Analysis Advance Access published February 26, 2015

BOOK REVIEW

The Innocent Eye: Why Vision Is Not a Cognitive Process

By Nico Orlandi

Oxford University Press, 2014. 272 pp. £29.99

In *The Innocent Eye*, Nico Orlandi argues that vision is not a cognitive process. In particular, she argues that forming subject-level visual representations that are available for reasoning should not itself be understood as a process of inference. This comes to the claim that vision (properly so-called) is a process that produces representations but is not best understood as a process that uses representations.

In cognitive science, the view that vision involves processes like inferences that operate on representations is most closely associated with the work of David Marr (1982) and his philosophical admirers.[1] According to Marr, early vision involves a series of computations that proceeds in stages, whereby a representation of the stimulation of the retinal sensor array is transformed into a representation of luminance boundaries, then of edges, then into a partial model of the world, and finally into a 3-D model of the world. But Orlandi observes that Marr's theory is not the only or most prominent approach employed by vision scientists. There is now some impressive evidence that vision works rather differently from the way Marr supposed, and Orlandi argues that the best alternative frameworks do not construe vision as a process of computation over representations.

To achieve her goals, Orlandi must convince us of at least two things. First, that early vision does not employ fancy kinds of representations and inferences, of the sort that would be readily recognized as representations and uncontroversially accepted as cognitive. That's the easy part. Second, Orlandi has to convince us that vision does not employ more rudimentary forms of representation and computation that, while not canonically cognitive, nevertheless may count as such (cf. Milkowski 2013). Much of the book is given to this second task, which itself involves two complementary forms of argument. One is the case-by-case evaluation of explanations of visual processes, arguing that they do not employ representations. The other is a general argument that the purportedly 'representational' states found in visual processing do not count as genuinely representational after all. Orlandi argues that such states fail to satisfy some widely accepted criteria on representations; and that they only count as representations on problematically 'deflationary' theories of representation.

The crux of my concern with Orlandi's argument is that what she identifies as a defect in 'deflationary' representational theories is supposed to be a feature of those theories. 'Deflationary' representationalist theories are those that build representations out of tracking relations in the world, such as Fodor's asymmetric dependency theory (1990) and Dretske's teleofunctionalist theory (1981, 1991, 2002).

1 Orlandi allows that the textbook interpretation of Marr may not be the correct one; but I will follow the textbook interpretation.

Analysis Reviews Vol 0 | Number 0 | 2015 | pp. 1–3 doi:10.1093/analys/anv002

On Dretske's version, many processes in the world carry information, so many states of things can serve as indicators of what is going on in the world. These are tracking processes. According to Dretske (2002), such processes can be intentional—they are about the information that they carry or about the things in the world about which they carry information. But this sort of intentionality is not sufficient for representation, as he believes that representation requires the possibility of misrepresentation. For Dretske, an information-carrying system, that is, a tracking or indictor system, becomes a representational system when it acquires the function of indicating. The importance of representations having the function of indicating is not to be overlooked. It is because they have the function of indicating that they can misrepresent, that is, they can fail to do what it is their function to do; and in virtue of that fact, representations can be decoupled from their objects.

Orlandi argues that decoupling is a criterion for representation, but that Dretske-style representations are not sufficiently decoupled. Representations 'are states that stand for something that is not *continuously* present to the organism – that is, not, at all times, impinging on the organism's senses' (122, emphasis in original). But, Orlandi argues, the detector states in the early visual system are not decoupled from their objects in this way; instead, they perpetually respond to changing stimuli. There is something sensible about this idea, but I don't think that Orlandi has the right theory. She allows that decoupling can occur whenever variation in the proximal stimulus is neglected, when the system 'takes' the stimulus to be stable despite impinging flux: 'My proposal is that this *taking* constitutes an early representational capacity. It is a capacity that involves abstraction. Achieving stability and constancy in perception are ways in which this abstraction takes place. The behavior-guiding states that stand for constant and stable properties are plausibly representations' (130–131, emphasis in original). But if this basic sort of abstraction is all that is required to count as a representation, then contrary to Orlandi, representation occurs at the earliest stages of visual processing. Indeed activity at the retinal transducers – the rods and cones – will count as representation. These cells are tiny abstractors, not responding to every impinging photon but summing and averaging over time, and abstracting from (in Orlandi's sense of 'ignoring') wide variation in the wavelength and intensity of the optical signal. It seems to me that Orlandi has made it too easy for early visual states to count as representations.

Of course Orlandi can and does argue that this kind of stability and abstraction is not yet sufficient for genuine representation. But then it is hard to see how she will draw a distinction between the kinds of abstraction and stability that make for representations and those that do not without importing substantial assumptions about representations and their objects. For example, Orlandi requires that representations be available to system-level processes, thus denying that there can be merely peripheral and subpersonal representations. But that is to raise the bar quite high on representations. The defender of representations can respond that even if those assumptions apply to fancy kinds of representation, they need not apply to the simple kinds of representation that are postulated in the early visual system.

Orlandi repeatedly worries that 'deflationary' accounts of representation obscure the distinction between representation and tracking, or the distinction between inference and association, and thereby trivialize the representationalist claim. But one might say that the whole idea of Dretske's program is to trivialize the representationalist claim at the periphery, in order to show how fancy representations can be built

up out of simple indicator states. The distinction between representation and tracking is not an exclusive one; the question, rather, is to explain which tracking states get to be representations.

As I largely agree with Orlandi about the state of vision science, our disagreement might seem to be merely verbal. She is sensitive to the worry that the argument concerns only how the label 'representation' can be applied. But I think the disagreement is substantial, and it concerns our methods for discerning the ontologies of the sciences. Orlandi argues that we do not 'need' to treat information bearing states and processes in the visual system as representations, because all of the explanatory work can be done by appeal to the functional rather than representational features of those states and processes. But it's unclear that cognitive scientists do or should embrace an explanatory exclusion principle to the effect that the availability of functional or mechanistic explanations rules out the legitimacy of representational explanations. Rather than asking whether it is explanatorily necessary to attribute representational content to processes in the early visual system, we might ask the more general question of whether it is fruitful to do so. And when we address the question of the fruitfulness of representational explanations of vision, we may wonder whether the answer will be determined exclusively by the examination of the mechanisms of human early visual processing. Representationalist explanations might be justified by their utility for explaining how those mechanisms of biases and constraints came to be, or what those systems have in common with other human perceptual systems and with physiologically distinctive visual systems in other creatures.

I am sympathetic to Orlandi's viewpoint. *The Innocent Eye* draws the attention of philosophers to research that they have mainly neglected, and challenges the computationalist consensus that has been mainly taken for granted since philosophers learned about Chomsky, Pylyshyn, and Marr. Advocates of representationalist and computationalist approaches are not without resources to respond. But it will not be sufficient to bang the table and insist that computational or representational cognitive science is the only game in town.

THOMAS POLGER
Department of Philosophy
University of Cincinnati
Cincinnati, OH 45221–0374, USA

References

Dretske, F. 1981. *Knowledge and the Flow of Information*. Cambridge, MA: The MIT Press.

Dretske, F. 1991. *Explaining Behavior*. Cambridge, MA: The MIT Press.

Dretske, F. 2002. *A Recipe for Thought. In* Philosophy of Mind: Classical and Contemporary Readings, ed. D. Chalmers, 491–499. New York: Oxford University Press.

Fodor, J. 1990. *A Theory of Content and Other Essays*. Cambridge, MA: The MIT Press.

Marr, D. 1982. *Vision*. Cambridge, MA: The MIT Press.

Milkowski, M. 2013. *Explaining the Computational Mind*. Cambridge, MA: The MIT Press.