

## *Vision, Self-Location, and the Phenomenology of the ‘Point of View’\**

John Schwenkler, Mount St. Mary’s University

---

### 1 Visual Self-Location

Vashti has been kidnapped. She was dragged from bed, blindfolded, and forced into the back of a van. Once the van came to a stop, she was led out, made to walk up a flight of stairs, and tied down in a chair. Now the blindfold is removed. Vashti opens her eyes and sees the walls around her: they are red. She looks out the window: the ground looks to be a long way down. She glances around the room: there is a desk, a lamp, and a beat-up sofa. Everything is covered with dust. Having thus taken in her surroundings, Vashti thinks to herself: *I am in a red room on some elevated floor. Across from me is a desk, and there is a lamp to my left and a beat-up sofa to my right. The room is not used very often, or else it would not be so dusty.* Vashti knows all this, and she knows it on the basis of what she sees. But exactly what *does* she see? Is what she sees *the very same thing* as what she takes to be the case?

Intuitively, there is one sense of “sees” in which Vashti clearly does not see all of the things she judges to be so. For on any sensible view of the contents of visual experience, Vashti should not be said to see, in the strictly perceptual sense, that the room she is in is rarely used. Vashti judges this to be the case *based on* the visually evident accumulation of dust, but in this judgment she goes beyond the content of her visual experience itself. Is the same thing true, however, of Vashti’s judgments that she is in a red room, up high, to the right of a lamp and the left of a sofa, and so on? Do these “self-locating” judgments relate to her visual experience in the same way as her judgment that the room is rarely used? These are the questions at the heart of this paper.

We can clarify the issue with a bit of philosophical terminology. Following Christopher Peacocke, let’s say that a belief is *representationally dependent* on a mental state just in case (i) the belief and the mental state share a common content, and (ii) the belief is formed “by taking the mental state ... at face value, in respect of this content” (Peacocke 2000: 264).<sup>1</sup> In our example, Vashti’s judgment that the room is seldom used is representationally independent of her visual experience, as the second of Peacocke’s conditions is not met: Vashti takes this to be the case *because* of what she sees, and what

---

\* I have been working on this paper for a long time, and presented versions of it to the dissertation seminar at U.C. Berkeley in 2008, the 2009 Berkeley-London Conference, ASSC 14 in Toronto, and the 4th Online Consciousness Conference. I am grateful to audiences and commentators on those occasions, and especially to Tom Avery, Robert Briscoe, John Campbell, Sascha Fink, Aaron Henry, Mike Martin, Alva Noë, Lynn Robertson, Sherri Roush, Kranti Saran, Miguel Sebastian, James Stazicker, Assaf Weksler, and several anonymous referees for comments and discussion relating to this material. It is with regret that I apologize for not having been able to address sufficiently all the criticisms of my argument that these and other readers have raised.

<sup>1</sup> Actually Peacocke’s definition differs a bit from this, as it defines only representational dependence simpliciter for a belief, and not representational dependence with respect to a given experience or other mental state. But for our purposes the account can easily be modified to cut a bit more finely.

she sees *justifies* this belief, but how often the room is used is not part of the “face value” content of her visual experience. The question whether a thought, experience, or other mental state has a given content is thus equivalent to the question whether there can be judgments with that content that are representationally dependent on such a state, i.e. whether there can be judgments with that content that are formed just by taking such a state at face value. Can self-locating judgments be related to visual experiences in this way? The aim of this paper is to show that they can.

(Before going any further, let me note that while throughout this paper I will consistently use the language of visual experiential contents, I do this simply for the sake of simplicity, and not because my argument hinges on any particular view of the metaphysics of perception. To the contrary, philosophers who deny that perceptual experiences have contents must also be able to make sense of the issues at stake in this paper, and if my argument is sound, then its conclusion will bear also on the question of what an acceptable theory of perception would have to look like that rejects the assumption that experiences have content: philosophers who reject this assumption will have to develop their own versions of the thesis I defend. Exactly what I mean by this should be clearer once the argument is fully on the table.)

Here is how my argument will proceed. I begin in Section 2 by describing two different views of the content of perception: the *Self-Location Thesis*, which holds that visual experience sometimes has self-locating content, and the *Minimal View*, which denies this. At the start of Section 3 I briefly describe my argumentative method, which is a version of Susanna Siegel’s “method of phenomenal contrast”, according to which an adequate theory of perception must be able to explain the differences between phenomenally different experiences. I then identify a kind of experience whose distinctiveness I believe the Minimal View cannot explain: this is the experience of what J.J. Gibson called “visual kinesthesia”, as exemplified most vividly in illusions ofvection. Section 4 then considers several different ways a defender of the Minimal View might try to explain this distinctiveness, and argues that none of them is successful. I conclude in Section 5 by relating my argument for the Self-Location Thesis to some more general questions about self-awareness and spatial representation.

## 2 Two Views

Certain philosophers seem comfortable taking it for granted that visual experience has self-locating content of the same sort<sup>2</sup> as judgments of one’s own spatial location. For example, Quassim Cassam writes in *Self and World* that “in egocentric spatial perception the objects of perception are experienced as standing in spatial relations *to the perceiver*”, and that such perception “can therefore be described as self-locating; in experiencing objects as spatially related to one, one literally experiences the bodily self as located in the perceived world” (Cassam 1994: 52-53). Similarly, Peacocke provides as an example of representational dependence the case of a self-locating judgment, made on the basis of the visual experience of a door, that one is in front of a door. In forming this judgment one is making a claim about one’s own location, and one is justified in this claim by one’s visual experience. Peacocke supposes further that this self-locating judgment can be formed simply by taking that experience at face value, in respect of a self-locating content that it already possesses:

---

<sup>2</sup> By “of the same sort” I mean only that *what is represented* is the same; it could be, however, that perceptual content is unlike the content of judgments in some other respect, say by being non-conceptual.

Consider ... the everyday case in which an ordinary person forms a belief with the content 'I am in front of a door', and does so for the reason that he sees a door ahead of him. His visual experience represents the door as bearing a certain spatial relation to him. This is so even if he cannot see or otherwise experience his own body on this particular occasion. It would still be true that, taking his experience at face value, he would judge that he is in front of a door. (Peacocke 2000: 264)

So according to Peacocke, that *I am in front of a door* can be an element of the contents of visual experience and self-locating judgment alike: simply in having a certain visual experience, I may already be in a state that (as it were) "says" something about *where I am*. In addition to the door and its spatial location, *my* location is among the things my visual experience represents, and moreover that experience represents my location *as my* location, since it possesses the very same self-locating content as the perception-based judgment that that is the (door-relative) location I am in. The same, it seems, will hold for many of Vashti's judgments: that *she* is in a red room, that *she* is far above the ground, that the desk and the lamp and the sofa are in such-and-such locations with respect to *her*. If Peacocke and Cassam are right, then in thinking all of this Vashti may simply be articulating what she sees, and not going beyond it in the way she does when, say, she judges that the room she is in is rarely used.

Let the *Self-Location Thesis* be the claim that, simply in virtue of its perspectival character, visual experience can include the location of the perceiver among its face value contents.<sup>3</sup> In saying that it does this "simply in virtue of its perspectival character" we are alluding to an idea that is implicit in Cassam's view, and that Peacocke makes explicit in the passage quoted above, namely that visual experience can be self-locating even when "the self" is entirely out of view: that is, even when the perceiver's body is nowhere in field of vision. And this idea has seemed to some philosophers to be simply incoherent, and has led them to reject the Self-Location Thesis. For example, here is John Perry:

I see a cup of coffee in front of me. I reach out, pick it up, and drink from it. I must then have learned how far the cup was *from me*, and in what direction, for it is the position of the cup relative to me, and not its absolute position, that determines how I need to move my arm. But how can this be? I am not in the field of vision: no component of my visual experience is a perception of me. How then can this experience provide me with information about how objects are related to *me*? (Perry 1993: 205)

According to Perry, since his body is out of view in the case he imagines, his judgment of where things are with respect to him cannot be representationally dependent on his visual experience; he cannot form this judgment simply by taking what he sees at face value. If Perry is right, then so long as her body is out of view, Vashti's judgments that she is in a red room, up high, and so forth are just like her judgment that the room is seldom used, in being representationally independent of what she sees. To say this is to reject the Self-Location Thesis; it is to deny that visual experience can be self-locating in the sense that Peacocke and Cassam suppose it (at least sometimes) is.

---

<sup>3</sup> Thus the thesis I defend employs a concept of "self-location" that is at more and less strict than its counterpart in Schnellenberg (2007): my thesis says more than hers insofar as it is concerned specifically with the *visual* representation of one's own spatial location; and it says less insofar as I do not discuss the representation of oneself as an *agent*.

Call the view endorsed by Perry, according to which visual experience *cannot* be self-locating simply in virtue of its perspectival character, the *Minimal View*. Against Peacocke and Cassam, proponents of the Minimal View will argue that in many cases it seems possible to characterize the perspectival content of visual experience in a non-self-locating way. Thus John Campbell writes:

... there is a basic distinction that we have to draw here between what I shall call relational and what I shall call monadic egocentric spatial notions. Relational egocentric notions are those that we use when we say, for example, ‘He is sitting on my left’, ‘The chasm yawned before him’, ‘Look behind you’, and so on. These notions specify the person whose right or left, up or down is in question. They are two-place notions: ‘x is to y’s left’, ‘x is below y’, and so on. Now in stating the spatial content of vision, we do not seem to need these relational notions. We do not need the general conception of something’s being to the right or left of an arbitrary subject. Rather, we need the more primitive monadic egocentric terms. These are notions such as ‘x is to the right’, ‘x is below’, and so on. ... [Ordinary human vision] represents things as ‘to the right’ or ‘above’ using the monadic egocentric notions, rather than the relational terms. (Campbell 2002: 184)

Campbell’s argument in this passage is an example of what I will call the *minimizing strategy* in the philosophy of perception. This strategy involves showing how an experience whose content admits a natural description through one set of concepts can be redescribed, perfectly adequately, with a different set of concepts whose commitments are less extensive. (Think again of how we understand Vashti’s judgment that the room is rarely used, or of the perennial philosophical project of trying to show how the visual world can be described as a two-dimensional array.) Thus Campbell holds that we can characterize the spatial content of visual experience adequately just in terms of monadic spatial predicates, without any use of first-personal indexicals like “I” or “my”.<sup>4</sup> This is a way of arguing that the face value content of visual experience is less than the Self-Location Thesis supposes: visual experience can justify judgments concerning where one is and where things stand in relation to oneself, but not by possessing any self-locating content of its own. In holding this, Campbell is on Perry’s side as a proponent of the Minimal View.

But are Perry and Campbell right to reject the Self-Location Thesis? Is it true that *every* visual experience in which the perceiver’s body is out of view has a content that can be articulated without any appeal to “relational egocentric notions”, to concepts that state the spatial location of the perceiver? Against this assumption, this paper will show that the Self-Location Thesis does a better job than the Minimal View of explaining the differences between the experience that usually accompanies self-motion (though of course we can’t yet call it “the experience *of* self-motion”, on pain of begging the question), and the experience of a moving world.

### 3 Gibson’s Challenge

According to the Minimal View, a person’s visual experience never involves, at least not just in virtue of its perspectival character, any representation of where *he or she* is

---

<sup>4</sup> In Perry’s (1993) terminology, monadic spatial predicates *concern* the location of the individual who employs them (because the propositions they figure in are made true or false depending on that individual’s location), but that location is not part of their *content*.

located, or where things are located with respect to *him or her*. Thus for one who endorses this view, Vashti can properly be said to see that *she is in* a red room only in the same sort of sense as, say, a detective who sees a broken window and footprints on the ground can be said to see that a house has been burglarized: in each case the subject's visual experience justifies an ensuing judgment, but the judgment is not formed just by taking that experience at face value. And this is supposed to be made plausible by the possibility, in keeping with the minimizing strategy, of taking an experience whose content philosophers like Peacocke and Cassam will characterize in self-locating terms, and redescribing it using something like Campbell's monadic spatial predicates: Vashti sees not that *she is in* a red room, but just that *there is* a red room (here); not that the rooftops are down below *her*, but just that they are *down below*; and so on. If the Minimal View is correct, then such a redescription is possible for every visual experience there can be. This paper will show, however, that there are cases where describing visual experience in a phenomenologically adequate way requires attributing to it a self-locating content.

Before launching into the argument, let me say a few words about how I am using the concept of a "phenomenologically adequate description" of an experience. When faced with competing descriptions of the content of an experience, how are we to decide whether any given description is adequate, let alone which among the adequate descriptions are to be preferred? In keeping with what Susanna Siegel (2007, 2011) calls the *method of phenomenal contrast*, I will assume in this paper that any adequate description of an experience must *at least* provide one of the best explanations of each of the ways in which that experience differs from other experiences that are different from it. (If there are several hypotheses that are able to do this equally well, then we will need some other criterion to decide between them, but that will not turn out to be necessary here.) Siegel's method is a way of testing whether candidate hypotheses concerning the nature of experience are able to explain the thing they are supposed to, namely what experience is like. Given such a hypothesis, we put before it a pair of experiences that differ in some respect, and ask whether the hypothesis provides a better explanation of this difference than do its competitors: if it cannot, the hypothesis should be rejected.

Bearing this method in mind, let us consider J.J. Gibson's discussion of a phenomenon that he terms "visual kinesthesia", in which certain invariants in the pattern of optical flow give rise to what he describes as a visual experience of self-motion:

Student pilots see where they are going on the basis of this invariant and get better with practice. Drivers of cars see where they are going, if they pay attention. Viewers of a Cinerama screen see where they are going in the represented environment. ... And all of them at the same time *see the layout of the environment through which they are going*. ... The doctrine that vision is exteroceptive, that it obtains "external" information only, is simply false. Vision obtains information about *both* the environment *and* the self. (Gibson 1986: 183)

Though Gibson puts the point in different terms than ours, what he says here about the "interoceptive" character of visual perception clearly supports the Self-Location Thesis – for surely it is impossible visually to experience oneself as in motion without being able visually to experience one's spatial location, as Gibson essentially indicates:

Vision, of course, is also *statethetic*, if one wants to be precise about words, in that it picks up *nonmovement* of the body and its members. But since nonmovement is actually only a limiting case of movement, the term *kinesthesia* will do for both. The point is that a flowing and an arrested optic array specify

respectively an observer in locomotion and an observer at rest, relative to a fixed environment. Motion and rest are in fact what an observer experiences with flow and nonflow of the array. (ibid.)

Thus according to Gibson, *what a person sees* when his or her eyes are subjected to certain patterns of optic stimulation will include apparent changes and constancies in his or her own spatial location – and this of course is exactly the sort of claim that the Minimal View requires us to reject. But is it possible to do this, while still explaining what makes these experiences distinctive? We saw in Section 2 how defenders of the Minimal View offer minimal descriptions of visual experiences whose contents are naturally characterized in terms of static spatial relationships between oneself and other things, by saying instead that what is *really* represented in such experiences is only that certain things are (say) “to the right” or “to the left”, and that the perceiver’s location itself is not a part of the face value content of visual experience. But Gibson’s cases are different than these: they are ones in which the spatial content of visual experience is changing, and this aspect of change over time has to be accounted for in any adequate visual phenomenology. This is *Gibson’s challenge* to the Minimal View: how is it possible to describe these experiences adequately in a non-self-locating way?

### 3.1 The Simple Response

Suppose first that we simply extend the rudimentary minimizing strategy recalled just above, saying e.g. that what a student pilot *really* sees in Gibson’s first case is not “where he is going” but rather that certain things are at one moment *far ahead*, then *a little closer*, then perhaps *above* or *beneath* or *on the right* or *on the left*, and so on. In virtue of having such a series of perspectival but non-self-locating visual experiences, the pilot is able to steer his plane appropriately and keep knowledgeable track of his spatial position. Nevertheless he has no visual experience of that position, nor does his visual experience represent the things he sees as bearing any spatial relations to him. Call this the *simple response* to Gibson’s challenge.

However, the simple response is insufficient to explain the phenomenal contrast between the experiences Gibson describes and others that are quite different from them. To see this, consider Stephen Palmer’s description of the visual illusion ofvection, here induced by placing a subject inside a rotating drum (see Figure 1 below):

... if you were seated inside a large, opaque, cylindrical drum with vertical stripes painted on it, ... and if the drum were rotating, you would soon perceive the drum as stationary and yourself as spinning in the opposite direction inside it. This experience of self-rotation is so compelling that many people become dizzy and nauseous [*sic*], very much as they would if they were absolutely rotating. In fact, however, they are quite stationary; only the cylinder around them is moving. (Palmer 1999: 505)

For simplicity’s sake, and to avoid begging any questions about its content, let’s call the illusory experience that Palmer describes experience “*I*”. A defender of the Minimal View must show that Palmer’s description of *I*, like Gibson’s description of visual kinesthesia, includes elements that can be eliminated through a proper minimizing strategy: for according to the Minimal View it *cannot* be that you “perceive ... *yourself* as spinning”, at least so long as your body is out of view. But if we try to redescribe *I* along the lines of the simple response – thus claiming, say, that what the subject inside the drum *really* sees is that a point on the drum that was first “to the right” and later

“straight ahead” is now “to the left” – the consequent description is neutral between the illusory visual experience *I* and another experience, which we’ll call experience “*V*”, in which the drum’s motion is seen for what it is: and as Palmer notes, before the onset of the vection illusion the subject inside the rotating drum will briefly have a veridical experience of exactly this sort. There are phenomenal differences between *V* and *I* that have to be accounted for, and the simple response is not enough to do this; the Minimal View must be developed differently if it is to withstand Gibson’s challenge.

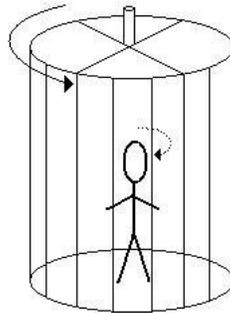


Figure 1: Visual illusion of self-motion (vection) induced by a rotating drum. The drum rotates counterclockwise (solid line), and a stationary observer placed inside it soon has an illusory visual experience that is naturally described as one of himself or herself as rotating clockwise (dotted line) and the drum as stationary. Based on Figure 10.3.2 from Palmer 1999: 505.

Significantly, and as Palmer notes, it would be wrong to think that the contrived nature of this particular setup, or the fact that experience *I* is illusory rather than veridical, means that it should be dismissed as somehow exceptional. Rather, illusions of vection simply bring into stark relief a very widespread phenomenon, which is an essential characteristic of our ability to perceive the world veridically:

Consider driving a car as an example. Your eyes register a highly structured pattern of image motion, which you correct interpret as indicating that you are moving forward in a stationary world. But how do you know that the world isn’t streaming past you while you remain stationary, as in a high-tech video arcade game? ... As in classical induced motion, the visual system assumes that the larger, surrounding world is stationary and that the smaller enclosed object (namely, you) is moving within it. In the driving example, this assumption is correct, so it produces veridical perception of self-motion through the world. In the cylinder example, however, the same assumption is erroneous, so it produces an illusion of self-motion. (Palmer 1999: 505)

Again, however, a proponent of the Minimal View cannot endorse Palmer’s descriptions of these experiences, and so must account for the difference between *V* and *I* in a way that does not credit either of them with self-locating contents. The difficulty is represented in the table below. The Minimal View says that the spatial contents of visual experience can be characterized adequately in the terms of Campbell’s monadic spatial predicates, i.e. terms like “to the right” and “to the left” as opposed to “to *my* right” and “to *my* left”. In this respect, *V* and *I* do not differ at all – yet clearly they are experientially different. So far, the Minimal View fails our test of phenomenological adequacy.

	MONADIC CONTENT
<i>V</i>	Motion from the right to the left
<i>I</i>	Motion from the right to the left

### 3.2 Amending the Simple Response

The simple response presented in Section 3.1 is not enough to explain how the difference between *I* and *V* can be accounted for within the constraints of the Minimal View. But suppose we add to it the observation that only after the onset of *I* does the *drum* appear to be stationary, while in *V* the drum appears to move from right to left. Thus we can expand our table as follows:

	MONADIC CONTENT	DRUM
<i>V</i>	Motion from the right to the left	Appears to be moving
<i>I</i>	Motion from the right to the left	Appears to be at rest

The idea here is that though *V* and *I* are identical with respect to their monadic spatial contents, the difference between them lies in the fact that the drum is represented as moving in the former experience, and as stationary in the latter. In this way, we might suppose, the difference between *V* and *I* can be accounted for within the constraints of the Minimal View.

But is the apparent position of the drum the *only* respect in which *V* and *I* are phenomenally different? It seems not to be: for in addition to this difference there is an aspect of *apparent change* present in *I* that is lacking in *V*, just as the apparent motion of the drum represented in *V* is not part of the content of *I*. That is, in *V* the position of the drum appears to be changing, while nothing else is; and once *I* sets in the drum appears to be stationary, but now there is a *new* aspect of apparent motion, which was not there before. However, the monadic spatial contents of *V* and *I* are the same, so this new aspect of apparent motion cannot just be the world's leftward motion across the visual field, as this was already present in *V*. It is this new appearance of motion in *I* that the perceiver identifies by saying that not only does the drum now appear to be motionless, but also that *he or she* now appears to be spinning. The method of phenomenal contrast requires the Minimal View to explain this further qualitative difference between *V* and *I*, and the simple response remains unable to do this.

Here is another way to make this last point. In addition to *V* and *I*, imagine a further experience *S*, which is the one that the perceiver has before anything in the setup has started to move, and thus in which nothing at all appears to be in motion. The content of *V* differs from that of *S* in that it includes the motion of the drum, while in this respect *S* and *I* are entirely the same. Now assume for *reductio* that *V* differs from *I* only insofar as *V*'s content includes the apparent motion of the drum, while in *I* the drum appears to be at rest. This entails that *I* differs from *S* only by including the appearance of leftward motion. Yet *V* includes this appearance as well: it “overlaps”, as it were, with the apparent motion of the drum. This means that on our hypothesis, the characteristic of *I* that makes it different from *S* is something that it shares with *V*: there is no experience of

motion in *I* that is absent from *V*. Yet this conclusion runs afoul of the phenomenology, and so the simple response remains inadequate.

	MONADIC CONTENT	DRUM
<i>S</i>	Unchanging	Appears to be at rest
<i>V</i>	Motion from the right to the left	Appears to be moving
<i>I</i>	Motion from the right to the left	Appears to be at rest

To sum up, then: whether or not the egocentric spatial contents of *unchanging* visual experiences can be described adequately in the terms allowed by the Minimal View, the phenomenon of vection suggests that some visual experiences are not susceptible to such an analysis. To defend the Minimal View against Gibson’s challenge, we have to find a way to characterize these experiences adequately without attributing to any of them first-personal contents involving the apparently changing (or unchanging) location of the self. In the following section I will consider several other attempts to do this, arguing against each of them in turn.

#### 4 Further Responses

Gibson’s challenge to the Minimal View is to explain how the experience of what he calls “visual kinesthesia” is different from the corresponding experience of a moving world. This section will consider several ways in which a proponent of the Minimal View might try to do this, in increasing order of complexity: first, I will ask whether the relevant differences between experiences like *V* and *I* might be cognitive rather than genuinely perceptual; second, I will ask whether these differences might reside in a non-visual modality; third, I will ask whether they might be “purely phenomenal” differences that are not at the level of content; and fourth, I will ask whether they can be explained by an aspect of visual content that is not self-locating. In each case the answer will turn out to be “No”.

##### 4.1 A Matter of Belief?

A first way to defend the Minimal View against Gibson’s challenge would be to argue that the difference we have isolated between experiences *V* and *I* – *viz.*, the “aspect of change” that is present in *I* but missing from both *V* and *S* – is really a difference in what the perceiving subject *believes*, rather than a difference in perceptual experience.<sup>5</sup> Someone who is attracted to this strategy might be motivated by Palmer’s talk of how we “interpret” visual feedback “as indicating that you are moving forward in a stationary world”, and thereby “know that the world isn’t streaming past you while you remain stationary”. Surely the subject in *I* does not *know* that he or she is spinning and the drum at rest, but perhaps he or she *believes* it; whereas in *V* the subject will believe that he or

---

<sup>5</sup> This is one way of developing an objection that James Stazicker raised in his commentary on this paper at the Online Consciousness Conference, though admittedly without attention to many of its nuances, as a fuller discussion of the distinction between perceptual and cognitive phenomenology would require a paper of its own. Suffice it to say that if this distinction breaks down in a case like this, the Self-Location Thesis is no worse off than the Minimal View.

she is motionless in a moving drum, and this is not the case in *I*. Could these differences in what is believed constitute the further difference between *V* and *I*?

The obvious problem here is that the illusory character of an experience like *I* is not at all dependent on what one believes, as it is possible to undergo this illusion while believing all the while that one's body is at rest. This is an illustration of what Gareth Evans (1982: 123-124) famously called the *belief-independence* of perceptual states, or the fact that whether a person is in such a state "is independent of whether or not he believes that the state is veridical".<sup>6</sup> In Palmer's example, the person placed inside the rotating drum may be entirely aware of the setup, and may not abandon the *belief* that he or she is motionless and the drum spinning even as the illusion reaches its peak. This is enough to show that the relevant differences between the experiences are genuinely perceptual, and not just a matter of belief or "cognitive phenomenology". Gibson's challenge to the Minimal View cannot be met in this way.

## 4.2 A Non-Visual Difference?

I have argued that the "aspect of change" in *I* that is missing from *V* is a difference in perceptual experience rather than belief, but might this difference lie in something other than the specifically *visual* dimension of the way things perceptually appear to be?<sup>7</sup> As an initial way to motivate this possibility, note that *V* and *I* are products of identical patterns of visual stimulation, which might make us think that things should be *visually* the same whether the perceiver is moving in one direction or the things around him or her are all moving in the opposite one. Thus, the thinking goes, it must be in some non-visual modality – like kinesthesia, and perhaps also feelings of nausea and dizziness – that the crucial experiential differences between *V* and *I* reside.

This motivation is inadequate, however, as it runs up against a number of basic intuitions about what it takes for an aspect of experience to count as visual. For example, when you look at a visually ambiguous figure like the Necker cube and its appearance undergoes a shift of aspect from having one apparently protruding side to having another, it is clear that despite the sameness of visual "input" your *visual* experience changes in character: perhaps you are also inclined to reach and grasp the figure differently or form different judgments about it, but these differences seem to be a *product* of the fact that there is a change in how the figure appears in visual experience.<sup>8</sup> Or again, in a much-

---

<sup>6</sup> This is not to say that one's beliefs can have no effect at all on one's perceptual experiences, but rather that they can have such effects only indirectly, and that what it is to experience something perceptually is not constituted by whether one believes that the perceptual state is veridical.

<sup>7</sup> Thanks to Sascha Fink, Sherri Roush, and an anonymous referee for pressing me to take this concern more seriously.

<sup>8</sup> Two points here, with thanks to anonymous referees for raising each of them. First, I am not saying that the phenomenon of the shifting Necker cube cannot be accounted for within the Minimal View: of course it can. But this phenomenon does illustrate the possibility of differing visual experiences that arise from identical visual stimuli, and so blocks one inadequate motivation for the response we are considering. Second, some philosophers hold that in addition to the change in visual aspect that accompanies the shift in the appearance of the Necker cube, there is also a respect in which the appearance of the figure appears to remain the same: that is, these philosophers think that the character of the experience does not admit a single, unambiguous description. (For such an analysis, see Merleau-Ponty 1962: 6.) And perhaps the same is true of the shift from *V* to *I*: that is, perhaps there is a respect in which they are visually the same in addition to being visually different. How can we account for this possibility on the supposition that there is a single thing that I have called the "face value content" of the experience? In fact I

cited experiment Shinsuke Shimojo and colleagues showed that an auditory stimulus that “blinks” off and then on again can make a visually unchanging stimulus appear to offset and onset (Shams, Kamitani, and Shimojo 2002). Phenomena like these show that what makes an aspect of perceptual experience count as visual is not whether it arises directly from optic stimulation, but rather whether it is an aspect of how things *look* (in some strict sense).<sup>10</sup> And it is just such a change in the look of things that subjects in Palmer’s experiment describe.

But there may be a more straightforward motivation for the appeal to non-visual aspects of experience as a way of responding to Gibson’s challenge. For it certainly seems as if differences in, say, kinesthetic sensations and feelings of nausea and dizziness are *among* the respects in which experiences like *I* differ from ones like *V*: so what reason is there to think that there are any differences in *visual* appearance in addition to these? As a first pass at answering this question, note that in illusions ofvection a person’s experience has a particular sort of *cross-modal unity*: unlike a case where, say, it looks like a person is unmoving even as his or her voice seems to retreat into the distance, in an illusion ofvection there is no such discordance among the various ways that things appear to one to be.<sup>11</sup> Rather, what makes an illusion ofvection so compelling is precisely that, in the absence of any significant perceptual input<sup>12</sup> suggesting a lack of change in one’s position, *everything* in one’s experience suggests that the surrounding world is stable while one’s own body is in motion: if things did not appear this way to vision as well, then the illusion would be much less vivid than it is. So it does not seem reasonable to explain the difference between *V* and *I* just in terms of non-visual factors.

A defender of the Minimal View may respond to this argument that even if experiences with cross-modal unity are more compelling than those that lack it, still the representation of self-motion in *I* may be entirely non-visual: it is just that this illusion

---

don’t think that my argument in this paper requires this supposition: that is, it seems possible to allow the existence of multiple contents (or levels of content) and define the Self-Location Thesis as the claim that some of them can be self-locating. One simple way to do this that would fit within the framework developed here would be to appeal to the existence of (either representational or non-representational) properties of visual experience that correspond to the positions of things in the field of view: the idea would be that it is in respect of these properties that pairs of experiences like *V* and *I*, and the different appearances of the Necker cube, are visually similar, while it is in respect of some “three-dimensional” content that they differ. (Note that this would not require holding that the aspects of experience corresponding to visual field position are always part of how things look; perhaps they turn up only when we adopt a certain introspective or painterly attitude.) For more on this idea, though in a different context, see section 4.3 below.

<sup>10</sup> Many complexities arise here; for a discussion of some of them see O’Callaghan (2008). Significantly, Gibson seems not to have recognized this; he thought that for an aspect of experience to count as genuinely visual, it must have been the consequence of information available in the ambient optic array. For a discussion of visual kinesthesia that rests on this assumption, see Lishman and Lee (1973).

<sup>11</sup> Indeed, it seems to be a guiding principle of our perceptual systems that appearances in different modalities are integrated with one another, as e.g. in the McGurk effect. Thus Lackner (1977) found that illusions of self-motion could be produced by a rotating sound field, though not in conditions where one’s lack of motion was visually evident. Similarly, Riecke, Feueressen, and Reiser (2009) found that auditory illusions of self-motion are enhanced by corresponding vibrotactile cues.

<sup>12</sup> Or top-down factors: here see Riecke et al. (2005).

would be even *more* compelling if self-motion were represented visually as well.<sup>13</sup> But if this were the case, then it should be possible to undermine the illusion of self-motion in *I* by focusing one's attention on visual experience (as opposed, say, to the way things sound or the absence of a sensation of the air moving against one's skin). In fact, however, the opposite seems to be true: the illusion is more vivid visually than in other modalities, and the way to undermine the illusion is by focusing on the non-visual dimensions of experience (think e.g. of what you can do when you experience the apparent motion of your moving train).<sup>14</sup> And this is just as we should expect: after all, it's input to the visual system that gives rise to sense of self-motion in the first place, and any experiences of dizziness or illusions of self-motion in other modalities seem to be a *product* of the fact that one's own apparent motion is experienced visually, as opposed to something we could appeal to in order to explain this visual illusion away. So the aspect of motion in *I* that is missing from *V* must be an element of visual experience.

### 4.3 A Purely Phenomenal Difference?

If it is agreed that the aspect of apparent motion in an experience like *I* that is missing from an experience like *V* is at least partly an element of visual perceptual experience, then one option for a defender of the Minimal View is to deny that the differences between the qualitative *characters* of *V* and *I* have to be explained entirely in terms of their respective representational *contents*. If we go this route, we can concede the presence of an aspect of motion in *I* that is absent from *V* without identifying this with a visually represented change in the position of anything in particular: instead, it can be identified with a purely qualitative aspect of visual experience itself.

In motivating such a response, the most natural place to look is to Peacocke's *Sense and Content*, where he argues as follows for the existence of non-representational aspects of visual experience:

Suppose you are standing on a road which stretches from you in a straight line to the horizon. There are two trees at the roadside, one a hundred yards from you, the other two hundred. Your experience represents these objects as being of the same physical height and other dimensions ... Yet there is also some sense in which the nearer tree occupies more of your visual field than the more distant tree. This is as much a feature of your experience itself as its representing the trees as being the same height. ... It is a feature which makes [Irving] Rock say that the greater size of the retinal image of the nearer tree is not without some reflection in consciousness, and may be what earlier writers such as Ward meant when they wrote of differences in extensity. It presents an initial challenge to the Adequacy Thesis [i.e., the view that we can characterize an experience adequately simply by giving its content], since no veridical experience can represent one tree as larger than another and also as the same size as the other. ... (Peacocke 1983: 12)

According to Peacocke, the visual experience of a faraway tree differs from that of a nearby tree of the same size in ways that go beyond any differences between those experiences' respective contents: that the one tree appears to be farther away than the other is *part* of what makes them different, but there is a further experiential difference

---

<sup>13</sup> Thanks here to Kranti Saran.

<sup>14</sup> Of course this might be wrong: it is an empirical question, rather than an a priori one.

that outstrips this difference in representational content, and is determined by the spaces that things take up in the visual field. It follows, he says, that there are aspects of visual experience that go beyond the way it represents the world as being. This account is controversial, but let's suppose for the sake of argument that visual experience does have purely qualitative aspects of the sort Peacocke identifies here: will this concession be enough to save the Minimal View?

Following Peacocke's lead, a defender of the Minimal View would respond to our objection along the following lines: "You have insisted that the only way to explain the aspect of change found in *I* but not in *V* is to say that only in *I* does the subject have a visual experience of *his or her* motion in an otherwise stationary environment. But this argument overlooks the fact that visual experience has purely qualitative features that go beyond the ways it represents things to be, a feature corresponding to the places of objects in the subject's visual field: thus in the illusory experience *I* there is, in addition to the apparent motionlessness of the surrounding environment, also the leftward visual field movement of the lines inside the drum. This is the feature of experience that makes Irving Rock say that despite the visual experience of objects' constant positions across movements of the perceiver, 'this angular motion [of objects with respect to a frame of reference defined by the subject's eyes] is in some sense perceived because we are aware that objects are changing their location in the field of view' (Rock 1983: 257). And it is only this qualitative feature that you are identifying when you note the visually apparent change in *I*; we need not treat this experiential change as a case of changing experiential contents."

So the defender of the Minimal View now argues that the aspect of change in *I* that is missing from *V* is not a change in the visually apparent location of the perceiver, but rather a purely qualitative difference deriving from the changing positions of things in the visual field. But as soon as we state the response in this way, it becomes clear that it does not advance the Minimal View even one step beyond the position we already rejected: for these allegedly "purely qualitative" changes in the character of visual experience are, *just like* the changing positions of things identified by Campbell's monadic spatial predicates, clearly shared by *V* and *I* alike. In *each* of these experiences the enviroing objects are changing their positions in the subject's field of view, and so according to a theory that postulates a purely qualitative dimension of experience of the sort in question *both* experiences will therefore involve corresponding experiences of leftward motion across the phenomenal visual field. Since this characteristic is had in common by the two experiences, it cannot be used to explain the aspect of change in *I* that is missing from *V*. Allowing the possibility of divergences between character and content does nothing to help the Minimal View to overcome Gibson's challenge.

#### **4.4 Some Other Difference in Content?**

This seems to be enough to show that we cannot respond to Gibson's challenge without allowing that illusions of vection (and corresponding veridical experiences) involve a change in the visually represented position of *something*, but might it be possible to defend the Minimal View by holding that that experiences like *I* do not represent one's motion *as self-motion*, as opposed to the motion of something more non-descript? This would count as a version of the Minimal View insofar as it eschews the use of explicitly first-personal language in stating the face value contents of visual experience, and so treats the relationship between, say, an illusion of vection and the judgment that I am moving as one of representational independence: for if the visually represented motion in the vection illusion is not represented as *my* motion, then the visual experience and the visually-based judgment will have different face value contents. But

the challenge for this version of the Minimal View is to specify what, if they are not self-locating, the relevant contents of visual experience could be.

To see the force of this challenge, consider a response to it that would obviously be inadequate, namely that part of the content of *I* is that *someone* (or: *something*) is moving, while this “someone” does not appear to be moving in *V*. Denoting the difference between these experiences in this way clearly respects the constraints of the Minimal View, but it simply pushes the problem back a step: for example, the veridical experience of an otherwise stationary environment in which a person *other than* oneself is moving is also a case in which “someone” appears to be moving, and so we need then to explain how the illusion ofvection is visually different from an experience of *this* sort. An adequate defense of the Minimal View will require describing illusions ofvection in a way that accounts for their experiential distinctiveness – i.e. the way in which the sort of visual experience that usually accompanies one’s own movement is phenomenally different from that of the motion of an “outer” object – without using explicitly first-personal language in saying where that distinctiveness lies.

Given this challenge, a strategy that can seem more promising is to say that in *I* it is *this person*, or perhaps *this body*, that appears to be in motion in an otherwise stationary environment. The idea here would be that the perceptual demonstrative “*this*” marks the experiential distinctiveness of the person (or body) that visually appears to move, thus explaining how such an experience differs from one where the thing in apparent motion is simply an object somewhere in the field of view. But there are two reasons why this response cannot work. First, it seems to be a condition on the meaning of a visual demonstrative like “*this*” that the object it refers to be experienced visually: yet in the kind of case we are considering the perceiver’s body is supposed to be entirely out of view. Second, even if we reject this intuitive demand, the use of the demonstrative “*this*” alone is not enough to mark what is distinctive about the object of apparent motion in an experience like *I*: for the experience of a moving object in the surrounding world, or even of one’s own body as seen “from the outside”, can equally well be described by using perceptual demonstratives, and even though the demonstratives are being put to different uses in the respective cases that difference is not exhibited, as it were, simply by citing demonstrative contents in this very generic way.<sup>15</sup> Again, what is needed is an aspect of visual experience that is non-self-locating but nevertheless proprietary, in the right sort of way, to the thing that occupies the perceiver’s location as experienced “from within”.

One last way for advocates of the Minimal View to try to address this demand would be to characterize the contents of visual experience partly in terms of the apparent location of something like *the origin of the visual reference frame*, or perhaps just *the point of view* itself.<sup>16</sup> On this account, what subjects in Palmer’s experiment wish to describe as experiences of *their* apparent motion are better treated as experiences of the apparent motion or rest of what Husserl (1989: 166) calls the “*zero point*” of the egocentric spatial framework. Strictly speaking, the contents of visual experience have nothing to say about the spatial properties of the perceiver, but they can involve the

---

<sup>15</sup> This is not even to mention the immense awkwardness of terms like “this person” or “this body” – used here as elements in phenomenological analysis, no less!

<sup>16</sup> Another variant might propose that it is the subject of experience whose location is experienced in these cases, but the objections that follow would apply with equal or greater force to such a suggestion. Tom Avery, who argued for a version of the Minimal View in a paper he presented at the 2009 Berkeley-London Conference, has proposed in discussion that the relevant content of an illusion ofvection might be that “*here* is moving”, but of course it is precisely the content of the spatial indexical “*here*” that changes in a case of self-motion in a way that makes a sentence like “Here is moving” simply nonsensical.

apparent motion of the visual point of view, together with constant locations and changing egocentric positions of the visible objects that surround it.

But can we really use concepts like those of visual reference frames or points of view in describing the contents of ordinary episodes of visual awareness? The problem here is that perception is a way of experiencing how things are *in the world*, whereas reference frames, points of view, and “zero points” instead pertain strictly to the nature of *experience*: they can be useful in psychological or phenomenological analysis, but such analyses regard them not as among the contents of experiences but rather as aspects of their distinctive modes of presentation. Defending the Minimal View in this way would require supposing that visual experience represents not just the structure of the world, but also the structure of visual experience itself: we must think that the point of view, or the origin point of the reference frame, is made manifest in visual experience *as* such a point, as opposed to a location occupied by a particular worldly thing, albeit one with a distinctive role in the experiencing subject’s life. And this would be a bitter pill to swallow. We cannot do justice to our phenomenological intuitions if we insist, with the Minimal View, that self-locating judgments are always drawn only by inference from visual experiences whose face value contents are entirely non-self-locating. Instead, things are as the Self-Location Thesis would have it: reports of one’s own motion can correspond to aspects of visual experiences every bit as basic to their contents as the apparent motion or rest of the things one has in view.

## 5 Self and World

The argument of this paper notwithstanding, the reader may not be able to shake the sense that the Self-Location Thesis simply *cannot* be correct, as the idea of perceptual self-location is just too mysterious, especially as applied to cases when one’s own body is literally out of view. In this last section I want to diagnose this puzzlement, and relate the Self-Location Thesis to some broader questions about the nature of spatial representation.

To the idea that, as Perry puts it, one’s own location cannot figure in the content of visual experience when one’s body lies outside the field of view and so “no component of my visual experience is a perception of me”, we should note that the possibility at stake here is not altogether different from a range of “amodal perception” phenomena that are quite common characteristics of visual experience.<sup>17</sup> For example, vision seems to represent opaque three-dimensional objects *as having non-facing sides*, and partially occluded objects (like a cat seen through a picket fence, say) *as continuous*, even though these aspects of the objects are entirely hidden from view. Similarly, the world itself is visually represented as extending beyond the bit of it that is seen at any moment; it does not appear to be bounded. Of course there is a *sense* in which, in cases like these, there is “no component of my visual experience” that is a perception of the items in question, and it is possible to bring this fact to one’s attention by adopting a particular sort of introspective attitude. But as Gibson notes, what one ordinarily perceives is not so limited: it is rather “an environment that surrounds one, that is everywhere equally clear, that is in-the-round or solid, and that is all-of-a-piece” (1986: 195). Visually representing

---

<sup>17</sup> Importantly, by calling this aspect of perception “amodal” I do not mean that it is not a part of some sensory modality (here, vision). Rather, the purpose is to distinguish perceptual completion phenomena that involve experiential “filling-in” (such as the Kanisza square, where we tend to “hallucinate” the unperceived contours of an object) from those where the missing parts of an object are experienced simply in their absence. (Thanks to an anonymous referee for leading me to clarify this point.) For a careful analysis of this distinction, and a defense of Gibson’s claim that amodal completion phenomena are properly perceptual, see Briscoe (2011).

one's body even when it is entirely out of view may reside at the far end of a spectrum of amodal completion phenomena, but that does not mean it is impossible; and moreover one's body usually *is* partly seen, and can always be brought into view more fully simply by redirecting one's gaze.<sup>18</sup>

I suspect, though, that the deeper source of the apparent mysteriousness of the Self-Location Thesis rests in a common, and often unexamined, way of conceiving the various modes of spatial representation, according to which the very nature of perception bars it from representing the perceiving subject as located in objective space. This is the kind of conception that Jean Piaget seems to be endorsing when he writes that the universe of a young child is “centered on a self, ignorant of itself”, and that it is only thanks to “the functioning of intelligence” that one's conception of the universe ever comes to contain “the personal body aware of its displacements in the unlimited series of permanent solids which have movements independent of the subject” (Piaget 1954: 235, 245). For Piaget, there is a sense in which the egocentric condition of the early infant is never fully overcome: he holds that “the purely perceptual point of view is always completely egocentric” (Piaget and Inhelder 1967: 193), and thus that it is only in conceptual thought that the world is represented as independent of oneself, or oneself as something located in it. On this quasi-Kantian<sup>19</sup> way of thinking, the idea that perception itself, as opposed to *thought about* the perceived world, could be self-locating is simply incoherent.

Piaget's day has come and gone, but there is something in his view that continues to exert a pull on us. For example, in introducing a volume on spatial representation, Naomi Eilan, Rosaleen McCarthy, and Bill Brewer (1993: 3) highlight the importance to self-consciousness of “[t]hinking of oneself from the outside, as one object among others” (emphasis added); and they flesh out the concept of an objective spatial world as follows:

By ‘the world out there’, the external world, we also mean the world as it is independently of our interaction and engagement with it: it is what is ‘there anyway’. The idea that we have such a conception of the world is sometimes expressed in the claim that we have a conception of the world as it is from no particular point of view, and that our conception of the world is, in this sense, objective. (ibid.: 2)

If *this* is what it means to represent the world as mind-independent and the self as an “object among others”, then there is no question that visual experience does not do these things: for it is in the very nature of vision to be perspectival. But there is no reason to assume that there cannot be a middle ground between a world represented in these ways and a world “centered on a self, ignorant of itself” – as if the only way to characterize the spatial content of perception, or for that matter any other mode of spatial representation, were to choose one of these two options.<sup>20</sup> Indeed, the simple fact of perception's spatial constancy – i.e., the fact that things in the world do not appear to move whenever we do – already shows that perception is not “completely egocentric” in the way that Piaget seems to have assumed. What is there, then, to prevent the possibility that perception may also be able to represent not just the stability of the world, but also the location of the self?

---

<sup>18</sup> For the suggestion that this sort of possibility is characteristic of amodal completion in general, see Noë (2005).

<sup>19</sup> It is only quasi-Kantian, because Kant himself seems to have denied the existence of any “purely perceptual” – that is, concept-free – level of experience.

<sup>20</sup> I do not mean to imply that Eilan, McCarthy, or Brewer assume this.

Of course there is a sense in which the perceived self cannot be represented just as *any old* “object among others” – or rather, insofar as it is represented in this way the perceived self will not be represented *as the self*, at all. (Think of unknowingly catching sight of yourself in a mirror, or of your arm in a tangled nest of limbs.) But it follows from the argument of this paper that just as thinking of oneself as “I” can be a way of conceiving of oneself objectively even though still “from a point of view”, our perception of the world, and thereby of ourselves, can remain perspectival without simply collapsing into the self-blind world of the Piagetian infant.

## References

- Briscoe, R.E. Mental imagery and the varieties of amodal perception. *Pacific Philosophical Quarterly* 92: 153-73.
- Campbell, J.J. 2002. *Reference and Consciousness*. New York: Oxford University Press.
- Cassam, Q. 1994. *Self and World*. New York: Oxford University Press.
- Eilan, N., R. McCarthy, and B. Brewer. 1993. General introduction. In Eilan, McCarthy, and Brewer, eds., *Spatial Representation: Problems in Philosophy and Psychology*, 1-22. New York: Oxford University Press.
- Evans, G. 1982. *The Varieties of Reference*. New York: Oxford University Press.
- Gibson, J.J. 1986. *The Ecological Approach to Visual Perception*. Hillsdale: Lawrence Erlbaum & Associates.
- Husserl, E. 1989. *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy, Second Book: Studies in the Phenomenology of Constitution*. Trans. R. Rojcewicz and A. Schuwer. Dordrecht: Kluwer.
- Lackner, J.R. 1977. Induction of illusory self-rotation and nystagmus by a rotating sound-field. *Aviation, Space and Environmental Medicine* 48: 129-31.
- Lishman, J.R., and D.N. Lee. 1973. The autonomy of visual kinaesthesia. *Perception* 2: 287-94.
- Merleau-Ponty, M. 1962. *Phenomenology of Perception*. Trans. C. Smith. London: Routledge.
- Noë, A. 2005. Real presence. *Philosophical Topics* 33: 235-64,
- O’Callaghan, C. 2008. Seeing what you hear: cross-modal illusions and perception. *Philosophical Issues* 18: 316-38.
- Palmer, S. E. 1999. *Vision Science: From Photons to Phenomenology*. Cambridge: The MIT Press.
- Peacocke, C. 1983. *Sense and Content: Experience, Thought, and Their Relations*. New York: Oxford University Press.
- Peacocke, C. 2000. *Being Known*. New York: Oxford University Press.
- Perry, J. 1993. Thought without representation. In his *The Problem of the Essential Indexical and Other Essays*, 205-26. New York: Oxford University Press.
- Piaget, J. 1954. *The Construction of Reality in the Child*. Trans M. Cook. New York: Basic Books.
- Piaget, J., and B. Inhelder. 1967. *The Child’s Conception of Space*. Trans. F.J. Langdon and J.L. Lunzer. New York: W.W. Norton & Co.
- Riecke, B.E., D. Västfjäll, P. Larsson, and J. Schulte-Pelkum. 2005. Top-down and multi-modal influences on self-motion perception in virtual reality. In *Proceedings of HCI International 2005*. Las Vegas, NV.
- Riecke, B.E., D. Feuereissen, and J.J. Reiser. 2009. Auditory self-motion simulation is facilitated by haptic and vibrational cues suggesting the possibility of actual motion. *ACM Transactions on Applied Perception* 6.

- Rock, I. 1983. Unconscious inference. In his *The Logic of Perception*, 240-82. Cambridge: The MIT Press.
- Schnellenberg, S. 2007. Action and self-location in perception. *Mind* 116: 603-631.
- Siegel, S. 2007. How can we discover the contents of experience? *Southern Journal of Philosophy* 45: 127-142.
- Siegel, S. 2011. *The Contents of Visual Experience*. New York: Oxford University Press.
- Shams, L., Y. Kamitani, and S. Shimojo. 2002. Visual illusion induced by sound. *Cognitive Brain Research* 14: 147-52.