

Action in Perception

A Statement

Presentation Notes for the DC Williams Colloquium on the Intentional Content of Vision and Grammatical Competence at the University of Maryland, College Park, April 2006

Alva Noë

UC Berkeley

Part One: *The trouble with visual representation*

I begin with a question that comes into focus when we take seriously the central idea of the version of the two visual-systems hypothesis that has been championed by Goodale and Milner. The idea is that there are two distinct visual capacities, *conscious seeing*, on the one hand, and *vision for the guidance of action*, on the other, each of which is subserved by anatomically distinct neural structures (the ventral and dorsal streams, respectively). The question I want to pose is this: how do these distinct visual capacities interact? How are they able jointly to operate in normal vision? That they *do* jointly operate is beyond doubt.

Consider what happens when you drop your purse and coins scatter across the floor. You look around and spot them, one by one. You find them with your eyes and pick them up. Vision guides you to them. You couldn't pick them up in anything like the way you do if did not see them. Or consider the famous visual form agnostic DF, who figures prominently in Goodale and Milner's discussions. DF has residual visual consciousness – she sees colors, and they show up for her, I gather as belonging to objects, as opposed to being free floating or unattached. DF is able to reach for and succeed in grasping objects of various shapes and sizes – a fork, say, or a screwdriver -- correctly molding her fingers and hand to reflect the visible size and shape of the object. What she cannot do, or rather, what she is unlikely to do is grasp these items in ways that are appropriate to their use or function, for she cannot identify what she sees, and so she is not visually sensitive to the function or meaning of what she sees. What DF is unable to do on the basis of what she sees helps us mark out the role performed by normal visual experience in guiding action. (As an aside: the fact that DF is not visually sensitive to what things are for – to what Gibson called their affordances – should put to rest the suggestion that the dorsal stream subserves the perception of affordances, a claim made, for example, by Jacob and Jeannerod (2003, 181).)

We can appreciate more fully the point that visual guidance of action – even visually guided reaching and grasping -- relies on visual consciousness by recalling that in daily life, optic ataxia seems to be a fairly minor impairment. 1) Most optic ataxics do not complain of action deficits in daily life. In Vighetto and Perenin's 1981 group study, half the patients complained of mild sensory difficulties and half seem not to have noticed any impairment at all, prior to their participation in systematic experimental

investigation. 2) It is not generally the case that optic ataxia is a deficit in visual guidance of action. “Most patients with optic ataxia are able to guide accurate actions toward objects in central vision” (Rosetti, Pisella and Vighetto 2003, 173). “Out of ten patients studied by Vighetto (1980; Perenin and Vighetto 1983, 1988), only three had a deficit in central vision” (Rosetti et al 2003). Even patients with bilateral optic ataxia “exhibit normal visuomotor performance in central vision in many experimental conditions involving” a grasping response. Problems show up when reaching into the periphery of the visual field or 3) when there is a requirement for fast, online modifications of reaching in response to the displacement of a target object. 4) Performance in reaching and grasping is much improved for familiar objects *even when they are in the periphery of the visual field* (Jeannerod et al 1994; Milner et al 2001). And of course most of the objects that we reach out and grasp in daily life are likely to be familiar.

Now these considerations have several upshots. First, we can rule out that the dorsal stream is the main support of action. The ventral stream supports action in a broad range of cases. Neither the phenomenon of visual form agnosia, nor that of optic ataxia, give us reasons to doubt visual experience plays a basic role in the guidance of action; in particular, there is no challenge to the idea that what we see causally influences our reaching and grasping, even if it is also true that we don’t need conscious visual attention to help us decide how to move our fingers when reaching for something. Second, we can see that it is misleading to suggest that there is a double dissociation of seeing and visually guided action. What is dissociable, rather, is visual consciousness, on the one hand, subserved by the ventral stream, and something like vision for the setting of some (size and shape based) parameters for reaching and grasping, as well as for making rapid, quick responses to objects and events around one. *This* dissociation is pronounced. Of course it isn’t very surprising either. Every typist knows that paying close visual attention to your fingers will disrupt rather than improve performance; and there is no need to look at the ground as you take a step.

These conclusions help us take a first step toward answering the question with which we are concerned: how does visual perception in the ventral stream – which is usually understood to represent or describe objects, their features, interrelations and location – manage to influence the dorsal stream in setting the parameters for reaching and grasping? To a first approximation, the answer would seem to be that the ventral stream enables us to *see* objects of interest, and so to select them as targets for reaching and grasping; the information about the selected objects represented in an experience is then passed on to the dorsal stream. Visual experience gives us the world as present; with the world made present in experience, we can act on it. On this way of thinking, the dorsal stream is analagous to the autofocus on a camera. The autofocus is not controlled by the photographer; it responds automatically to the situation. But you need a photographer to find objects worth taking pictures of and to aim the camera in the right direction. For ease of reference, I am going to refer to this simple picture of how seeing and dorsal visuomotor guidance are related as the AutoFocus Conception. The AutoFocus conception has two parts: 1) the ventral stream produces representations of objects present in the environment; ; 2) the dorsal stream is then able automatically to target and callibrate itself appropriately to the represented objects.

The AutoFocus conception raises three issues.

The first I mention only to set it aside. It is the representation problem. We don't really understand how the brain produces representations of objects, of their features and interrelations, of their locations, in the ventral stream or anywhere else. So it just isn't that clear how the story is supposed to go.

The second is what I call the commensurability problem. It is widely appreciated and has been stated nicely by Bridgeman (in press). "It is not clear how the cognitive system [the ventral stream] communicates a particular object as a target, to a system [the dorsal stream] that cannot distinguish different targets." The problem is this: The ventral stream builds up a representation of the object in object-centered coordinates, but the dorsal stream is a kind of robotic zombie that is built to detect and respond to the true size and shape of things; whatever the dorsal robot represents, it represents it in egocentric, not object-centered, terms. So we can ask, how is it that the dorsal system ever manages to pick up the very thing the ventral system is looking at? Suppose, to use Jeannerod's example (Jeannerod 1997, 80), that we want to eat the lone red apple among the many green ones. We see the red one; we want it; and we manage to reach precisely for it. How do we carry this off?

A suggestion made by Jeannerod (1997, 80) and endorsed by the philosopher John Campbell (2002), and I think by Goodale and Milner (2004) too, is that what the ventral stream needs to do is to tell the dorsal stream *the location of* the object on which it is to act. The ventral stream passes along the object's coordinates to the visuomotor system. The dorsal stream doesn't need to be sensitive to the object itself, merely to its location.

This proposal raises two issues, one merely technical and one more substantive. The technical question issue is this: How does the dorsal stream get the real-world, egocentric location information from the allocentric visual representation of the object made available by the visual consciousness module? Goodale and Milner address this issue when they point out that the dorsal and ventral stream share a common retinotopic base camp in primary visual cortex; they suggest that "Once a target has been highlighted on a retinal map [thanks to back-projections from the ventral stream], it can be converted into any other coordinate system that the dorsal stream might need to use" (102). This seems right, as far as it goes. But this proposal to the technical problem doesn't address the more substantive issue. Remember: the dorsal stream wants to know how and in what manner to move the hands to reach the target; the dorsal stream is not interested in the relative size and position of the thing. The problem is that, at least on what would seem to be the standard conception of visual experience, that's all that the ventral stream provides. Jacob and Jeannerod (2003, 197) make this conception of experience explicit when they explain:

Consider a visual representation of a mug to the left of a telephone. In the visual percept, the location of the mug relative to the location of the telephone is coded in allocentric coordinates. The percept has an iconic or pictorial content that is both informationally richer and more fine grained than any conceptual representation of that state of affairs. For example, it cannot depict the mug as being to the left of the telephone without depicting how far the former is from the latter. Nor can it depict the relative distance of the mug from the telephone without depicting the

orientation, shape, texture, size and content (if any) of the mug, the orientation, shape, size and color of the telephone.

My concern here with this passage is not with the contrast drawn here between pictorial, analog content and conceptual nonanalog content. What is of interest, rather, is the way the pictorial content is specified, i.e. the idea that seeing provides a representation of *a* cup to the left of *a* telephone. It just isn't clear to me how this sort of pictorial representation of a scene can be used to specify *where* the thing is, any more than looking at something through binoculars can help you see how far away from you things are.

It may be, then, that what we need is an alternative to the standard conception of visual experience. This conclusion is reinforced by a consideration of the third problem I want to raise about the AutoFocus Conception.

This third issue I want to mention is a problem about perceptual presence. We have considered that a natural way to understand the role of visual experience in guiding action is that, intuitively, it provides us with a sense of the presence of objects and so enables us to reach out and grasp them. The problem that arises now, and for the second time, is that it doesn't seem that visual experience, as it is understood on standard conceptions, can discharge this role. The standard conception, as we have seen, is that the content of an experience has something like this form: an object satisfying such and such a description stands in thus and such relations to other things at a certain place in the environment; or, to use Jacob and Jeannerod's example, an experience has a content like this: a mug is to the left of a telephone. The problem is that when you see a phone and a mug, or whatever, it isn't the case that you merely see *an x* or *a y* satisfying a certain description. You see these very particular things themselves, and you see them even if you don't attend to or notice their features. Moreover, when you reach out for the phone say, on the basis of what you see, what you reach for is *the very thing* that you see, namely, *the particular phone itself*. Experience seems to play the role of enabling us to lock onto things themselves; it does not merely provide us with general descriptions. And it is in thus giving us the things themselves as present that experience seems to play a role in making action possible. – Can the standard conception of visual experience accommodate this function of experience? The worry here is familiar to philosophers. It is hard to see how one can distil *reference* to a particular thing from a merely descriptive representation of it. But what we need from experience is reference to the particular thing.

A defender of the standard view of experience and the AutoFocus Conception is like to respond to this worry as follows: It is precisely what makes the research of Goodale and Milner so astonishing, and so important, that it requires us to accept the consequence that the dorsal stream succeeds in targetting its object *even though*, as it turns out, visual experience does not lock onto the object *in particular*. Dorsal-based adjustments of grip aperture, for example, are simply autonomous in respect of what we see. Seeing may not deliver reference to the object of sight. But that's just so much the worse for seeing; it's no skin of the dorsal stream's back, for the the dorsal stream can get the job done its own way.

In fact, so the defender of the standard view will continue, what the data shows us is not only that we don't not need to experience the presence of particular things in order to reach out and grasp them, but further, that the feeling of visual presence itself is nothing but an illusion. Indeed, perhaps it is an illusion generated by the fact that we can reach out and pick things up!

If this is right, then the structure of visually guided action in normal vision is identical to that of visually guided reaching and grasping in blindsight. In blindsight the patient does not *see* or consciously experience that for which he reaches. He reaches not on the basis of what he sees, but rather on the basis of what he is told to do, what the experimenter instructs him to do. He relies on the experimenter's description of what is there for him. The blindsight patient lacks a sense of the presence of what he sees, and so he is not guided by that presence. On the standard view of experience, it would seem, visual experience stands in just the same relation to our reaching, grasping system that the instructions of the experimenter stand in relation to the blindsight patients' visuomotor system – the ventral stream of the normal and the experimenter each offer the visuomotor system (or, for that matter, the reasoning thinking system) a more or less useful description of what is there. On this approach, what differentiates the seeing case, from that of blindsight, is that in the former, but not the latter, there is an accompanying (probably illusory) feeling of presence!

I find all this very unsatisfying. Here are four reasons to reject the AutoFocus conception and the standard view of experience.

1. If the defense of the standard view is right, then we ought to find, but we do not find, that there is a kind of visual analog of Capgras' Syndrome (Hirstein and Ramachandran 1997). In Capgras Syndrome, a patient might come to be convinced that his wife is an imposter. He grants that she looks *exactly* like his wife, but he is convinced it's not really her but a look-alike. According to Ramachandran, patients with Capgras are perceptually normal; for some reason they are missing the *feeling of familiarity* that usually accompanies our perceptual judgments. Dennett has speculated along similar lines about some kinds of obsessive/compulsive disorder. The obsessive/compulsive checks the door, establishes that it's locked, but finds that he must check again and again, not because he hasn't confirmed that the door is locked but because, for some reason, he isn't getting the familiar *feeling* of having confirmed it. Well, if the way of thinking about experience that we are considering is right, then we ought to find patients in whom, as a result of lesion, presence and representational content are in fact dissociated.

2. A second reason for rejecting this defense of the standard view is that it now carries the burden of making intelligible the claim that you can factor experience into a representational content and a dissociable feeling of presence. Mohan Matthen (2005) suggests that we are all familiar with cases of visual experience minus the normally associated feelings of perceptual presence. This is what happens when we *visually imagine* a couch that is not before us, say. The content of our experience – that there is a couch with such and such qualities in such and such relation to a background – may be exactly the same as that of the corresponding experience of seeing the couch. In the latter case, however, but not the former, we typically have the added feeling of the presence of the particular couch. Pictures provide Matthen's second example. When you see a picture you have a visual experience with the content that is depicted in the picture, but you don't

have a sense of the presence of that which is thus depicted. For example, when you see picture of the Eiffel Tower you have an experience that represents it thus and such, but you don't have a sense of its actual presence to you in the picture. What supplies representational content with presence, according to Matthen? The dorsal stream. Motion guiding vision, Matthen argues, is nondescriptive and deictic. It secures reference to particulars and makes it the case that that which is represented is made to feel vividly present. A consequence of Matthen's view is that damage to the dorsal stream would in fact alter visual consciousness by making it more like picture-consciousness or like the visual imagination. As I have already indicated, this upshot is untenable. Optic ataxics sometimes fail to know that they have any impairment whatsoever, let alone an impairment in their visual consciousness of this order of magnitude. It is part of the burden of this view to offer an account of the feeling of presence, one that treats presence as independent of the dorsal stream, on the one hand, and representational consciousness, on the other. Unless we are given such an account, it is hard to take the view seriously.

3. A further problem with the AutoFocus Conception and the standard view of experience, one related to this previous one, is this: it is just *not* the case that the neurological data supports the claim that visual guidance of action is autonomous of the ventral stream. We discussed this at the outset. Optic ataxia is not a condition of general impairment in visual guidance of action. Moreover, as Goodale and Milner explain, while grip aperture seems to be dorsal, grip force seems to be ventral. Given this, it seems that visual guidance of action needs information about the presence of objects that can only be supplied by visual consciousness.

4. Finally, other things being equal, I would prefer an account of vision, of our sense of the world's presence, and of our ability to act on the basis of what we see, that does not represent us as wildly confabulated and deluded.

Fortunately, we have another live option. Instead of thinking of the ventral stream as delivering picture-like visual experiences that are useless for purposes of telling us where things are or whether they are there, let us suppose that seeing is a way of being in touch with particular things. The content of a visual experience is not: a cup is to the left of a telephone; rather, it will be something more like, *this cup here is to the left of that telephone*. Seeing is a way of locking onto things; it is a way of being related to them.

If visual experience is in this way non-descriptive, it becomes immediately clear why blindsight *feels so different* from normal visually guided action. It isn't that in normal seeing we have an associated (illusory) feeling of presence that we misguidedly think makes a difference to how we act. It is that in normal seeing we act on the basis of the presence of particular objects or situations, whereas in blindsight we act only on the basis of general descriptions, descriptions that leaves our relation to what is described unspecified and indeterminate. What makes healthy intact vision different from blindsight is that vision results in contact with objects themselves, but blindsight is forced to make do with mere descriptions.

This way of thinking about seeing also helps us see how to solve the commensurability problem. When we act on what we see, we aren't in the position of needing to convert one kind of description of where a thing is into another. Our predicament is not like that at all. That framing of the problem presupposes the standard

view of experience that we are giving up. Think of seeing the object as like pointing to it with the eyes. The dorsal stream doesn't need to be told *where* to reach; it already knows that it needs to reach where you are looking, to reach for what you are looking at. Here we take a leaf from the page of the roboticist Dana Ballard. Metaphorically, it is as if visual consciousness takes us by the hand and leads us to the object.

As far as I know, there is nothing in the empirical data on visual form agnosia, blindsight or optic ataxia that speaks against this idea that seeing is a kind of visual reaching to the object with the eyes. The ventral stream is known to support general spatial orientation; optic ataxia is not, as a general rule, an impairment of pointing or the directing of the eyes. Of course, there are optic ataxics who are thus further impaired. We find this in Balint's syndrome. But as Goodale and Milner explain, there is good reason to think of Balint's as involving damage to both systems. Moreover, in Balint's we get dramatic distortions of visual consciousness (e.g. simultagnosia).

Indeed, there is some data to support this proposal that visual awareness of objects is basic and that it does not depend on the visual system's ability to represent locations in allocentric coordinates. This is the idea advanced by Zenon Pylyshyn (2003); it gains support from his work on multiple object tracking. What Pylyshyn shows, first, is that the category of the *object* is fundamental in perception. For example, we are quicker to recognize differences or similarities between features belonging to one and the same object than to distinct objects. Second, he argues that we can pick out and keep track of objects without binding features to locations (in the way that Treisman's feature integration theory requires). Part of what justifies this claim is the fact that we are able to keep track of objects in multiple tracking experiments even when (as is in fact the case) we are unable to keep track of the object's features (even when those features change without our noticing). Moreover, it seems that we can keep track of objects that spatially overlap, ie. that share locations!

The crucial thing about object-directed experience in Pylyshyn's sense is that it is an achievement of consciousness. This isn't something that Pylyshyn says; he doesn't like to talk about consciousness. He prefers to talk about attention. Let's call it attention then. The point is that we attend to, we keep track of naked individuals, and having kept track of them, we are able to experience their properties, and also to act on them. If I am right, the sensitivity to objects is the foundation of perceptual judgment and visuomotor skill.

For what it is worth, phenomenological considerations support this relational view of experience.

First, when you see the phone and the cup, you have a sense of their presence there; not as represented in your head. The world is present to us in experience; it is not represented by us in experience.

By the way: this is Gibson's central idea – that seeing is the achievement of contact with the world, rather than the representation of the world. (Consider Fodor and Pylyshyn's 1981 criticism of Gibson.)

Second, Phenomenological reflection demonstrates that perceiving is, for us, an encounter with situations and things; it is not, for us, an encounter with mental images or some other kind of interior data of sense. This claim about the world-involving character

of perceptual experience -- that experience is an encounter with things and situations -- is *not* compatible with any old metaphysical or empirical picture of perception and its nature. In particular, it is not compatible with any internalist and representationalist view like Searle's. For in presenting perceptual experience as a kind of *involvement* with or *entanglement* with situations and things, the phenomenology presents experience as something that could not occur in the absence of situations and things. Phenomenology reveals perceiving, then, to be a condition whose nature depends *essentially* on the presence and involvement of the world encountered. If there were no object, or no situation, then there could be no contact with or involvement with them, which is just to say that there could be no perceptual experience.

The perceptual phenomenology by itself, taken at face value, entails that a physical duplicate of me would not be in the same experiential state that I find myself in, if he were not, like me, confronted by the very same situation. The phenomenology is silent as to whether my duplicate would be having *some* experience – perhaps he would be having an experience that he cannot discriminate from the experience that he would have if he were in my situation (Martin 2004) – but what the phenomenology does commit us to is the proposition that the experience of my duplicate is not (indeed, could not) be an experience of the same *basic kind* as my experience. Searle's view, in contrast, is committed to just this consequence.

Philosophers sometimes suppose that phenomenology is *free standing* in the sense that phenomenological facts are logically and conceptually independent of empirical or metaphysical facts. The phenomenological facts can be settled, on such a conception, without settling any other questions about the natural world. Phenomenology, so conceived, is a matter of how things seem to us. How things seem to us leaves open how things are beyond the limits of our consciousness. This autonomy of phenomenology has a methodological upshot: it is possible to undertake a phenomenological investigation without making any empirical or metaphysical assumptions. We can bracket science and metaphysics when we do phenomenology. On one reading, this was Husserl's view. The purpose of the phenomenological reduction, or *epoché*, for Husserl, on this reading at least, is to neutralize the full cast of empirical and metaphysical commitments.

I reject pure phenomenology. If I am right, there is no way Searle and Gibson can agree on the phenomenology. Phenomenology makes metaphysical commitments.

Now, it might be objected, *how could you think that reflection on your experience could give you insight into basic facts about nature and our place in the world?* But this challenge betrays a misunderstanding. Phenomenology, or the investigation of experience itself, on the conception I am proposing, is only, if it is anything at all, a kind of investigation of the world. In particular, it is an investigation of the world in so far as the world – the things and situations in which we find ourselves – gives contour to human experience. Crucially, then, the basis for my phenomenological assertions is not, say, reflection on how things are with me subjectively, or how things are in my own case. My grounds for thinking that experience is an encounter with things and situations is an evaluation of the nature of *perception and its nature*, an evaluation subject to re-evaluation in light of anything else we might know or come to learn about perception and its character.

And of course, crucially, I may be wrong, even *wildly* wrong, about the character of experience. There is no particular reason to think that it should be easy reliably to discern the nature of perception and the nature of our perceptual relation to the world. The immediate point is that it is precisely getting right about how things are that forms the burden of phenomenology.

My point is weak, but it is important: phenomenology, understood correctly, isn't representational. And that's a reason to take seriously that it is not representational, a reason that Kelly should take seriously. Another reason to take it seriously, we have just seen, is that it gives us the means to resist arguments from hallucination. The fact of hallucination shows no more than that we may be hallucinating; it cannot compel us to think of experiences as representations in the head.

Part Two: *The problem of perceptual presence*

Now we come to a final problem. The view of experience that I am now defending would seem to come up against the problem of perceptual presence, but in a different but no less difficult way. Whereas the problem for the standard view was how do we get particularity out of descriptions of features, our problem now is how do we get full-blooded objecthood – objects with features – out of naked visual reference? How is it that visual objects can show up for us as propertied and as situated in complex visual scenes? I'll try to explain how.

Consider: solid opaque objects have visible and invisible parts. When you look at a tomato, you can't see its back side. When you look at a cat behind a picket fence, you can only see the parts of the cat that show through the fence. Nothing surprising here. What could be surprising about the fact that you can't see the occluded parts of the things you see? Euclid epitomized the thought: when an object is seen it is never perceived at once in its entirety.

There is a way of thinking about perception – a reasonable way – that comes into conflict with these Euclidean truisms. According to this reasonable way of thinking about perceiving, when you look at a tomato, or at a cat behind a tree, you have a sense of the presence, a perceptual sense, of the presence precisely of a *cat* or of a tomato, that is to say, of a voluminous, ovoid-with-a-furrow, red fruit.

It was in defense of this reasonable way of thinking about perceiving that Strawson argued that we distort or misdescribe our experience of the cat or the tomato if we describe them as the experience of cat parts or tomato parts, let alone of cat or tomato sense data. To characterize our familiar perceptual experience, the experience we have when we take our experience at face value, we must describe it precisely as the experience of the cat or of the tomato.

And so we have a standoff between the Euclidean on the one side and the reasonable philosopher on the other. Each side can appeal to phenomenology, that is, the insights of each side are grounded in everyday reflection on experience. And yet each side tends dogmatically to deny the insights of the other side. The Euclidean insists that we don't really see the tomato or the cat, that we go beyond what is given when we say

that we do. And the reasonable philosopher, no less dogmatically, denies that our experience is limited by perspective in the way the Euclidean insists.

Now, my topic today is a problem that I call the problem of perceptual presence. And this problem only comes clearly into focus when we recognize that both sides in this standoff are, in a sense, right. The fact is, you do experience the presence of the cat and the tomato even though you don't see them all. The problem we then face is this: in what could your sense of the presence in perception of the cat or the tomato consist if it does not consist in the fact that you see them all?

It is important not to confuse this problem with a nearby, closely related problem epistemological problem about perception. This nearby problem is the question of whether we are entitled, on the basis of what is given, to judge things to be the way they perceptually seem to be. Our problem is prior to this. Our problem is precisely with the question, *what is given?* As I said at the outset: our question is: what do we see? This is a problem about content. The fact is that it seems that we can always answer this question two ways, that there is a kind of duality in the content of our seeings. In a way the problem concerns what we can think of as presence in absence. The hidden parts of the cat are out of view, hidden from view, absent from view, and yet they are perceptually present. In what can the perceptual presence of that which is absent consist?

Now it is important to see that we don't make progress toward a solution to our puzzle if we say something like the following: we don't see the cat as a whole, or the tomato as a whole. Rather, we *infer or judge or think* their presence. After all, we know what cats are, what tomatoes are, we have those concepts. Maybe we somehow cognitively fill in our experience using this background conceptual knowledge?

Now, I think there is something surely right about this sort of proposal. I'm very sympathetic to the idea that our experience depends on our background of understanding. I'll return to this later. But this simple appeal to inference and conceptual mastery cannot be the whole story, and for two reasons.

First, remember, our problem is precisely that of understanding in what our perceptual sense of the presence of the hidden parts of the tomato or the cat consist given that we can't see them, that they are hidden from view. It is of no help to be told that we don't see them. Of course we don't see them. That's just the point.

Second, there is a difference, surely, between *thinking* that something is present, and having a *perceptual* sense of its presence. This is phenomenological bedrock. At least it is for now. I think there is a hallway on the far side of this wall, but it doesn't now *look* to me as if there is. It feels present, but it is not perceptually present. In contrast, the back of the tomato, or the cat, *look* present. They are perceptually present. Indeed, something can *look* present even when one knows that it is not present. So, for example: Here we know perfectly well that the card B is not continued behind card A, but it looks as though it is. Psychologists sometimes call this amodal completion and they do so to distinguish this sort of phenomena from, say, phenomena such as illusory contours where you seem to see something that is not present. Here you seem precisely *not* to see something that is not present. But the crucial point is that this not seeing is *modal*; it is visual. B looks to extend behind A.

The upshot of this is then that even if in fact we do make use of our knowledge of tomatoes or cats or whatever when we experience things as being this way or that, this appeal to background conceptual understanding can't be the whole story. We still need a further explanation of what the distinctively *visual* sense of the presence of the unseen features consists. This is the challenge we face.

Now, before we go on to consider a solution to the problem, I want to call attention to the fact that the problem is actually one of very great generality. A whole range of perceptual phenomena illustrate the problem of perceptual presence. Consider some examples:

Circular coin

Trees along the road side

Color constancy

The experience of detail

Each of these phenomena illustrates the phenomenon of presence-in-absence.

Now, I said at the beginning that the problem of perceptual presence only comes clearly into focus when we realize that in a way both the sense datum theorist and the Strawsonian reasonable philosopher are right. To appreciate this we need to recognize that, in a way, perceptual experience is much more complicated than we are inclined to suppose. Perceptual experience presents us with the world and it presents us with how the world perceptually seems to be from here. A satisfying account of perception needs to explain how the silver dollar can look both circular and elliptical, how the trees can look to be the same in size and different size, how the wall can appear uniform in color across its surface and also variable in color across its surface. The content of perceptual experience is two dimensional. This needs explaining. To account for what I am calling the problem of perceptual presence is to provide such an explanation.

Philosophers have tended to chicken out when it comes to the two-dimensional character of perceptual experience. They deny the legitimacy of one or the other dimensions of content. Sense-datum theorists deny that we visually experience the coin's circularity or the tomato's sameness of size. At best perhaps we infer the presence of those features on the basis of what is present to consciousness. A certain kind of direct realist, in contrast, is tempted to deny that there is any sense in which the coin looks elliptical or the tomatoes appear to differ in size.

Among the chickening out strategies, one stands out as particularly influential. We find it developed in the work of Peacocke and also in the more recent work of AD Smith. Smith, for example, insists that coins, when seen from most angles, do not look elliptical. This is simply not true, he says. They look round, or perhaps "round and tilted away from you" (172). Smith does not deny – just as Peacocke doesn't – that there is a definite experiential basis to the claim that the coin presents itself by means of an elliptical perspectival property. He acknowledges, reasonably enough, that when a person insists that the coin *looks* elliptical to us even though we *know* it is round, she isn't simply mistaken. She is calling attention to a salient aspect of our distinctively visual experience when looking at the coin.

However, Smith does want to deny that this phenomenal aspect corresponds to the way the coin looks. He stakes out a position according to which what is elliptical is not the way the coin looks, but a sensation, that is to say, a quality of our experience of the coin. In this way the view preserves the idea that there are two dimensions along which the phenomenal character of experience gets fixed, but one of these gets kicked “upstairs and inside,” so to speak. – What justifies this claim?

Smith emphasizes that one feels no inclination – not even “the tiniest bit” -- to take a coin *to be* elliptical when one sees it (182). He asserts that “After all, if something really does perceptually look elliptical to me, I shall, if I notice the thing, and if I have no countervailing information to hand, take the thing to *be* elliptical, for I have nothing else to go by.” This may be so. However, it leaves open that what explains the fact that normal adult perceivers feel no inclination to take the coin to be elliptical on the basis of its visual appearance is the fact that they do have an abundance of countervailing information. Most of us know that coins are round, and we know that round, flat things change their perspectival shape as our spatial relation to them varies.

As a matter of fact – contrary to what Smith’s view would lead us to expect -- there is considerable evidence that genuinely naïve perceivers – perceivers truly lacking the relevant countervailing information – *would* take round coins presented at an angle to be elliptical. Cheselden (discussed in von Senden) describes the astonishment of a young boy who had undergone cataract surgery at the way a coin changed its shape as it moved; Helmholtz describes a similar case of a boy astonished at the changing shape of a locket. More recently, in a similar vein, Valvo describes a patient who, shortly after undergoing surgery to remove cataracts, perceived as black holes what he later found out to be windows of a house across the street. Large objects far away looked to this post-operative patient like small holes nearby.

Smith gives an additional reason for thinking that coins don’t usually look elliptical (or, *mutatus mutandis*, that the trees don’t look different in size). The quality of our visual experience that leads theorists such as myself to say (mistakenly) that the coin looks elliptical is one that, obviously, changes as our spatial relation to the coin itself changes. But normal perceivers do not experience the shapes and sizes of what they see to be changing. This shows that our experience of ellipticalness is not an experience of the coin’s *being* elliptical.

Well, I certainly accept that mature perceivers can experience a coin as elliptical without coming to think that it is elliptical. I grant, after all, that there is genuine size and shape constancy. But this does not support the claim that the coin does not look elliptical. The question after all is not whether there is size and shape constancy – I stipulate that there is – but whether there is also *inconstancy* with regard to apparent size and apparent shape. Nothing Smith says rules that out.

Part Three: *A solution to the problem of perceptual presence*

When you approach an object, it looms in your visual field. When you move around it, its profile deforms. When you blink, sensory stimulation from the object is

disrupted. In these and many other ways there are patterns of dependence of sensory stimulation, on the one hand, and movement on the other.

Perceivers are familiar with these patterns of “sensorimotor dependence”. We expect changes in such things as apparent size, shape and color to occur as we actively move about.

Our possession of these skills is in part constitutive of our ability to see (and, more generally, to perceive in other modalities). This is confirmed by the fact that we can disrupt a person’s ability to see by causing changes in the patterns of sensorimotor contingency, even as we leave the rest of the perceptual apparatus intact.

This is what occurs, for example, when one puts on inverting lenses of the sort used by Stratton [1897] and Kohler [1951] in their well-known experiments. Eventually one masters the new patterns — one accommodates — and vision is restored.

This example illustrates the ways in which the ability to see (or to perceive) depends on the perceiver’s mastery of the way sensory stimulation varies as a function of movement.

Note from the standpoint of this account, it is possible to offer an account of the sensory modalities. Perception is an activity that depends on the possession and exercise of a very special kind of practical know-how, namely, the implicit understanding of the way sensory stimulation varies as a function of movement. When you perceive an object, your relation to it is mediated by patterns of sensorimotor dependency with which you are familiar. When you see it, your relation to it is mediated by patterns of distinctively visual sensorimotor dependence, i.e. patterns of dependence of sensory stimulation of the eyes on movement of eyes, head and body. Hearing depends on different kinds of sensorimotor contingencies. Touch on yet others.

There is a lot more to be said to motivate and explain this approach; this is the topic of my book *Action In Perception* (The MIT Press, 2005).

Let us now return to the problem of perceptual presence.

Do you take yourself, when you open your eyes and look, to be aware of the whole scene before you, in sharp detail, all at once? The correct answer to this, ought to be: yes and no. Yes, in so far as you take yourself to have a sense of the presence of a richly detailed world. But *no*, in so far as it does not seem to you as if you actually see every bit of detail. There is no such thing as seeing all the detail at once, just as there is no such thing as seeing the tomato from all sides all at once.

Phenomenologically, the world is given to perception *as available*, or as accessible.

To solve the problem of perceptual presence – comprising as it does, a broad range of phenomena – we need to make explicit this feature of the relevant phenomenology.

We visually experience the scene before us as densely detailed without seeing all that detail, just as we visually experience the tomato as voluminous and three dimensional even though we don’t see all of it. The presence of the detailed environment – of the occluded parts of the tomato, of the uniform color of the wall’s surface –

consists, then, not in our feeling of immediate contact with those features, but in our feeling of *access* to those bits of detail. The detail is present *now*, though absent (unseen, out of view, partially occluded, etc.), because we *now* possess the skills needed to bring the relevant features into view.

And not just any old kind of access. The scene is present to me now as detailed, even though I do not now see all the detail, because I am now able – by the exercise of a repertoire of perceptual skills – to bring the detail into immediate perceptual contact. For example, I need but move my eyes, or move about, or direct my attention here or there, to bring the relevant detail to focus. The detail is present because it is, as it were, *within reach*.

The basis of our feeling of access is our possession of the skills needed actually to reach out and grasp the relevant details. As we have seen, we are familiar, as a general rule, with the ways our sensory experience changes as we move. Moving the eyes, blinking, turning the head, moving the body – all this produces familiar kinds of sensory change. Familiarity with the ways sensory stimulation changes as we move is the ground of our perceptual access. Perceivers *know how* to gain access, to make contact, with the environment around them.

This is the key to the problem of perceptual presence: our sense of the perceptual presence of the detailed world does not consist in our representation of all the detail in consciousness *now*. Rather, it consists in our access *now* to all of the detail, and in our knowledge (itself practical in character) that we have this access. This knowledge takes the form of our comfortable mastery of the rules of sensorimotor dependence that mediate our relation to the world, to the surrounding detail, the cat, the wall. My sense of the presence of the whole cat behind the fence consists precisely in my knowledge, my implicit understanding, that by a movement of the eye or the head or the body I can bring bits of the cat into view that are now hidden. And so for the tomato: My relation to the strictly unseen portions of the tomato is mediated by familiar visual laws of sensorimotor dependence. The presence of the tomato to me as a voluminous whole consists in my knowledge of the sensory effects of my movements in relation to the tomato.

Worry: My relationship to the room next door — however strongly I believe or know or assume or feel that it is present — is not a perceptual relation. My relation to hidden parts of the cat, or to the far side of the tomato, in contrast, *is* perceptual, even though I don't actually see these items. For my relation to them is mediated by patterns of sensorimotor dependence. My relation to them is affected by bodily movements. You may wonder whether this account can be quite right. In particular, you might object, my visual relation to the room next door is no less mediated by patterns of sensorimotor dependence than my relation to the tomato is. *What* I need to do to produce the relevant sensory change may differ (e.g. I would need to walk out into the hall and through the door of the room next door); nevertheless, in both cases my relation is mediated by patterns of sensorimotor dependence.

I have been trying to show how a conception of visual experience as relational and nonrepresentational (a conception that takes support from both Pyslyhyn and Gibson, ha ha!) is one that can explain perceptual presence. The key is appreciating that our

perceptual powers are in part constituted by sensorimotor skills that make the world accessible to us.

Now this *actionist* approach – developed by me and O'Regan in our 2001 BBS paper and by me in my 2004 book *Action in Perception* – has been criticized by Pylyshyn himself. Pylyshyn, commenting on our target article, points out that “*most* things we see are things we cannot act upon directly, such as the words in the target article. Of course what we find out through vision may lead to new beliefs and so may eventually affect what we do, but this is not the sort of behavior that people have in mind when they speak of visually guided action...Much of what we see guides our action only indirectly by changing what we believe and perhaps what we want” (Pylyshyn, 2001, 999). I agree entirely. The claim was never that seeing is *only for* guiding behavior, or that we *only see* that which guides behavior or on which we act. The point is that our ability to experience the world depends on our appreciation of the way our sensory relation to things is and would be modified by what we do, by our movements.

Is the actionist, enactive, sensorimotor approach compatible with what Goodale and Milner have taught us about the double dissociation of deficits in dorsal and ventral lesions? Goodale expresses skepticism about this in his commentary on our target article. He stresses that DF “shows nearly perfect visuomotor control in the absence of any evidence that she actually “sees” the form of the object she is grasping” (Goodale 2001, 984). But then, I interpret him as asking, how can we think that the visual awareness is a matter of visuomotor skill? This question rests on a misunderstanding. It isn't our view that visual awareness depends on visuomotor skill, that is, on the ability to reach and grasp what you see. Our view is that it depends on an understanding of the sensory significance of movement, e.g., on such facts as that moving the eyes changes what you see.

Block (2005, his JPhil review of my book) draws on work by Goodale and others to argue that a proper appreciation of the dorsal ventral story undercuts the actionist view. The dorsal stream, he reasons, is the basis of visuomotor know-how. A distinctive feature of this know-how is that it is effective in peripheral as well as in central vision. In contrast, conscious vision seems to be only effective in central vision. But if visual activity includes dorsally based visual know-how, then visual activity “simply doesn't reflect the phenomenology of conscious vision.” And Block adds that “It is hard to see how skilled activity of exploring the environment could *fail* to involve visually guided action.”

But this is exactly wrong. The dorsal stream supports visually guided activities of reaching and grasping and making quick online adjustments. The exercise of those skills is not constitutive of visual consciousness, on my view. What is constitutive of visual consciousness, what gets exercised in our visual explorations, is our understanding of the ways sensory experience varies as we move.

Block bolsters his claims with reference to findings of Goodale and Murphy (1997). As Block explains: “They presented 5 rectangular blocks to subjects at various positions in the visual field ranging from 5 to 70 off the line of sight. They compared accuracy of perceptual discrimination of one block from another with accuracy of grip via a device that measured the aperture between thumb and forefinger as subjects reached out

to pick up one of the blocks. The basic finding is that subjects' grip accuracy is roughly the same at 5 and 70, whereas perceptual discrimination is vastly worse at 70 than at 5."

The thing is, reaching for blocks in the periphery of the visual field is one thing; reaching for unfamiliar objects is another. To my knowledge it has not been shown that there is no drop off in our ability to handle novel or *unfamiliar* objects in the periphery of the visual field. As a general rule, there's no need to foveate familiar objects; we've seen them before. But how often do we handle *unfamiliar* objects without, first, taking a closer look and foveating them? It has been shown, I gather, that optic ataxics show less difficulty with the reaching and grasping of familiar objects than with unfamiliar ones. As I noticed earlier, this suggests, I would say, that optic ataxics rely on foveal ventral vision.

[Add brief paragraph or two on Jacob's critical discussion of me in his paper for this conference.]

I have been arguing that to make sense of the role of experience in guiding action we need to rethink the role of action in experience. In particular, I am suggesting that in order to account for the facts as we know them, we need to appreciate that there are two very distinct ways in which we might understand the relation of perception and action. One way – the focus of Goodale's research – concerns the manner in which visual representations, whether in the ventral or the dorsal stream, serve to make possible (or fail to serve to make possible) reaching, grasping, online adjustments, etc. Another way – the focus of my work, and that of O'Regan – cuts across the dorsal ventral story and suggests that visual experience itself depends on access to the world, an access that requires that perceivers be sensitive to the sensory significance of what they do. If I am right, then visual experience, subserved as it may be by the anatomically distinct ventral system, is a system whose functioning depends on perceiver's implicit knowledge of the sensory significance of movement. Movement and action are at the foundation of visual consciousness.

Part Four: *A final problem: nonveridicality*

We do need a new model of what veridicality or nonveridicality can amount to, if we are to be non-representationalist. Unless we are to reject these categories outright. On a representationalist conception of perceptual content, the question is, in effect, one of matching or satisfaction. How must things be for things to be the way this experience represents them as being? What can we say about the nature of veridicality if we reject the idea that perceptual experiences have representational content?

To take steps towards such an account, consider the visual experience of stars in the night sky. When you look up in the night sky, you don't actually visually experience *the stars*; what you see, rather, are points of light in the night sky, points of light you reasonably take to be stars (or to be marks or signs or traces of stars). The stars themselves do not enter into your experience. We can see why this is so by considering that, for one's experience of the night sky to be veridical (as we would like to think that it is), there is no requirement that stars be points of light, or that they really look like points of light, or that they be located where we seem to see them. That's why the fact that stars

may have gone extinct millions of years ago does not put the lie to our current experience of the night sky.

The case of seeing stars is one where perceptual constancy breaks down. This breakdown introduces nonveridicality. Perceptual constancy also breaks down, for example, when you look down from the height of a very tall skyscraper at cars and people below. The people look to be the size of ants! That is to say, you can't really experience *the people* from the top of the skyscraper. After all, there's nothing in the least ant-like about people (in respect of size). That they look the size of ants is intelligible, of course. This means, roughly, that what you see takes up about the same amount of visual field as an ant would when looked at from a normal upright position. What can't be denied is that this is an *incorrect* experience of the people. For what *would* make such an experience veridical? The actual presence of ant-sized people!

These cases of breakdown in perceptual constancy can be contrasted with cases where there is no such breakdown. Consider the familiar example of the silver dollar seen from an angle. You do experience the coin's circularity (that is, you encounter it) even though the coin looks elliptical from your viewpoint. Despite the discrepancy between the coin's shape and the way it looks, there's no element of nonveridicality in your experience of the coin. Elliptical is just the way circular coins look when seen from an angle. Likewise, when you experience a wall as uniformly colored, even though there are visible variations in brightness across the surface of the wall, as a result of uneven illumination, you do not *misperceive* the wall's color.

But now the puzzle begins to take shape. What licenses us to say that there is no misperception in the case of the coin and the wall, but there is in the case of the stars and the people down below? In both cases, you might insist, the very same sorts of causal processes join us to distal events and objects. If *elliptical* is just the way circular coins look from an angle, and if *variegated in brightness* is just the way a uniformly colored wall looks when light falls on it in this way, then why can we not say that *ant-like* is just the way people look from this height, or *point-of-light-like* is just the way stars look from this distance? How can we defend against the charge that we are simply prejudiced and so arbitrary in our attitude to large distances (or time scales)?

Traditional thought about perception has been governed by an optical projective conception of seeing that is closely allied to the idea that experiences are representations. If you take this conception of experience for granted as a starting point, then indeed it is impossible to find a principled way of distinguishing between the sort of perceptual failure we experience when looking down on people from great heights and the normal, veridical variations in perspective that we experience when looking at objects from different angles. I would suggest that this is a reason – a further reason – to give up the optical-projective conception of visual experience.

Now, we have considered different ways of failing to establish and maintain contact with things in perception. I have argued, for example, that if you don't *understand* what you see – in the sensorimotor sense I have laid out – then sensory stimulation caused by an object won't rise to the level of contact. In the context of the present discussion, we come upon another form of *failure* to make or preserve contact. When we look up at the night sky, I argue, we just don't succeed in making contact with

the stars although we do succeed in making contact with the lights in the night sky. The stars are just too far away! Ditto for the people down below on the ground. In contrast, for the skillful perceiver, a glimpse of a coin from an angle suffices to enable him or her to grab hold of the coin visually, as it were. Of course if you move the coin far enough away, it ceases to look circular, or elliptical, and becomes little more than a spot. This reveals not that coins seen from such and such a distance look like spots, but rather, that at such and such a distance, one's ability to see the coin – to maintain contact with it – breaks down.

[It is relevant that most cases of nonveridical perception are cases of *merely partial* nonveridicality. You misperceive the spoon as bent in water, for it isn't bent. But in thus misperceiving the spoon, you do succeed in seeing the spoon. In exactly the same way, you may fail to visually experience the stars, but not because you are hallucinating. You *do* experience the lights in the night sky. You *are* in contact with them. For example, when you move your eyes away, they go out of view. You modulate your relation to them in this and many other ways. For the vast majority of cases of nonveridical perception, the world *is* at hand, and is present, thus, as a partner in the experience, as content for the experience.

Perfect hallucination (to use Mike Martin's phrase) is altogether different from ordinary misperception. For in perfect hallucination, the mistake is more radical. It is not that you misidentify what you are in contact with; it is that you take yourself to be encountering the world (to be having an experience, with content), when you are not. The mistake, then, is that you take yourself to be *in contact* with something, when you are not. And there is no need, as argued earlier in this paper, to explain that possibility by supposing that you are, in such a case, *really* in contact with something else. There need be nothing there, and so no real experience, no real content.]

On an enactive, non-representational approach to experience, then, an experience is nonveridical when you do not make contact with what you might have thought you do. Perfect hallucination is only misleadingly described as nonveridical perceptual experience. For in such a case there is no genuine experience of the world at all. There only seems to be.