictions whose sole purpose in what follows will be to illustrate our proposal concerning the measurement of visual representations' accuracy. Thanks to an anonymous referee for inviting us to clarify this point.

Letting r_{SE} stand for a given region of S_{TE} , we can postulate that a perceptual strategy will be a *realist* one if $FQr_{SE} = r_{SE}$ —i.e., if our visual representations perfectly mirror the environmental stimuli that give rise to them. A strategy will instead be a *critical realist* one if there is at least a subspace of S_{TE} (call it S'_{TE}) within which, as it were, realism holds—i.e., within which, if $r_{S'E}$ is a region of S'_{TE} , and $S'_{TE} \subset S_{TE}$, then $FQr_{S'E} = r_{S'E}$. A strategy will finally be an *interface* one if $S'_{TE} = \emptyset^{10}$ (cf. Fig. 1).

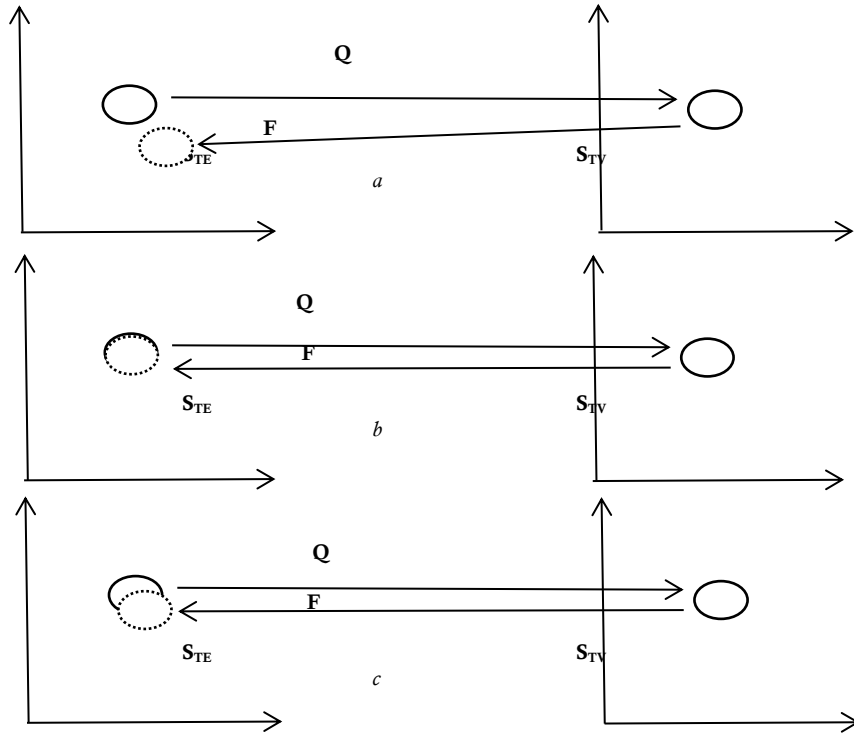


Fig. 1 – A certain set of stimuli—represented by the ellipse in the state space of the environment (S_{TE})—causes the changes described by function Q , i.e., a certain modification in the visual field of the organism (S_{TV})—the ellipse on the right. Such modification in turn constitutes an attempt to represent (F) the initial set of stimuli—the dotted ellipse on the left. There will hence be three possible situations: *a*: Interface strategy, *b*: Realist, and *c*: Critical Realist.

In light of the above, we can now think of the accuracy of our visual representations as a correspondence between the two state-spaces S_{TE} and S_{TV} . In particular, our framework will allow us to measure such accuracy through the *distance* d_{EV} between the (objective) conjunct probability measure on S_{TE} —i.e., μ_{EE} —and the conjunct probability measure on S_{TV} —i.e., μ_{VV} . This last point perhaps requires

¹⁰ However, an organism implementing an interface strategy will still be sensible to *environmental discontinuities*, and it will therefore preserve a residual representational capacity. Our definition is only meant to capture the idea that the representational contents of a perceptual system implementing such a strategy will be so far removed from a completely accurate representation of the environment as to have virtually zero accuracy. Thanks to an anonymous referee for inviting us to clarify this point.

some clarification. Perceptual strategies have earlier been defined relative to the state-space of our “best possible theory” T_E of the environment—i.e., S_{TE} . It must be kept in mind, however, that neither an individual visual representation belonging to S_{TV} , nor an individual stimulus belonging to S_{TE} are themselves directly accessible for us. As a consequence, the only viable way to assess the relevant distance (d_{EV})—and to thereby decide which one of the three perceptual strategies is actually being implemented—will be to rely on the conjunct probabilities of distinct stimuli, and of distinct visual representations respectively. In operational terms, then, the relevant question will not have the form: What is the probability that the organism will experience a visual representation of red, given that a red distal stimulus is being instantiated? But rather: What is the probability that it will experience two adjacent visual representations (e.g., a green and a red one), given that two corresponding adjacent distal stimuli—a green and a red one—are being instantiated? It is the answer to this latter question that will in fact give us a measure of the visual system’s accuracy.

The distance d_{EV} can then be normalized so that, when two measures are the same, its value will be “0”, and when two given elements x_E and y_E of an algebra defined on S_{TE} are such that “ $\mu_{VV}(x_V, y_V) = 1 - \mu_{EE}(x_E, y_E)$ ” its value will be “1”. At this point it will be reasonable to posit that a critical realist strategy will determine a value of $d_{EV} \leq 0.5$, an interface strategy will determine a value of $d_{EV} > 0.5$,¹¹ and a realist strategy will hold when $d_{EV} = 0$. With this formalism in place, let us now move on to consider H&C’s main argument for the purported inaccuracy of our visual representations by focusing on the interplay amongst the perceptual strategies defined above.

3. The Case for Interface

Evolutionary game theory is arguably the best way to predict the evolution of a discrete phenotypic trait whose fitness depends on its frequency within a population.¹² The general idea is that a trait’s fitness could be *affected* by its frequency. Consider, for instance, the random appearance, on a butterfly’s wing, of a pigmented region which just so happens to mimic the eye of a snake. This random mutation will presumably have the immediate effect of decreasing the butterfly’s chances to be eaten by a bird, thereby increasing its fitness. The mutation in question, however, will only have this effect (i.e., misleading birds into believing that a butterfly is a snake) if it makes its appearance in a limited number of butterflies.¹³ In our present case, the trait will of course be a perceptual strategy coexisting with other strategies, and whose fitness will therefore also depend on the frequency of its rivals. As we anticipated above, H&C hold that fitness-maximization is bound to have a negative impact on the overall accuracy of our visual representations, as an interface strategy, in their view, would clearly outcompete—and hence, in the long run, drive to extinction—a critical realist one.¹⁴ In order to substantiate

¹¹ For the sake of simplicity, we are here focusing on *binary* features only (such as, e.g., black/white). With respect to such features, it seems reasonable to assume that getting them right 50% of the time is tantamount to having zero information about the environment.

¹² Cf. Rice 2004: 263.

¹³ We would like to thank an anonymous referee for encouraging us to clarify this point.

¹⁴ Cf., e.g., Mark et al. 2010: 504.

this claim, they ask us to consider the following evolutionary game in which all three of our perceptual strategies—i.e., *realism*, *critical realism*, and *interface*—compete with each other.

The playing field features three different territories, and only one resource whose values range from 1 to 100. Utility—which is proportional to fitness—is represented by a Gaussian with its peak at 50, and it is therefore not proportional to the quantity of resource to be found on each territory. Now, whereas the *realist* strategy will gather all of the available information, the *critical realist* one will instead only rely on three visual representations (e.g., three different colors standing for different resource quantities), and the same will be the case for the *interface* strategy. The difference between the two latter strategies lies in the way in which the three colors are used (cf. Fig. 2).

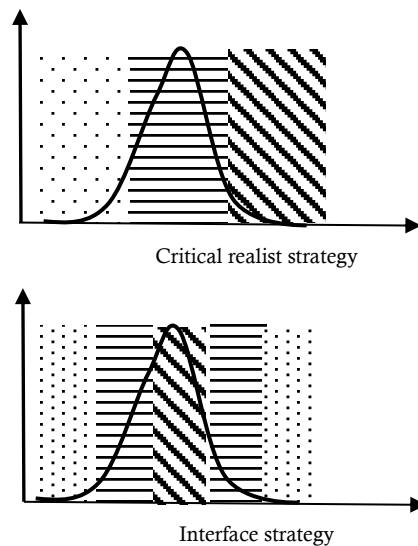


Fig. 2 – Critical realist strategy (above); Interface strategy (below). On the x-axis the quantity of resource; on the y-axis the utility. The difference between critical realist and interface strategies is expressed through diverse distributions of colors. The different colours are represented through the diverse types of filling: points, horizontal lines and diagonal lines. It is evident that the use of colors in the interface strategy is more useful—in terms of fitness—than its counterpart in the critical realist strategy.

As we can see from the two graphs in Fig. 2, while the critical realist strategy will disregard the utility curve and simply associate the three colors with the increasing quantity of the resource, the interface strategy will keep track of utilities only. Now, as resource quantity and utility are non-monotonically related, each strategy will incur the costs associated with the process of gathering information about the environment and calculating its corresponding utility. It follows that the interface strategy will soon outcompete the critical realist one.

On closer inspection, however, this stage seems clearly and intentionally set to put critical realism at a disadvantage. Indeed, by keeping perceptual complexity fixed, the interface strategy will obviously have a running start. And yet, as we shall presently see, additional considerations may easily turn the tables on the

interface strategy. As it has been objected, for instance,¹⁵ organisms often tend to homeostasis, and hence have an interest in knowing whether the quantity of a given resource happens to be above or below a certain threshold. When this is the case, the critical realist strategy will have an advantage over its interface counterpart. While this objection seems to point in the right direction, in the next section we will argue that H&C's view of human perception is beset by a more fundamental limitation of their model.

4. Making Room for Change

To shed light on what we regard as the main drawback of H&C's model, let us now consider the following simple case. Suppose that the organisms competing in our game are sparrows, and the resource are little worms. Given that worms evolve, we can easily imagine that a small and apparently inconsequential random mutation will at some point significantly decrease the size of a small number of individuals in their population. And we can further imagine that—as the sparrows' foraging strategy tends to zero in on bigger preys—the new trait will spread rapidly across the worms' population. This latter fact will in turn obviously alter the ratio between the utility of the resource and its quantity (expressed in number of worms). The point now is that, under the imagined circumstances, a sparrow implementing an interface strategy will accordingly still “think” that the same number of worms is needed in order to maximize utility, and will hence end up lagging behind in terms of fitness.¹⁶ Its critical realist competitor, on the other hand, will “know better” and accordingly move to an area where either more or bigger worms are to be found.

Cases similar to the above, we believe, clearly suggest that slight modifications in the environment can bring about serious disadvantages for organisms implementing an interface strategy. Indeed, by completely disregarding apparently useless information—such as, e.g., worms' size—the organisms in question will be utterly unresponsive to possible environmental modifications that do however have a significant impact on utility. Our main point is hence that, given a static environment, a strategy targeting utility will clearly outcompete one aimed at representing reality. In a situation where the environment changes, however, the opposite will be the case.

This can be shown by means of a very simple model in which an interface strategy will initially prevail over a critical realist one and yet this trend will reverse in due time because of modifications in the environment. According to the model in question, in other words, when the environment is held fixed and each

¹⁵ Cf. Anderson 2015.

¹⁶ The reason is that, immediately after the environmental change has taken place, a sparrow implementing an interface strategy will still lack the information that, in order to maximize utility, it will need to eat more worms. This is due to the fact that it will presumably take generations for a visual system implementing such a strategy to *retune* to the new utility distribution. As a real-life example of this dynamic, we can think of the extinction of dinosaurs after an asteroid hit the Yucatán Peninsula 66 million years ago thereby causing vast and sudden environmental changes. Their extinction was due to their incapacity to change rapidly their genetic code to face the new situation. Thanks to an anonymous referee for inviting us to clarify this point.

organism has the same number of bits at its disposal, a strategy aimed at increasing utility will outcompete one aimed at representing reality. As the environment changes, however, the opposite will be the case.

Let us consider a genus ω , divided in two species ω_{IF} and ω_{CR} —each implementing an interface, and a critical realist strategy respectively—and let us suppose that the environment within which the two species compete features two resources, x and y , whose density happens to fluctuate over time with a slight difference of phase. Let us then additionally suppose that the overall resource utility is not proportional to the mere *sum* of x and y 's density, but that it also depends on a further term related to the *difference* between their phases. Consider now the different ways in which ω_{IF} and ω_{CR} will respectively go about gathering information. Whereas ω_{IF} will approach this task just by assessing the exact resource utility of an initial environmental situation, ω_{CR} will instead at least keep an approximate track of the density fluctuation in the two resources. It can be shown that, in a similar setting, ω_{IF} would initially outcompete ω_{CR} , as its perceptual strategy will for a while do a better job at tracking utility. As time goes by, however, ω_{CR} ' rough estimate of x and y 's density fluctuations—i.e., its relative responsiveness to environmental changes—will prove extremely valuable, as it will allow for a much better long-term assessment of their utility. At the end of the day, then, ω_{CR} will be better off than ω_{IF} from an evolutionary point of view.¹⁷

While this simple model is admittedly limited in scope, the assumptions upon which it rests seem quite reasonable. We take those assumptions to be the following:

- (1) Environments change.
- (2) Many environmental features display an oscillating pattern.
- (3) Utility is not in general the mere sum of two such features.
- (4) A constant utility function is not appropriate to represent utility in a changing environment.
- (5) Knowledge of the environmental features' variation, while itself insufficient to locate the real utility function, nonetheless seems a reasonable starting point to assess utility in a changing environment.

5. Conclusions

If perception is the only way to acquire information about our environment and it turns out to be not even partially accurate, then investigating *Homo sapiens* and its environment would amount to merely inspecting our subjectivity. Yet modern science's moral and cognitive mission also consists of pursuing fallible and revisable attempts at formulating justified hypotheses about *Homo sapiens*, its origins and the world it inhabits. Many cultural milieus encourage the idea that empirical science cannot make any justified claims about the external world, independently of the way in which that world is perceived or thought of. If perception were completely inaccurate, this idea would be reinforced. We believe, however, that whether and to what extent human perception accurately represents the world is an epistemological matter which can be empirically investigated at least indirectly by using evolutionary mathematical models. We showed the limits of H&C's attempts at establishing the negative impact on fitness of an accurate representation

¹⁷ Cf. Angelucci et al. 2021 for the mathematical derivation of this result.

of the world. Our model is clearly only a sketch at this stage, and it certainly requires further development. Indeed, we are confident that, given reasonable assumptions concerning what should count as an accurate perceptual representation, it should be possible to empirically investigate the comparative fitness of different perceptual strategies along the lines suggested by H&C. We also believe, however, that such investigation should carefully take into account modifications in the environment.¹⁸

References

- Anderson, B.L. 2015, "Where Does Fitness Fit in Theories of Perception?", *Psychonomic Bulletin & Review* 22, 6, 1507-11.
- Angelucci, A., Fano, V., Ferretti, G., Macrelli, R., and Tarozzi, G. 2021, "Evolutionary Dynamics and Accurate Perception: Critical Realism as an Empirically Testable Hypothesis", *Philosophia Scientiae*, 25: 157-78.
- Brogaard, B. 2014 (ed.), *Does Perception Have Content?*, New York: Oxford University Press.
- Chemero, A. 2009, *Radical Embodied Cognitive Science*, Cambridge, MA: MIT Press.
- Ferretti, G. (forthcoming), "For an Epistemology of Stereopsis", *Review of Philosophy and Psychology*.
- Ferretti, G. and Zipoli Caiani, S. 2019, "Between Vision and Action: Introduction to the Special Issue", *Synthese*, Special Issue, DOI: 10.1007/s11229-019-02518-w
- Geisler, W.S. and Diehl, R.L. 2003, "A Bayesian Approach to the Evolution of Perceptual and Cognitive Systems", *Cognitive Science*, 27, 3, 379-402.
- Hoffman, D.D. and Manish, S. 2012, "Computational Evolutionary Perception", *Perception*, 41, 9, 1073-91.
- Hoffman, D.D., Singh, M., and Mark, J.T. 2013, "Does Evolution Favor True Perceptions", in Rogowitz, B.E., Pappas, T.N., and de Ridder, H. (eds.), *Proceedings of the SPIE 8651, Human Vision and Electronic Imaging*, XVIII, 865104, DOI: 10.1117/12.2011609
- Hoffman, D.D., Manish, S., and Prakash, S. 2015, "The Interface Theory of Perception", *Psychonomic Bulletin & Review*, 22, 6, 1480-1506.
- Hutto, D.D. and Myin, E. 2013, *Radicalizing Enactivism: Basic Minds without Content*, Cambridge, MA: MIT Press.
- Mark, J.T., Marion, B.B., and Hoffman, D.D. 2010, "Natural Selection and Veridical Perceptions", *Journal of Theoretical Biology*, 266, 4, 504-15.
- Marr, D. 1982, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, San Francisco: Freeman.

¹⁸ We would like to thank the audience at the 3rd conference of the P.R.I.N. "The Manifest Image and the Scientific Image": *Models, Structures and Representation*, held in Urbino in June 2022, where the ideas contained in this article were first presented. We also thank three anonymous reviewers for helping us improve on an earlier version of this manuscript. We acknowledge support by the Italian Ministry of Education, University and Research through the PRIN 2017 project "The Manifest Image and the Scientific Image" prot. 2017ZNNW7F_004. Gabriele Ferretti also acknowledges support from a Humboldt Fellowship, hosted by Professor Albert Newen at the Institute for Philosophy II, Ruhr University Bochum, Germany.

- Nanay, B. 2013, *Between Perception and Action*, Oxford: Oxford University Press.
- Noë, A. 2004, *Action in Perception*, Cambridge, MA: MIT Press.
- Nowak, M.A. 2006, *Evolutionary Dynamics: Exploring the Equations of Life*, Cambridge, MA: Belknap Press.
- Pautz, A. 2010, "Why Explain Visual Experience in Terms of Content", in Nanay, B. (ed.), *Perceiving the World*, New York: Oxford University Press, 254-309.
- Pizlo, Z., Li, Y., Sawada, T., and Steinman, R.M. 2014, *Making a Machine That Sees Like Us*, New York: Oxford University Press.
- Rice, S.H. 2004, *Evolutionary Theory: Mathematical and Conceptual Foundations*, Sunderland: Sinauer.
- Siegel, S. 2006, "Which Properties Are Represented in Perception?", in Gendler, T.S. and Hawthorne, J. (eds.), *Perceptual Experience*, Oxford: Oxford University Press, 481-503.
- Stich, S.P. 1990, *The Fragmentation of Reason: Preface to a Pragmatic Theory of Cognitive Evaluation*, Cambridge, MA: MIT Press.
- Trivers, R. 2011, *The Folly of Fools: The Logic of Deceit and Self-Deception in Human Life*, New York: Basic Books.
- Yuille, A. and Bülthoff, H.H. 1996, "Bayesian Decision Theory and Psychophysics", in Knill, D.C. and Richards, W. (eds.), *Perception as Bayesian Inference*, Cambridge: Cambridge University Press, 123-62.