

# What Should a Theory of Vision Look Like?

Anne Jaap Jacobson

*This paper argues for two major revisions in the way philosophers standardly think of vision science and vision theories more generally. The first concerns mental representations and the second supervenience. The central result is that the way is cleared for an externalist theory of perception. The framework for such a theory has what are called Aristotelian representations as elements in processes the well-functioning of which is the principal object of a theory of vision.*

*Keywords: Block; Constitution; Neuroscience; Noë; Normativity; Object Constancy; Perception; Philosophy of Mind; Supervenience; Vision*

## 1. Introduction

This paper supports two major revisions in standard philosophical views about what a theory of vision should look like. The first provides an account of neural representations as what we will call ‘Aristotelian representations’, in contrast to ‘Fodorian representations’, which are representations in the sense usually meant by philosophers. The second is the claim that a theory of vision ought to be normative. A central result is the framework for an externalist account of perception.

## 2. Representations

It is a useful simplification to see Western philosophy as developing two main accounts of the mind’s relation to the world around it. The familiar one, which prevails in contemporary philosophy, has the mind directed toward objects

---

Anne Jaap Jacobson is Director at the Center for Neuro-Engineering and Cognitive Science at the University of Houston.

Correspondence to: Anne Jaap Jacobson, Professor of Philosophy and Electrical and Computer Engineering, Department of Philosophy, University of Houston, Houston, TX 77204-3004, USA. Email: [ajjacobson@uh.edu](mailto:ajjacobson@uh.edu)

outside it. It has truth-evaluable states with content that is about the world; such states are ‘mental representations’ (Crane, 2003; Elder, 1998; Fodor, 1975, 1990, 2006; Peacocke, 1992; Ramachandran, 2004; Rowlands, 2006).

A second and far older picture has states of the world getting into the mind. In its earliest and most problematic form, perception of a colored object involves getting that color inside one. Aristotle’s account is the best known of the ancient ones, and so we can say that on this model, the mind has ‘Aristotelian representations’. In contrast, the account familiar today ascribes to the mind what we can call ‘Fodorian representations’.

We can find the older view in ancient philosophy:

By a ‘sense’ is meant what has the power of receiving into itself the sensible forms of things without the matter . . . (Aristotle, *De Anima*, II.12, 46). The thinking part of the soul must therefore be, while impassible, capable of receiving the form of an object; that is, must be potentially identical in character with its object without being the object. Mind must be related to what is thinkable, as sense is to what is sensible. (III. 4,55)

And in medieval philosophy:

Intelligible forms are related to the intellect as sensible forms are related to the sense. But a sensible form is not that which is sensed; rather, it is that by which a sense senses. Consequently, an intelligible form is not that which is understood, but that by which the intellect understands. (Aquinas, *Summa*, Pr 1, Q.85, Art 2, p. 433)

And in recent cognitive neuroscience:

**Representation:** Properties of the world that are manifested in cognitive systems (mental representation) and neural systems (neural representation). (Ward, 2006, p. 33)

Theories using Aristotelian representations have two important dimensions along which they may differ: *What* gets realized in the mind and *how* it is realized. For Aristotle and Aquinas, forms are realized in the mind and the way they are realized is the way appropriate to their being so realized; i.e., they are not realized in the mind in the way they are realized in the material world. This picture will not survive into contemporary cognitive neuroscience. What is common in the members in this range of theories is the idea of a representation as something like a copy, example, or sample.

Accordingly, we can think of the contribution of an Aristotelian representation as like that of an example. Consider learning some movement in Tai Chi: it is very difficult for the novice to learn to perform one from reading a verbal description and fairly easy to do so from seeing an example. Or one might give a companion animal a sample of a new food to see if she likes it; what is important in these cases is the realization or re-realization of relevant properties. There is a big difference between finding out about something through descriptions or other verbal information, and finding out about it by having an instance of the relevant properties, though in human discourse, we may need a verbal setting to know which of the features the sample

realizes are relevant. Fodorian representations have their place in a linguistic model of the mind, while Aristotelian representations work with a different model, one in which the mind gets tokens of properties of the world manifested in it.

As we will see below, Aristotelian representations do not provide aboutness or satisfaction conditions, which is one reason why, understood as the topic of resemblance theories of representation, they are rejected as inadequate Fodorian representations. Nonetheless, there are good reasons for using the word ‘representation’ in talking about them. First of all, they occupy a position in thousands of years of philosophy that is analogous to the one that is today occupied by Fodorian representations. They are what mediates the mind’s contact with its environment. Secondly, historical philosophers call them representations. Hume (2000) is clear about this:

The full examination of this question is the subject of the present treatise; and, therefore, we shall here content ourselves with establishing one general proposition, *That all our simple ideas in their first appearance, are derived from simple impressions, which are correspondent to them, and which they exactly represent.* (T 1.1.1.7; SBN 4)

And the relation between an impression and the representing idea is a particular kind of resemblance; that is, Aristotelian co-instantiation:

The first circumstance that strikes my eye, is the great resemblance betwixt our impressions and ideas in every other particular, except their degree of force and vivacity. The one seems to be, in a manner, the reflection of the other; so that all the perceptions of the mind are double, and appear both as impressions and ideas. When I shut my eyes, and think of my chamber, the ideas I form are exact *representations* [italics added] of the impressions I felt; *nor is there any circumstance of the one, which is not to be found in the other* [italics added]. In running over my other perceptions, I find still the same resemblance and representation. Ideas and impressions appear always to correspond to each other. (T 1.1.1.3; SBN 2–3)

Accordingly, Aristotelian representations have the prior claim to the term. Relatedly, they are a much closer fit to the OED’s definition of representation than are Fodorian representations. Finally, and most importantly for us, they are well entrenched in common discourse and are often what is referred to by the use of ‘representation’ in cognitive neuroscience. They are representations in a different sense of the word, but it is one that has widespread use.

The thesis that the representations that neuroscientists appeal to are—or are often—Aristotelian representations needs explanation and defense. It needs explanation because there are different things it could be claiming. One thing it might assert is that neuroscientists routinely explain ‘representation’ in these terms. As far as I know, such a claim is false. A second version is that seeing them as Aristotelian representations provides a consistent interpretation of what the neuroscientists are saying. This is also false, principally because some neuroscientists, unlike Ward quoted above, have endorsed philosophers’ formal theories of representations. My thesis is rather that when one looks at how the term ‘represent’

and its cognates function in the theories, the interpretation of them as standing for Aristotelian representations is often better than the interpretation of them as standing for Fodorian representations. And it is better because the representations themselves function as Aristotelian representations. That is, their contribution is explained better by seeing them as Aristotelian representations.

In what follows, I am going to continue to simplify the issues by speaking as though a theory ascribes just to Aristotelian representations or just to Fodorian ones. It may be that the better theory eventually invokes Aristotelian representations in some cases and Fodorian ones in others. This is not an issue we will try to resolve here, since this paper does not attempt to address as one the various areas in cognitive neuroscience.

I have argued in other papers for an interpretation of uses of ‘representation’ as ‘Aristotelian representation’ (though not always by that name) in both recent neuroscience and in the history of philosophy. Jacobson (2003) looks at several areas of cognitive science, while Jacobson (2007) concentrates on human mirroring systems and argues that ‘action representations’ and ‘neural representations of emotion’ as used in the recent literature are better understood as Aristotelian. The latter paper also initiates an examination of the ways in which the representations discussed can contribute to our social knowledge without being Fodorian representations. Jacobson (2006) briefly explores the Humean use of ‘represents’ in an Aristotelian sense.

It is, however, interesting to take a further look at some of the literature. It is not just that philosophers tend to assume their own interpretation of ‘representation’ when they see it. There is the further fact that when neuroscientists explain what representations are, they may well explain them as Fodorian representations, though interestingly enough not always consistently. Thus (Ramachandran, 2004) tells us:

The first step we must take toward understanding perception is to forget the idea of images in the brain and think instead of transforms or symbolic representations of objects and events in the external world. Just as little squiggles of ink called writing can symbolize or represent something they don’t physically resemble, so the action of nerve cells in the brain, the patterns of firing, represent objects in the external world. Neuroscientists are like cryptographers trying to crack an alien code, in this case the code used by the nervous system to represent the external world. (p. 24)

Most of the passage appeals to a language based conception of representation that is very Fodorian. Not all of it, however. “Transforms” is a mathematical term and it applies to the result of subjecting something to a transformation. Since it is mathematical, we are considering a quantitative account of the “patterns of firing” to be derived from quantitative accounts of “objects and events in the external world.” And this is at least close to the Aristotelian idea that sensory features get realized in our sensory system in a different way from how they are realized in the environment.

Ramachandran has given us an overt contrast between symbolic coding and resemblance, and he says that the patterns of firing are like the first, not the second. But he also gives us a third option: The patterns of firing can be described as

transformations of features in the external world. This picture does not fit the symbolic coding model, since typically “little squiggles of ink” are not mathematical transforms either. But the idea that patterns of firing are transformations of external features is a familiar idea in the Aristotelian tradition.

In addition, it is not unusual to find Aristotelian representations accompanied by a linguistic model. For example, in “Cortical *representation* [italics added] of the sensory dimension of pain” (Hofbauer, Rainville, Duncan, & Bushnell, 2001), we find talk that is strongly indicative of Aristotelian representations, along with talk about ‘coding’, which, if taken as it standardly is in philosophy, seems to take us to the linguistic model Ramachandran inconsistently invokes. Such an interpretation cannot be correct, however, since the article is clearly about how the dimensions of pain itself are *realized* in the brain; it is *not* about neural events that manage to be *about* pain.

In contrast, Palmer (1999) consistently explains representation in terms that make them Aristotelian representations. That is, his representations are isomorphic to, or have the same structure as, what they represent (p. 658). Palmer’s work is significant in the present context, since he is cited by Block (2005) as a proponent of a standard internalist view. Nonetheless, as we will see below, his views clash with Block’s at several important junctures.

Further, the cognitive neuroscientific literature on mirroring systems is replete with a use of ‘representation’ that is overtly explained in terms of the Aristotelian picture of the re-instantiation of properties. Thus:

What neural mechanism underlies the capacity to understand the emotions of others? Does this mechanism involve brain areas normally involved in experiencing the same emotion? We performed an fMRI study in which participants inhaled odorants producing a strong feeling of disgust. The same participants observed video clips showing the emotional facial expression of disgust. Observing such faces and feeling disgust *activated the same sites* [italics added] in the anterior insula and to a lesser extent in the anterior cingulate cortex. Thus, as observing hand actions activates the observer’s *motor representation* [italics added] of that action, observing an emotion activates *the neural representation* [italics added] of that emotion. This finding provides a unifying mechanism for understanding the behaviors of others. (p. 655)

More generally:

In recent years, abundant evidence from behavioral and cognitive studies and functional-imaging experiments has indicated that individuals come to understand the emotional and affective states expressed by others with the help of the neural architecture that *produces such states in themselves*. (Decety and Jackson, 2006, p. 54)

Thus there is a general interpretation of the representations spoken of in the mirroring literature as re-realization of salient features. That makes them Aristotelian representations.

The anger one picks up from a companion or the pain represented in one’s brain are not, respectively, about the companion or about the pain. They are anger

and pain. They are also not, strictly speaking, true or false; one's beliefs about the companion's anger might be false, but one's felt anger itself is just there. The lack of aboutness and satisfaction conditions could seem to amount to a major disadvantage for any theory seeking to understand mental states in terms of Aristotelian representations. For example, it certainly seems a requirement on a theory of vision that one can see individuals and so one can have states that are in some sense about the individuals. Similarly, one can see what obtains, and so one has states that are truth-evaluable.

If we think our picture of the mind should consist in delineating computations in a way that leaves it an open question what they are realized in, the loss of aboutness and truth may seem particularly unacceptable, since we will lack an account of what the computations are computing. However, computational neuroscience carries with it an account of what the computations are computing without appealing to Fodorian representations; that is, the computations are computationally described neural processes, some elements of which are transformations of environmental features.

Further, if we take seriously the picture of human beings as creatures embedded in particular niches, we can ask whether the brain does need aboutness and truth. For example, Stich (1990) has made a strong case for saying that the brain needs useful states, which are not necessarily true ones. And Montague (2006) sees the developed brain in part as a response to demands for efficiency, with the consequence that precision starts to look prohibitively expensive. In many cases of action the brain does not need 'aboutness' in the sense of determinate reference, which means the supposed importance of truth can be seen to be lessened. This is so because the brain is tethered to its environment in a myriad of individuating ways; the idea that we should have to tag a neural transform of chocolate with a reference to an individual in order to explain why one reaches for one piece rather than another mistakenly intellectualizes many of the ways one moves through one's environment.

To say that the brain does not need truth is not to say that humans do not need or have truth-assessable states; however, they have to supervene on more than internal Aristotelian representations, which are not truth-assessable. In particular, vehicle internalism is a thesis about Fodorian representations; if we reinterpret 'representation' in an area of cognitive science as referring to Aristotelian representations, then we open the door to an externalist account of the vehicles of mental content.

There are projects in the history of philosophy that have undertaken to explain how cognitive states can have features not possessed by internal states. For example, for Hume statements ascribing abstract ideas are not made true solely by what is in the mind, but they are made true by that plus one's use of language. Hume's explanatory project generalizes: Aristotelian representations do not have anything like semantic content, so the ascriptions of states with reference or truth values are not made true solely by such inner states; rather, they are made true by that plus the environment, including the presence of linguistic communities. This is a heavily externalist project.

Finally, we should note that the historical philosophers employing Aristotelian representations often explicitly insist it is not possible for an Aristotelian idea to be a

false representation of some feature. 'False representation' is an oxymoron if 'representation' means 'Aristotelian representation'. Thus, Locke says:

Any idea then which we have in our minds, whether conformable or not to the existence of things, or to any idea in the minds of other men, cannot properly for this alone be called false. For these representations, if they have nothing in them but what is really existing in things without, cannot be thought false, being exact representations of some thing: Nor yet, if they have any thing in them differing from the reality of things, can they properly be said to be false representations, or ideas of things they do not represent. (ECHU, Pt. 2, Ch. 32.20, p. 433)

And Hume (2000):

Our ideas are copied from our impressions, and represent them in all their parts. When you would any way vary the idea of a particular object, you can only encrease or diminish its force and vivacity. *If you make any other change on it, it represents a different object or impression* [italics added]. (T 1.3.7.5; SBN 96)

Rather, the philosophers maintained, the seemingly false idea represents something else. It is true that one's experience might instantiate neural versions of properties that are not there in one's environment and that that might make them misleading because they may lead to false beliefs. Nothing in the account presented here precludes our saying that the misleading experience results in false beliefs.

Why take on the debt of such a project? There are at least three reasons for doing so. First of all, as we have seen, a significant amount of the discourse of cognitive neuroscience is about Aristotelian representations. While some philosophers are not fond of letting scientists decide on matters of interpretation (Appiah, 2007), there is a world of understanding to be gained by trying to see the picture of the mind actually being presented in cognitive neuroscience, rather than reinterpreting the findings in ways we find easier to understand.

Secondly, and relatedly, cognitive neuroscience is addressing a new question; namely, how does the brain give rise to our psychology? (Montague, 2006; Montague and Quartz, 1999; Zeki, 1999) Neural events and processes are described in terms that include the location and patterns of neural firing, along with references to chemical and structural features. How to understand the mind in these terms is a very different task from understanding how a mind independently characterized in semantic terms can get realized in the patterns of neural activity; that is, from understanding mental processes as computations over Fodorian representations.

The new task brings with it a methodological problem that Aristotelian representations can help solve. To see a creature as having a mind seems to involve seeing it as more than instantiating a causal chain set off by features in the environment; in some sense, a minded creature reacts to its environment. Contemporary philosophy of mind has one major answer to what mind-states relating a creature to its environment are like; they are Fodorian representations. The alternatives appear to be either behaviorism or eliminativism, and neither aims to explain how the brain gives rise to the mind. But cognitive characterizations involving Fodorian representations bring in the older linguistic model of the brain

developed before cognitive neuroscience hit its stride. The task of looking at how the brain gives rise to the mind's features becomes instead that of incorporating a pre-existing account of mental states.

Aristotelian representations are paradigmatic mental states in a different, non-linguistic model of the mind. As features realized in the brain, they fit immediately into the neural story without any puzzling residue about a content having causal powers. Relatedly, an account invoking Aristotelian representations without Fodorian representations comes with far fewer theoretical commitments, since Fodorian representations leave theorists with the fairly heavy task of figuring out how some neural episode can have content and truth-conditions, and the attendant difficulty that a theory about the assignment of content may place such an assignment far outside of neuroscience's domain. For example, on the widely well-regarded historical-teleological accounts of representations with content and satisfaction conditions, content is due to past facts about the evolution of the involved traits and is not a matter of current neural functioning (Neander, 2008).

Aristotelian representations are, then, part of a different model of the mind that is actually employed in cognitive neuroscience. As realized in the brain, they are mind-states at least as understood for thousands of years of philosophy of the mind. And their theoretical commitments are significantly less than those of the alternative Fodorian representations.

The philosopher's picture of Fodorian representations as causal elements in the brain exerts strong pressure on us to adopt a vehicle-internalism. Whatever external debts a mental state's content may bring, the state itself looks to be realized in the brain. Vehicle externalism is well served by Aristotelian representations because, while they can be wholly internal causes, they are not content-bearers; hence, the possibility is open that the content-bearers for our mental states are not brain states or processes. In what follows, we will see the other component of this paper's argument for vehicle externalism, and it has been important to attempt to mitigate the pressure towards internalism Fodorian representations carry. In addition, as we will see after we have examined the case for normativity, Aristotelian representations constitute an important part of the framework that the two components together provide.

### **3. The Normativity of Vision Science**

This section will argue for the normativity of vision science and connect that claim with an externalist view of vision's supervenience base. It is important to note that the argument establishing the externalist supervenience thesis does not establish the stronger enactive theory that Alva Noë and others have argued for (Hurley, 2001; Noë, 2004). After we have considered the thesis, we will consider some of the material needed to take us toward the stronger theory, but our focus will be on supervenience. This section concludes with a consideration of Ned Block's arguments (Block, 2005) against Alva Noë's theory and, in particular, against Noë's commitment to a supervenience thesis endorsed here.

There are three strong arguments for the view that the topic of vision science is well functioning vision. The first argument appeals to what vision scientists say are the sorts of general questions they are addressing. For example, Palmer's book on vision science (Palmer, 1999), which Block takes to be an instance of "the standard view," lists among the central questions:

How can we *determine* whether an object is large and distant or small and close?  
How do we *perceive* which regions in a visual image are parts of the same object?  
How do we *know* what the objects that we see are for? (p. 4)

That is, vision science aims to explain such successes.

A second argument appeals to a condition of success on vision science that is implicit in Palmer's goals. If one has a theory that does not distinguish between well and poorly functioning vision, one does not have an adequate general theory of vision at all. Nor is the distinction between well and ill functioning vision a statistical distinction; it may well be that the majority of people in the world have defective or impaired vision; that does not mean that a theory of vision is going to count such vision as well-functioning on the grounds that it is statistically normal.

The third argument appeals to an analogy. Artifacts may exhibit some of the features of living organisms, and so clocks form a telling analogy for us to start exploring the differences between conceiving of vision neuroscience as concerned simply with internal functioning and conceiving of it as concerned with explaining vision. However, the analogy is also limited in ways that will be important to note.

As Wittgenstein (Wittgenstein et al. 1975, p. 69) perhaps implies, you can have two identical clock faces at some moment, only one of which tells the time, since only one of the clocks is working. Like organisms, clocks do things and we are interested in their doing them successfully (compare Smith, 1988). Suppose, then, that we have two clocks:

**Clock 1:** plugged into a functioning outlet, time set in synch with the TV cable monitor, the speed of the dial is correctly calibrated, etc. Let's suppose that it is now noon and the clocks hands both point to 12.

**Clock 2:** Its hands also both point to 12. However, it hasn't been plugged in for years. Suddenly, right at noon, in one of those astonishing philosophers' "low probability chance fluctuations" [cf., Block, 2005], all the atoms match those of Clock 1 and 2 reduplicates the internal state of 1 for 2 seconds.

Was it telling the correct time for those two seconds? The answer "no" is justified on the grounds that a clock's telling the right time depends on a non-chance connection between what it says and what is true.

Telling the correct time does not supervene just on the internal states of the clock, for two distinct reasons:

1. Clocks have wide content; moving a clock can change whether it tells the correct time.
2. Functioning well *is constituted* in part by a non-accidental connection between changes in a clock's face and local time.

The first gives us a familiar form of externalism, content externalism. The second, however, gives us the less usual form of externalism, vehicle externalism, because it maintains that telling the time is in part constituted by states outside the clock face.

Should a science of clocks be focused on how clocks work correctly (science<sup>1</sup>) or merely on what the internal structure and operations of some clocks are (science<sup>2</sup>)? I shall take it as obvious that we do not have an adequate science of clocks if the science cannot tell us how well-functioning clocks manage to perform the function of telling the time correctly.

The analogy with clocks decisively puts in question the internalist's claims about supervenience. A world in which clocks are not calibrated correctly is one in which they do not tell time. Correspondingly, well-functioning vision, the proper object of vision science, does not supervene on interior states alone. A world in which one has the same internal processes in the brain with largely false visual reports is not a world in which one has well functioning vision.

At the same time, the clock analogy radically fails to provide an adequate picture of the calibration that vision requires. Without a better idea of what is involved in visual calibration, we will fail to have a robust conception of vision's supervenience base. One unfortunate consequence would be that vision continues to seem to tick along with only the occasional mind-world interactions.

A major manifestation of visual calibration is object constancy. For example, our movements do not produce the illusion that objects around us are moving and changing, even though the retinal inputs are changed by our movement. And this means that seeing the world as we see it is a function not just of our retinal inputs but also of our movements. It is essential, for example, that well functioning vision gives us a reliable ability to distinguish between cases in which we turn our head to the left and cases in which an object we are concentrating on turns to the right. We can set up a situation in which the retinal inputs are very similar in two such cases, but the cases will not even begin to look similar to the perceiver. In such extremely common cases of seeing an object as stationary while we move, as opposed to seeing one move while we are stationary, correct calibration requires the coordination of the one's movement and retinal input.

Possession of object constancy is part of the possession and exercise of sensorimotor knowledge and skills; such possession and exercise are causally important to vision, but they are also *essential* to correctly functioning vision. It is this essentialist thesis, made possible by the observation that vision science is the science of well-functioning vision, which goes beyond the causal claims and gives us a claim about what constitutes vision.

We can start to get an even fuller account of the calibration demands on vision if we consider the functioning of (part of) dopamine reward systems (Montague and Quartz, 1999). When we receive a reward, we get a burst of dopamine which over time transfers from the reward to the earlier conjuncts of the reward, and affects our sensory experience of them. Thus, deprived smokers may find cigarette packages leap out at them, while others of us never fail to spot a book store. Our experience can leave us attuned to spotting desirable things. If the formerly rewarding experience

becomes less rewarding, then if we are functioning well, the experience of the earlier conjuncts also changes. We may not be grabbed by the same things now as we were as ten year olds. Hence, we interact with our environment in such a way that it changes our experience of it as we explore it, and our so doing is very significant for our survival. This interaction is wholly lacking in the clock case.

The clock analogy can also leave the impression that identifying parts of the brain is rather like identifying parts of a watch, as though with a little shaking the structurally distinct pieces just fall out. However, the taxonomizing in vision science is in fact often not structural. Thus, for example, the division of the occipital cortex into different areas is done not solely on a structural basis, but rather includes the ways in which the stages and aspects of grasping one's world visually are located physically.

Though the normativity-driven taxonomizing that goes on in biology has been there at least from Aristotelian times, cognitive science's use of the computer model may have led it initially in a different direction. But whatever the reason, the old views of cognitive science contrast with the newer views of cognitive neuroscience (Montague & Quartz, 1999):

Early investigators thought that the really important problem was to find the functions or computations being implemented by the brain independent of the specifics of their implementation using biological components. This view is now seen as impoverished because as structures constructed by evolution, *most creatures are tightly woven into particular environmental and social niches, and are the 'answers' to manifold questions posed by their environs* [italics added]. (p. 87)

Ned Block's starting point is quite different. It is what he calls "the standard view," which he characterizes with two theses:

1. The minimal constitutive supervenience base for perceptual experience is the brain and does not include the rest of the body.
2. Although motor outputs and motor output instructions affect perceptual experience, much of perceptual experience can be understood in abstraction from such causes. (2005, p. 270)

His arguments for these theses are twofold: (1) The correct claim that causation alone does not confer metaphysical necessity and the assertion that once one has the causation, claims about constitution are explanatorily vacuous; and (2) An appeal to intuitions about a thought experiment. The rejection of constitution has been disputed above. The claims about constitution are in place, since they perform the additional task of situating the causes as essential to well functioning vision.

The appeal to intuition about a thought experiment brings in a form of the argument from hallucination. In Block's version, it involves the contrasting kittens (Held and Hein, 1963), one of whom could actively explore its environment, and the other of whom remained passive and had severe visual deficits as a result. The accuracy of the conclusions originally drawn by Held and Hein is not relevant for our purpose of adjudicating between Block and Noë, since the latter agree that the passive kitten lacks sensorimotor knowledge or skills. Block denies that this shows that

having such skills is essential to seeing on the grounds that Noë is thereby wrongly committed to *denying* that if the two kittens are in the same brain states they will have the same experience. As he says:

But the issue is not whether neural stimulation can *produce every* experience but whether if the relevant brain state were to come about—*somehow*—the experience would be instantiated. . . . To illustrate the point, one might ask the question: suppose it was arranged so that the passive kitten in Held's and Hein's experiment had the same brain goings on as the active kitten. The upshot of Noë's view is that the passive kitten still will not have the same experience as the active kitten. . . . Any state of the brain that can be brought about by normal perception could perhaps occur—although with very low probability—by chance fluctuation. (Block, 2005, p. 266)

The point of this thought experiment is to show that the “brain's goings on” actually achieved through coordinating movement and vision might have come from a very different source; in such a case, Block claims, the resulting experience would be the same. Hence, the sensorimotor causing cannot be essential to having the experience.

It is hard to see how Block can be right about the sensorily deprived kitten if “same experience” means “same visual experience.” Seeing, along with remembering and registering the correct time, are success notions and they are notions that require that the success is not purely accidental, a matter of “chance fluctuation.” Thus, however close someone today comes to being in the same brain state Queen Victoria was in after Prince Albert's death, it is impossible for one of us to remember that death.<sup>1</sup> Similarly, a chance fluctuation might give the kitten the same brain state, but that does not mean the kitten actually sees or has the same visual experience. Hence, Block must mean “the same *inner* experience,” whatever that might mean or be.

Presumably the idea is that there is an illuminating analysis of vision that makes reference to this internal state, as the argument from hallucination is supposed to show. Block has told us that “much of perceptual experience can be understood,” (p. 270) by ignoring motor input, and this idea of an inner experience which would be shared among creatures with the same brain goings on would seem to play a role in that understanding.

There are two strong points against taking an internalist construal of “the same experience” to provide a key notion of experience used in vision science. First, vision neuroscience lacks a criterion for the reoccurrence of an experiential state, either in oneself or in another, which is independent of having the same overall neural state (Palmer, 1999). Hence, statements about the same inner experience count as operationalized only as statements about neural duplicates. But in practice the criterion of having the same overall neural state cannot be applied. There is no way to tell whether two people have the same overall neural configuration. Philosophers have considerable interest in what obtains when two creatures in different environments or with different histories get into the same neural configuration. The considerable prevalence of ‘arguments from hallucination’ and discussions of wide and narrow content attests to it. But in so far as vision science is elucidating what the nature of vision is, the idea of qualitatively identical experiences in different

circumstances is not useful. Science has to deal with the real world and the practical limitations it imposes on studying very complex creatures. Taking philosophical thought experiments to such a context imports concerns that cannot be shared.

In addition, the idea that vision science would be concerned with the purely internal neural causes of such an experience ignores important aspects of its methodology and its distinctions. Psycho-physical experiments and the distinction between bottom-up and top-down processing both fail to have a role, given the idea that the transformed kitten has what vision science is interested in. One implication of this incongruity between philosophy and vision science may well be that some philosophers have a defective understanding of what vision neuroscience is doing today.

While these arguments show Block's view departs from that current in vision science, we should add that his and others' notion of "the same experience" should receive more philosophical scrutiny than it typically does in discussions implicitly or explicitly employing the argument from hallucination. That is, the claim typical of the argument, that for each case of seeing we can understand the possibility that we actually have "a hallucination that one would not be able to distinguish from such a genuine perception of one's environment" (Wedgwood, 2006, p. 314) may fit how we ordinarily remember and think about visual experience much better than it does the experience itself. As Palmer (1999) points out, our perception of our environment is in fact constituted by a succession of highly selective takes on a scene.

Complex visual scenes like the ones that we normally look at contain a staggering amount of information, far more than we can be aware of at one time. As a result, we have to sample visual information over time in a series of distinct perceptual acts, each one of which is inherently selective. (p. 532)

(Cf., Noë, 2004.) One's eyes saccade as one sees a scene and one's perception of it is not correctly identified with any one of these very partial takes. The philosopher's inner visual experience may in fact be a somewhat fictive construction out of these successive takes, one that memory may provide for us as a single record with much more detail than any one instance could have.

#### **4. Conclusion**

Aristotelian representations are key components of states in a system whose well-functioning is the topic of vision science. Since they do not have content, they fall short of providing the truth-makers for statements ascribing to us perceptual states with truth-assessable content. However, as we have seen at least since Hume's time, there are considerable resources to be drawn on in content attributions. It is just that they are not, or not solely, in the brain.

In contrast, many philosophical theories of perception do largely appear to link understanding perception with characterizing the truth-makers for statements ascribing a perception to someone. In addition, as the argument from hallucination indicates, they may pick up assumptions shaped by centuries of philosophical work

on problems of epistemic justification. Each of these background elements is very far from the project of providing a theory of well-functioning perception.

Should philosophy be engaged at all with the project of providing a theory of well-functioning perception? From the point of view of those who are so engaged, disengagement gives one an impoverished picture of perception. In addition, it makes it less likely that philosophy will be able to assume a critical interpretive stance toward the profound changes in our ways of conceiving human beings that cognitive neuroscience is creating.

## Note

- [1] This point is made with reference to John Bickle's reductionism in Jacobson (2005).

## References

- Appiah, K. A. (2007, December 9). Idea lab—The new new philosophy. *The New York Times*.
- Aristotle. (2004). *On the Soul*. (J. A. Smith, Trans.). Whitefish, MT: Kessinger Publishing.
- Aristotle, & Ross, W. D. (1956). *De anima*. Oxonii: E Typographeo Clarendoniano.
- Aquinas, St. Thomas. (1947). *Summa Theologica*. (Fathers of the English Dominican Province, Trans.). New York: Benziger Brothers, Inc.
- Block, N. (2005). Action in perception. *Journal of Philosophy*, 102(5), 259–272.
- Crane, T. (2003). *The mechanical mind: a philosophical introduction to minds, machines, and mental representation*, (2nd ed.). London, New York: Routledge.
- Decety, J., & Jackson, P. L. (2006). A social-neuroscience perspective on empathy. *Current Directions in Psychological Science*, 15(2), 54.
- Elder, C. L. (1998). What versus how in naturally selected representations. *Mind: A Quarterly Review of Philosophy*, 107(426), 349–363.
- Fodor, J. (1975). *The language of thought*. New York: Crowell.
- Fodor, J. (1990). *A theory of content and other essays*. Cambridge, Mass: MIT Press.
- Fodor, J. (2006). *The revenge of the given*. Unpublished manuscript.
- Held, R., & Hein, A. (1963). Movement-produced stimulation in the development of visually guided behavior. *Journal of Comparative and Physiological Psychology*, 56(5), 872–876.
- Hofbauer, R. K., Rainville, P., Duncan, G. H., & Bushnell, M. C. (2001). Cortical representation of the sensory dimension of pain. *Journal of Neurophysiology*, 86(1), 402–11.
- Hume, D., Norton, D. F., & Norton, M. J. (2000). *A treatise of human nature, Oxford philosophical texts*. Oxford; New York: Oxford University Press.
- Hurley, S. (2001). Perception and action: Alternative views. *Synthese*, 129(1), 3–40.
- Jacobson, A. J. (2003). Mental representations: What philosophy leaves out and neuroscience puts in. *Philosophical Psychology*, 16(2), 189–203.
- Jacobson, A. J. (2005). Is the brain a memory box? *Phenomenology and the cognitive sciences*, 4(3), 271–278.
- Jacobson, A. J. (2006). Doing it his way: Hume's theory of ideas and cognitive science. In A. Brook (Ed.), *The pre-history of cognitive science*. Basingstoke: MacMillan/Palgrave.
- Jacobson, A. J. (2007). Empathy, primitive reactions and the modularity of emotion. *Canadian Journal of Philosophy*, Suppl. 32, 95–112.
- Locke, J. & Fraser, A. C. (1894). *An essay concerning human understanding*. Oxford: Clarendon Press.

- Montague, P. R., & Quartz, S. R. (1999). Computational approaches to neural reward and development. *Mental Retardation And Developmental Disabilities Research Reviews*, 5, 86–99.
- Montague, R. (2006). *Why choose this book? how we make decisions*. New York: Penguin Group.
- Neander, K. (2008). Teleological theories of mental content. In E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Retrieved from <http://plato.stanford.edu/archives/fall2008/entries/content-teleological/>.
- Noë, A. (2004). *Action in perception, representation and mind*. Cambridge, Mass: MIT Press.
- Palmer, S. E. (1999). *Vision science: photons to phenomenology*. Cambridge, Mass: MIT Press.
- Peacocke, C. (1992). Anchoring conceptual content: Scenarios and perception. In J. Ezquerro (Ed.), *Cognition, Semantics and Philosophy* (pp. 293–319). Norwell: Kluwer.
- Ramachandran, V. S. (2004). *A brief tour of human consciousness: from imposter poodles to purple numbers*. New York: Pi Press.
- Rowlands, M. (2006). *Body language: representation in action*. Cambridge, Mass: MIT Press.
- Smith, B. C. (1988). The semantics of clocks. In J. H. Fetzer (Ed.), *Aspects of artificial intelligence* (pp. 3–19). Dordrecht, Boston: Kluwer Academic Publishers.
- Stich, S. P. (1990). *The fragmentation of reason: preface to a pragmatic theory of cognitive evaluation*. Cambridge, Mass: MIT Press.
- Ward, J. (2006). *The student's guide to cognitive neuroscience*. New York, NY: Psychology Press.
- Wedgwood, R. (2006). The internal and external components of cognition. In R. J. Stainton (Ed.), *Contemporary debates in cognitive science* (pp. 307–325). Blackwell Publishing Company.
- Wittgenstein, L. Rhees, R., Hargreaves, R., & White, R. (1975). *Philosophical remarks*. Oxford: Blackwell.
- Zeki, S. (1999). *Inner vision: an exploration of art and the brain*. Oxford, New York: Oxford University Press.

Copyright of *Philosophical Psychology* is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.