Sensing, Objects, and Awareness: Reply to Commentators

Austen Clark
Department of Philosophy
103 Manchester Hall U-2054
University of Connecticut
Storrs, CT 06269-2054

I am very grateful to my commentators for their interest and their careful attention to *A Theory of Sentience*. It is particularly gratifying to find other philosophers attracted to the murky domain of pre-attentive sensory processing, an obscure place where exciting stuff happens. I can by no means answer all of their objections or counter-arguments, and some of the problems noted derive from failures in my original exposition. But a theory is a success if it helps spur the creation of better successors. By those lights this one seems to be succeeding admirably. Would that every author could receive such commentaries!

Two themes resound through the commentaries, and I will focus most of my efforts on them. The first is the relation between sensory feature-placing and the perception of objects; the second the relation between feature-placing and our conscious perceptual experience. Cohen and Matthen lodge numerous objections under the first heading; Rey and Levine post most of theirs under the second one. Of course each author also adds objections unique to themselves, and I shall address as many as can be fit into the space-time region allotted to this reply.

I. Feature-placing and objects

Cohen offers a clarifying and useful distinction between two claims constitutive of feature-placing. First, that perceptual representation picks out some individuals and attributes various features to them (Sensory Individuals) and second, that those individuals are space-time regions (Sensory Places). He is quite right that this distinction was not clear in the book, and that both claims need argument. Both he and Matthen largely agree with the arguments for Sensory Individuals, but disagree with parts of the argument for Sensory Places, including the conclusion. Cohen argues that typically the individuals picked out by visual perceptual representation are "visual objects", while Matthen argues that they are "material" objects.

I think in fact there is less disagreement here than meets the eye, though it is quite important for us philosophers to produce an explicit representation of what that disagreement is, since if it remains merely implicit in our discussions it will surely retain its near invisibility among psychologists. The disagreements arise from the intersection of two conceptual issues: (a) when does "early" vision end? and (b) in order to represent something *as* an object (a "visual object" or a "material object" respectively), what must one's representation include?

Suitable tinkering with answers to (a) and (b) can reduce the appearance of disagreement to practically nothing. The hypothesis on the table is that feature-placing is a form of representation common to vertebrate *sensory* processes, and found in any modality that can solve the Multiple Properties problem. Sensory processes were identified as developments in a given modality from receptors up to tertiary association areas in the neocortex. The paradigmatic example was identified as what psychologists call "early vision". The traffic in areas that are in the ambit of early vision is almost exclusively visual; once we get multi-modal associations, or outputs to motor areas, we are no longer in the domain of "early" visual processes (see Clark 2000, vi-vii, 200-201, 205). Visual feature-placing, then, is a kind of representation alleged to characterize early vision. It is not alleged to apply later than that.

Both Cohen and Matthen admit that a fair amount of the processing in early vision is transacted in a location-based way. Cohen even agrees that it is plausible to think that the individuals bearing the features represented by early processing are space time regions (Cohen 2004, section 3.3), but he says this concession is "quite small". Matthen too admits that locations play a role in early vision, "where the scene has not as yet been segregated into objects" (Matthen 2004, [section IV, 2nd para, ms p. 15]); but he argues that that role cannot be one sustaining reference.

So a simple way to reduce the appearance of disagreement almost to nil is to endorse the happy suggestion of Levine that "both Clark and Pylyshyn are right"; e.g. that *early* sensory processes pick out space-time regions as bearers of the features they attribute, and later ones pick out objects. The latter take more work and are more sophisticated. This in fact was a guiding idea of *Theory of Sentience*,

and it is explicit as well in Treisman's feature integration theory of attention. All of the feature-maps in her model are pre-attentive; and it is the operation of selective attention on multiple maps that allows, for the first time, the construction of "object files" (Kahneman, Treisman & Gibbs 1992). The latter representations are the first ones that could be considered to represent their referents *as* objects.

So what is wrong with this happy picture? The fundamental problem is that the claim that space-time regions are the only individuals picked out by early vision is vacuous unless one gives some independent characterization of what "early" vision includes. Furthermore, under practically any characterization one likes, there is evidence for at least something that seems like "object-based" processing within that domain. The most striking is the evidence for object-based effects on the selections made by selective attention; those selections perforce occur pre-attentively, and yet they pretty clearly require sensory individuals distinct from space-time regions. So even the claim that within early vision all the referents of representations are space-time regions is under threat. There may be ways to delimit "early" so as to exclude the object-based processing that seems also to be "early", but none is immediately obvious.

The other option is to push back against the premise that there is "early" processing that is genuinely object-based; it seems to be so, but perhaps one can cook up all the appearances of object-based processing using ingredients that are nothing but higher-order features. (My contribution to this issue makes various suggestions along those lines.) Just as the "early" in "early vision" is ill-defined, so too is the notion of an object. The conditions sufficient to convict a representation of the crime of being object-based are not at all clear; if this were a court of law, the law on which the conviction is based would be thrown out as unconstitutionally vague.

II. That reference cannot be to regions

The ambiguity of objecthood is best displayed through an analysis of the relevant arguments, and so without more ado it is time to turn to the specific arguments of Cohen and Matthen that space-time regions *cannot* serve as the targets of predication. These in any case I must address. There are three prominent ones: the Gabor patches argument, Matthen's argument from the phi phenomenon, and Cohen's argument from demonstratives.

II.A. Gabor patches

The goal of the Gabor patches argument is to show that one can identify and discriminate two distinct objects even though they have the same location. Such a demonstration would show that perceptual representation employs principles of individuation beyond those of space-time location. Of course the feature-placing hypothesis can concede *that* conclusion immediately; it does not deny that other principles of individuation are employed in stages beyond the early ones. So it is only if the Gabor patches argument shows that space-time regions could not serve as sensory individuals in early vision, or that space-time regions are not the only sensory individuals in early vision, that the argument poses a problem.

Common sense has some difficulty with the notion of two perceptibly distinct objects cohabiting exactly the same place, and it is interesting to notice that all the proposed candidates are rather insubstantial (Cohen's mists, Matthen's reflections, and patterns of light on a screen). The Gabor patches provide the most compelling example, but even here one can throw some doubt on the premise. If is true that one Gabor patch is "superimposed" on the other (if there are edges that form T junctions, for example), then one must appear to occlude the other, and the premise that these two appear to occupy the same location would be false. One would appear to be on top of the other. It is difficult to tell from the pictures in the experimental report (Blaser, Pylyshyn, & Holcombe 2000), but there do appear to be T junctions, so that one of the patches does seem to lie on top of the other. If it is to be a genuine counter-example, this cannot be: the two must co-habit exactly the same space.

The authors report that the patches appear to be transparent, so perhaps this is simply an artifact of photo-reproduction. Let us assume that here the words are worth more than the pictures, and that in fact the patches appear to lie in the same depth plane. Such a case would pose a strong challenge to models of location-based *selection*, showing that some other principle must be used to segregate "one" Gabor patch from the other. It is fairly clear what this principle is: coherent motion. That is, at any moment, all the parts of the pattern that makes up one Gabor patch rotate simultaneously, at the same speed and in the same direction. A Gabor patch is essentially a set of parallel stripes; as its "spatial frequency" changes, the number and width of the stripes varies. It can drop to just one stripe (when "spatial frequency" is lowest) but even then we have two fuzzy edges, in parallel. As spatial frequency increases the stripes narrow, and other stripes are

added, still in parallel. The rotation is coherent in the sense that the edges and stripes in one patch always remain parallel to one another.

It is remarkable that one can pick out one of two such patterns even if they appear in exactly the same place, and track it as it changes color, size, direction, and speed. On the other hand, shape-from-motion is a well-known and powerful effect; when a bunch of bits all move coherently, they are relatively easy to pick out and group together, even if they are disconnected from one another and lie in a background of things moving in all sorts of directions. (This is why a rabbit is well-advised to freeze if a predator approaches, even if said rabbit is hiding in the bushes.) In any case this principle of segregation is clearly distinct from any location-based principle. It provided one of the first counter-examples to Treisman's model of conjunctive search: Nakayama and Silverman (1986a, b) demonstrated that on a screen that had bits moving in different directions, the one red thing moving leftwards (for example) could be picked out in constant time no matter many distractors (green things moving leftwards or red things moving rightwards) there were. Treisman (1988) suggested that the different flow patterns allowed a kind of segregation similar to segregation in depth, modified the model, and later (1993) treats coherent motion as a "surface-defining" property. Items on the wrong "surface" can be ignored; the modification added a new kind of inhibition to the model. In short the Gabor patches example poses a problem to pure location-based selection models, but it is one that has already been recognized.

One final observation: note that if a Gabor patch qualifies as a "visual object" or "material object", then such objects can be rather insubstantial. Strictly speaking such a patch is nothing more than a pattern of light on a video screen; when such a patch "moves", nothing really moves, but the pattern of light on the screen changes, so that something appears to move. Furthermore, what defines "one" such patch is a pattern: the multifarious application of the relation "is parallel to" among all its edges. So processes must find all the edges on the screen and then find the ones that are parallel to one another. It is hard to see how that job could be done with just four or five visual indices, whereas edge detection and parallelism are image based properties that are grist for the mill of feature-placing. So it seems to me that we need feature-placing even to define what makes one Gabor patch "one", after which perhaps a visual index could be attached to that one (rather complicated) "visual object". That is the idea: visual feature-placing does the early image-based data crunching, after which indices, names, and kinds can be applied to the entities thereby revealed.

II.B. Phi phenomenon

Matthen argues that the phi phenomenon shows that "places *cannot* be the subjects for qualities" (Matthen 2004, [section III, ms p. 10]). Again, if the stress is on the word "the", so that we read "the subjects for qualities" as "the only subjects for qualities" then feature-placing is not inconsistent with this claim. Place-times are sometimes subjects, but later there are others too. But the argument has a stronger reading, concluding that space-time regions can never serve as subjects: "Thus, it is objects not places of which qualities are predicated by the visual system" (Matthen 2004, [section III, ms p. 10]), and with this reading I must disagree.

Matthen's argument is that a visual system in which locations are the only subjects of predication could not distinguish between two scenes which we can distinguish. In the first a light goes on at location 1, followed after a short interval by a distinct light going on at location 2. In the second scene, a light goes on at location 1 then moves to location 2. Clearly one can arrange things in such a way that if the only individuals represented are location 1 at the beginning moment, and location 2 at the end, then these two scenes could not be discriminated from one another.

Feature-placing can agree with that conclusion as well, but it does not follow there are no place-times which could serve to differentiate the two scenes. We have the spatio-temporal region *between* locations 1 and 2. Here it is critical to remember that the regions are spatio-temporal; they are extended in both space and time. In one scene something happens in the intervening swath; in the other nothing does. We can represent more than just the end-points. So even if a system is limited to place-times as subjects, it can still represent the difference between Matthen's two scenes.[1]

Even if this particular example fails, there is a deeper issue that Matthen's criticism brings to the fore. Motion perception provides a prima facie counter-example to the thesis that sensory individuals are space-time regions. After all, one defining characteristic of a space-

---

1   The difference is represented by representing movement in the intervening swath, or not. Of course sometimes the two physically distinct scenes are *not* discriminable from one another; with suitable adjustment of times and distances we get the phi phenomenon, or merely apparent motion, in which the visual system cannot discriminate the two physically distinct scenes. Both are represented the same way: as if there is motion in the interval between the endpoints.

time region is that it does not move. Its occupants might, but it cannot. So if one perceives something to move, one must (it seems) be perceiving individuals other than space-time regions.

Actually the major premise of this argument ignores a major complication. Even though space-time regions cannot move relative to one another, the perceptible spatial relations between space-time regions *can* change. They change because the transducers of one modality can move relative to those of another. A simple eye movement displaces visual locations relative to auditory ones, and the shift requires sophisticated and rapid "remapping" if the poor creature is not to get utterly confused (see Driver & Spence 1998). Regions identified visually can change spatial relations to regions identified auditorily. So in that etiolated sense, regions in one modality *can* move relative to those in another.

Nevertheless, there is indeed a problem posed by motion perception. There certainly are episodes in which one perceives one thing to move, slowly and distinctly, from one place to another; and those do indeed require sensory individuals other than place-times. The only sorts of episodes of perceptible motion that would be amenable to a feature-placing account are those in which one has an impression of motion without an impression of a distinct thing that is moving: of "motion, thereabouts", without "that thing moved". But there are indeed examples. Motion in the visual periphery qualifies: it can grab your attention even though you did not see what it was that moved. Features such as "glistening", "bristling", and "shimmering" require motion over time, yet the motion is at the level of texture or of surface gloss, and not of a specific object. Gibson's "optic flow" patterns provide another example: the flow patterns can give a powerful impression of movement (your movement) even though you do not perceive any thing to be moving. But the paradigm example is found in one prominent feature in the phenomenology of motion: when things move so fast they get blurry. If the basketball pass is fast enough (if the player is from UConn) you lose precise tracking of the ball, and see instead an orange blur. Similarly the hands move fast enough that you can no longer see or count distinct fingers; maybe there is webbing between them. For all you can see, the ball loses form and turns momentarily into a blurry elongated orange entity, a streak; and then later, after its streaky phase ends, it settles back into something with distinct borders. That blur, that streak of something moving, even though you cannot see exactly what is moving, is the prime candidate for a motion feature divorced from distinct objects. It fills a four dimensional region, extending from one place-time to

another.

Matthen says "The visual system represents scene 4 as a case of a single thing moving from one place to another" (Matthen 2004, [section III, ms p. 10]). Well, sometimes. I'm not sure this description accounts for the cases in which we see motion as a blur. Is a blur a single thing moving from one place to another? The problem is that when things move that fast, you can't tell what they are. It could be a ball, or it could be an granular orange cloud, or it could be ... any of many other things that would make that blur when they move fast enough. One perceives a four dimensional trajectory--a streak starting thereabouts and ending hereabouts. Useful for determining which way to duck! But it is one feature, placed four dimensionally, not necessarily requiring the work of individuation and reidentification.

So perhaps feature-placing can include some episodes of perceptible motion even though it does not and cannot represent subjects that move from one place to another. Matthen's criticism suggests a final deep objection to this idea. It might appear that in order for motion to be perceptible, one must at some level identify and track at least one object over time. Unless some thing (here and now) is identified, by motion perception systems, as the *same* thing as what was *there then*, one would not have the perception of (or the representation of) something moving from there to here. So even if we have a "motion feature" that is placed by "feature placing", there is an underlying level of object identification required in order to generate that feature. It is required for motion to be perceptible at all.

This is a powerful objection, but in the end I don't think it quite works. Specifically, I think there are other cues to motion that are prior to and independent of the identification of objects. In particular, reidentification (that this thing here is the same as what was there) is not always required.   What are these cues? The pattern of occlusion at leading and trailing edges, changes in shading and in surface contours, flow patterns in textures, Gibsonian optic flow, and so on. These are at a low level, not requiring anything so sophisticated as identification and subsequent reidentification.

II.C. The argument from demonstratives

Suppose you believe that sometimes the reference of a demonstrative in ordinary language is fixed only if it is accompanied by an appropriate sort of perceptual episode. These so-called "perceptual demonstratives" are assigned a value only if the speaker or hearer perceives the intended referent.   Such episodes are tremendously

interesting not only to philosophy of language, but also to philosophy of perception, because they tell us something about the character of perceptual representation. In particular, in order to provide a value for a demonstrative, perception must be able to refer to that value; so such episodes provide a second and independent reason for thinking that perceptual representation must have referential capacities.

Suppose furthermore that you believe in the possibility of what can be called "proto-reference". This is defined by the conjunction of two conditions. First, that in some parts of the perceptual system, the ways in which parts at that level of the system pick out their referents is less sophisticated than the full apparatus of individuation provided in a natural language. Typically they are still representing objects-- typically their targets are, *pace* Matthen, the same things that come to be represented as manifest material objects--but in initial stages the system is not representing anything *as* an object. It might be representing nothing more than a discontinuity of intensity, or a feature-at-a-place, or a feature-cluster, or an oriented surface, or a proto-object. Even the discontinuity of intensity turns out eventually be an edge of an object, but in initial stages it is not represented *as* such an entity.

The second condition necessary to endorse the possibility of proto- reference is to allow that sometimes the perceptual system can pony up a proto-referent--the value of a variable in one of these earlier levels of representation--to fix the reference of a demonstrative in ordinary language. This is the more controversial condition. The idea is that sometimes glimmers of the simpler, more primitive systems leak through and provide values for perceptual demonstratives. "This" sometimes refers to something less than (something not represented perceptually as) a fully individuated manifest material object. And so an anaphorically bound "it" can do likewise. Some examples:

That is a mouche volante, and it is caused by "floaters" in the eye ball.

This is an after-image. It is roundish, yellowish towards its edge, and orange towards its center.

You hear that? It is an aural harmonic.[2]

That is a Kanizsa triangle, and its edges are called "subjective contours".

"It" is the closest thing to a free variable language provides, and

sometimes it can have its value assigned by relatively primitive perceptual referential systems. Archaic genes are sometimes expressed; humans are occasionally born with a full coat of body fur. Perhaps "it" sometimes attaches itself to an archaic entity, a vestige from a simpler ontology.

The point of these preliminaries is to share the pain of Cohen's argument about demonstratives; I think anyone who accepts the possibility of proto-reference has to partake. It can be minimized somewhat (the apparently appalling consequences can be mitigated) but any theorist who thinks (for example) that "visual objects" could sometimes secure the reference of a demonstrative in ordinary language will also face the same consequences Cohen identifies.

What are those consequences? Suppose "this" has its reference fixed by a feature-placing stage, so that it refers to a space-time region. Then, Cohen points out (2004, [section 2.2, 4th para from end, ms p. 9]), sentences such as "This is a hand and it is my favorite body part" or "This is my wife and I love her" appear to proclaim affection for particular space-time regions--the ones occupied by hand and wife, respectively. Furthermore, either "this is a hand" has the same underlying form as "this is red" (in which case it predicates handhood of a region) or we insert an ad hoc "occupation relation" into the analysis of certain predicative VPs when they occur with perceptual demonstratives, making those predicates systematically ambiguous.

Ouch! Let us apply some palliatives. The first is to note that if you believe in any kind of proto-reference, then you will take a sentence such as "this is a hand" or "this is an inkstand" to refer to two distinct things, not just one. Suppose for example you think that "this" names a sense-datum, which is quite distinct from an inkstand. Then sentences like "this is an inkstand" are not as simple as they seem. Suppose we say "this is an inkstand and it is a good big one"; are we committed to the claim that the latter clause predicates good and big of a sense datum? Here is what G. E. Moore said about this sentence:

> When I make such a judgment as "This inkstand is a good big one"; what I am really judging is "There is a thing which stands to *this* in a certain relation, and which is an inkstand, and that thing is a good big one"--where "*this*" stands for this presented object. I am referring to or identifying the thing which is this inkstand, if there be such a thing at all, only as the thing which stands to this sense-datum in a certain relation; and hence my judgment, though in one sense it may be said to be about the inkstand, is quite certainly also, in another sense, a judgment about this sense-datum. (Moore 1918, 13-14)

If "this is an inkstand" refers to two distinct things, then the

---

2   An impression of harmonics of a sound, induced by the resonant frequency inherent to the ear itself. Also called "subjective harmonics". See Buser & Imbert 1992, 87.

subsequent anaphora can be bound to either one. One choice will be less painful: it allows us to avoid the need to feel affection for space-time regions.

The same bifurcation arises for any kind of proto-referent, not just the currently unpopular sense-data. It arises if proto-referents are not invariably identical to ordinary objects. If I am reading him correctly Cohen allows this: "visual objects" are not invariably identical to ordinary objects, and they need not satisfy all the "standard ontological criteria" of objects proper (Cohen 2004, [section 5.2, 3rd para, ms p 19]. I agree with Cohen that this is not a problem for perceptual theory per se; there is no reason to impose ontological standards on the work of psychologists if that work is aimed only at explaining perceptual phenomena. For that job they already have more than enough workplace regulation. But the situation changes if we think ordinary language ever refers to such entities. If it does, then we seem forced into an analysis of such occasions that has something of the form suggested by G. E. Moore.

Here's how. First, we should acknowledge with Cohen that most visual objects will turn out to be identical to objects proper; the "visual object" telephone is (thankfully) just the telephone, identified in purely visual terms. The visual object is not conceived to have auditory attributes, but (an empirical discovery) it turns out that it does. That very thing can also ring.

The problem arises only for "visual objects" that do not also fall under the heading "objects proper", if (furthermore) they are targets of a demonstrative in ordinary language. But it seems to me that both conditions are sometimes met. To use some examples derived from the commentators:

> This is a Gabor patch, and it is on top of that other one. No, sorry, it is transparent.

> (In a phi phenomenon test:) This was a duck, and now it is a rabbit.

> (In an MOT experiment): This one flashed, then it went under the bar. Now it is a circle.

> You see it? That's a reflection, and my favorite tree is in the same place.

> The photographer from Associated Press was the one hidden by that yellow orange afterimage. It was caused by his flashbulb.

> This is a subjective contour, and it forms an edge of a Kanizsa triangle.

All the targets here are "visual objects" (or, as Matthen would put it,

"material objects"): they can be bound to a visual index, they are locatable, they can move, and they can be tracked as they move. But none of them meet the ontological standards of objects proper; Moore (1993) would point out that none of them are to be "met with in space" even though they are "presented in space". (None of them preclude other things from being in the same place.) All of these examples require some Moore-like maneuvers in the analysis: we take "This is an *F*" to refer to two things, standing in a certain relation to one another; or we take the subsequent "and it is *G*" to be systematically ambiguous, and to have a second interpretation in the context of perceptual demonstratives bound to visual objects; or both.

Many of these visual objects are in fact pictures of objects--patterns of light on a video screen, named for the object they picture. (This is where Rey's commentary is extremely useful; intentionalism may give us another way of interpreting talk about such things.) For example, the duck that turned into rabbit was never a duck and never a rabbit: at best "it" is a video picture of a duck that became a video picture of a rabbit. So either we add an aphonic phrase ("a picture of") in places in the analysis (on the lines of the aphonic "is occupied by" in feature-placing); or the complements "is a duck" and "is a rabbit" are systematically ambiguous in the context of demonstratives referring to visual objects. Similarly, the "objects" tracked in multiple object tracking are typically *pictures* of objects: patterns of light on video screens. Such a thing cannot literally be "occluded by" anything on the screen, even if that is the natural way to describe what happens. Thus the pain felt in feature-placing spreads, to be shared by others.

The last three examples are particularly close to our target, since they seem to secure reference to an "object proper" via reference to a mere "visual" object. You pick out the favorite tree by finding the reflection; the AP photographer was the one hidden by the after-image; a "Kanizsa triangle" is found only if you find something that has three subjective edges. (By definition a Kanizsa triangle is *not* a triangle, though it appears to be one!) Each seems to implicate two entities standing in relation, like the inkstand and its sense-datum. With visual objects named in the noun phrase one alternative is to allow the predicative VPs to go ambiguous (so that we understand the "place" of the reflection or the afterimage in a different sense than the ones occupied by trees and photographers, and we allow the word "edge" to acquire a second sense, in which it refers to something that merely appears to be an edge). The other alternative is to insert aphonic qualifiers to avert the ambiguities. The "place" of the reflection or afterimage uses "place" in the same sense, but is

understood as the place where the putative entity *appears to be* located; likewise the word "edge" does not acquire a second sense, but when talking of Kanizsa triangles is understood to be talk about an *appearance of* an edge. So these three examples manifest all the appalling features that Cohen has identified. There are indeed problems here, but they are not specifically my problems.

III. Sensing and consciousness

A second big theme echoed in multiple commentaries concerns a problem (or really a set of problems) that one would think would be mine. To put it broadly: What is the relation between what I have called "sensory processes" and consciousness? One would think a book on sentience would answer this! But I fear that many, many distinct problems are huddled together, hiding, underneath this seemingly innocuous query. Indeed some of them are separately named by various commentators: relations to conscious experience, relations to phenomenology, the first-person perspective, the nature of qualia, the explanatory gap, the presentation of appearances, and how experiences seem to one who has them (e.g. why does the reddish character of experience seem to be an intrinsic property?). Answers to one sometimes constrain answers to others, but sometimes not.

One connection between sensing and consciousness is relatively clear: If a creature is awake and sentient then it is a conscious creature. So spelling out what it is for a vertebrate to be sentient is a way to spell out one set of conditions that suffice for creature consciousness. There might be other ways to be a conscious creature, but active exercise of the capacity to sense something is clearly one way that qualifies. So a theory of sentience is a theory of *creature* consciousness. This is the only connection between the two topics that is overt in *Theory of Sentience*.

What about state consciousness--the whole fun business of being aware of things, experiencing stuff, having conscious mental states, having conscious experiences, being consciously aware of things, or having experiences that seem a certain way? These disjuncts are not all equivalent! Here answers are not overt, and this is a source of consternation among some readers. One reason that the book provided no answers is that it did not address these questions: the domain of early sensory processing is treated as a domain distinct from that of conscious experience. Certainly one can give an account of sensory representation without giving an account of conscious experience. But

I seem to have failed to communicate clearly what the theory includes in its domain; this was a major failing in the original exposition. As noted above, visual sensory processes were limned to include only "early" vision. If those processes are pre-attentive they clearly *do not* include everything one might include in visual phenomenology. Rey is right: they do *not* explain the first person perspective. They do *not* explain the predicative structure of conscious visual experience. I think they will be part of the explanation, but they certainly are not all of it.

What then *are* the relations between sensory processes and the whole fun business of state consciousness, in all its varieties? This is a big, new, interesting, question, and I now think that the relations are far more remote than ordinary language suggests. Sensing red is a necessary condition for being aware of something visibly red, and for being aware of seeing something red, and for being conscious of the experience of seeing something red (and so on); but I suggest there are no sufficient conditions in this vicinity at all. That is, I would argue that one can sense red without being aware of anything red, and without being aware of seeing anything red. This I think is the best way to interpret the discovery of chromatic discrimination in blindsight, the efficient discriminations made in the dorsal channel, and the findings from hemi-neglect; but to spell it out takes much separate argument (parts in Clark, forthcoming). If all that argument is correct, sensing can be entirely divorced from state consciousness. So sensing does not entail any of the fun business, under any of the headings. Being aware of what one senses takes more than merely sensing it.

III.A. Phenomenal and phenomenology

In some dialects the word "sensation" entails consciousness of something; it entails some variety of state consciousness. If one has a sensation of red one must (in this dialect) be conscious of red, or (perhaps, but less likely) conscious of seeing red. This is one reason to prefer the more neutral term "sensory processes": those can happen without one being conscious of what they represent, and without one being conscious of them as they occur. Nevertheless it is natural to assume that sensing, however named, has some powerful connections with phenomenology, and thence with consciousness; and it is useful to explore some of the old, subterranean plumbing that connects these notions. Caution: wear your rubber boots.

One connection is etymological: the word "phenomenology"

derives from the same root from "phenomenal", as in "phenomenal properties". The latter characterize how things appear: how stimuli feel, or look, or, in general, seem--to use three prominent verbs of appearance. They generally are attributed to the things one senses. The water of the Dead Sea feels slimy. The distant mountains look blue. And so on. The going hypothesis is that when one is being appeared-to in any such way, it is because one is in a sensory state that represents the world to be that way. That sensory state is in one's head, and it has its own properties distinct from those it represents the world to have. As a matter of convenience, I suggested we follow Galen Strawson's (1989) terminological proposal and call the properties of the mental state "qualitative properties", while the properties it represents the world to have--the characteristics the world seems to have to the creature who has that state--are "phenomenal" properties.

The problem though is that the words "phenomenal" and "qualia", like the word "sensation", and all the other words in the vicinity, have both lower-order and higher-order interpretations. By "higher-order" I mean simply that they implicate some mental states that are about other mental states. If the explanatory gap attaches to the word "qualia", then there will be lower-order and higher-order versions of the explanatory gap. On the line I have just sketched, phenomenal properties characterize how the *world* appears, and explanations of those characteristics of appearance are to be sought in how the particular sensory processes work. The subjective contours found in Rey's example of the Kanizsa triangle are perfect examples of this variety of phenomenal property; the fact that the diagram yields that appearance is to be explained by appeal to edge detection mechanisms culminating in cells in V2, which fire just as they would if there were a real luminance edge at that place (see Zeki 1993, 264; Peterhans & Heydt 1989). It is a little trickier to explain why this triangle also appears brighter than its background, but that too is an example of this sort of "phenomenal property" and (at least in principle, if not yet in practice) it too can be explained once we understand how the mechanisms of early sensory processing work.

Subjective contours are registered pre-attentively; massive numbers can be ignored in parallel, as some unique configuration "pops out", and grabs attention from all those distractors (Davis & Driver 1994, 1998). I hear no squawk of protest from my concepts if I describe the case as one in which one is being appeared-to in edgelike fashion, even though one is not at the moment aware of those apparent edges. The qualia involved would be ones of which one is entirely unconscious. But some philosophers require awareness of x if x is to be phenomenal or qualitative; on such a line phenomenal properties are never found in mere sensing, but require both that one senses something and is aware of or conscious of what one senses. These would be one higher-order step up from mere appearance. Call them episodes of "sensory experience": episodes in which one senses something and is aware of what one senses. A step up from *that* requires not only awareness of what one senses but awareness of sensing it. One achieves awareness of one's own sensory experiences. Not only does one experience the Dead Sea, one is also conscious of experiencing it. These philosophers go for a swim, feel the sea, are aware of the sea, and are aware of feeling the sea. Then, in a particularly unfortunate turn, the term "qualia" comes to characterize how those *experiences* appear to one who is aware of them. What is it like to swim in the Dead Sea? "Slimy" is not an appropriate answer. That is how the *sea* feels, what the sea resembles. You need to describe how your *feeling* of the sea feels. The sea feels slimy, but your experience of the sea feeling slimy feels, well, *kablootie*. That's how the experience appears to one who has it.

Now consider two complementary questions, from Levine and Rey, about the explanatory gap. Levine asks

> what is the connection between the various levels of perceptual representation and what's present in conscious experience? When I consciously attend to the looks and feels of things, is this what theories like Pylyshyn's and Clark's are telling me about? (Levine 2004, [section 4, last para of ms, ms p. 20])

I think an account of sensory processing *can* tell us about the looks and feels of stuff (which stuff is later identified as things), but that it alone cannot tell us about attending to looks and feels, much less consciously attending to them. That is, the qualities of which you later become conscious, in post-attentive processing, are within the ambit of this account; but selective attention, and being conscious of stuff, are not.

The account in *Theory of Sentience* addresses the explanatory gap only in its lowest-order manifestation: why does this stimulus look red? Why does that one look green? (As Matthen charmingly puts it, why there is a green region at all?) This is basically the gap between primary qualities and our ideas of secondary qualities, and it can be cast in terms of being appeared-to. The puzzle lies in getting from physics to the kinds of appearances described by "that looks red" or "that looks green". Notice that the terminus can be described in a

feature-placing language. This explanatory gap could be closed if we could derive sentences in a feature-placing language from the underlying physics.

Rey asks about the higher-order manifestations of an explanatory gap, in which (at the very least) one is not only being appeared-to but is conscious of being appeared-to in a certain way. The gap he describes requires the first person point of view. The explanatory gap that "worries the qualiaphiles" he says is one that necessarily involves "first-person phenomena"--why appearances "present themselves" as they do, for example. (Rey 2004, [section 2.3.(i), ms p. 4]) One prominent version he mentions: explaining why experiences of color seem to be experiences of an intrinsic property.

I think Rey is quite right that the account in *Theory of Sentience* does not and cannot solve these higher-order versions of the explanatory gap. The lower-order version can be expressed in a feature-placing language, but such languages cannot provide even a first-person pronoun, much less the first-person point of view. An account of sensory processing will not account for all the vicissitudes of consciousness; I would be happy to account for the humble phenomena of being appeared-to redly.

It seems clear to me now that pre-attentive sensory processing on its own cannot even account for being *aware* of something red; the latter requires distinctions introduced only with the operation of selective attention, and that's post-attentive. So Rey is right that this account does not explain, and does not even address, the first-person point of view. He is right that I have not explained how appearances present themselves to someone who is aware of them. He is right that we need additional materials to explain why experiences of red seem to us to be experiences of an intrinsic property.

In a similar vein, Siegel (quoted by Cohen) is right that sentience limited to feature-placing will not suffice to explain a subject's "visual experience"; the latter implicates attention and other processes beyond mere sensing. The account does not aim to explain the "predicative structure of conscious visual experience", as Matthen puts it (2004, [section IV, 2nd para, ms p 15]). Some of the misleading advertising can be traced to the ambiguities in all the available vocabulary. In one sense the account can rightfully claim to explain facts of "sensory phenomenology": it aims to explain phenomenal properties, the structure of appearance, à la Goodman. But more typically "phenomenology" is taken to include in its domain only those phenomena of which one is conscious; it is delimited to or defined by introspective access. Feature-placing should not pretend to explain

that kind of phenomenology. I apologize for the misleading terminology. (Mistakes were made!)

III.B. The relational account of qualia

Suppose we restrict the application of the terms "qualia" and "qualitative properties" to those properties of mental states in virtue of which stimuli look red or look green (for example). Both Levine and Matthen raise some very useful criticisms of the "relational account" of such properties proposed in *Theory of Sentience*. In the next few sections I will consider these criticisms, starting with the ones that can be answered relatively easily, and ending with ones that can only be answered with great difficulty, if at all.

First, Levine asks:

> what discovery is it that Clark can point to that lends the kind of support to the relational type-identity thesis sufficient to overcome the apparent possibility of a red-only experiencer? (Levine 2004, [section 2, ms. p 8])

Levine notes and rightly rejects the possibility that the argument is based purely on a modal intuition (that the relations between colors are essential to them). Instead he is right: it is empirical findings. I take the experiments of Wallach and Gilchrist (and for that matter of Land 1985, and also now of Whittle 2002) to be quite revealing. They strongly suggest that our intuitions that chromatic qualities are intrinsic qualities are just wrong. Granted, the alternative has difficulties as well, but those experiments ought to deliver a shock to one's preconceptions sufficient to raise some fundamental questions about them.[3]

A second set of issues is charmingly described by Levine:

> the next question is what to make of the property we in fact experience - the redness in the experience itself. ... in the end, what does it mean to say that what the visual system is doing is color coding spectral reflectance differences in the world? What has the colors in this case? As in the old Pepsodent commercial, you wonder "where the yellow went?" (Levine 2004, [section 3, next to last para. of that section, ms p 15])

---

3   I hope, by the way, the argument is not simply special pleading on behalf of materialism; I think Wallach's results pose problems as well for a dualist who thinks chromatic qualities are intrinsic non-physical properties. The ontology is not at issue; the issue is the form that description of such qualities has to assume if we are to explain the experimental results.

To this last question I can only hope that the yellow stays right where it is! That is, the account is intended to provide a theoretical identification, a reductive account of chromatic phenomenal properties, and it is not meant to be eliminative. So to the penultimate question--what has the colors?--again the intended answer is: the same things that have them now, whatever they are. (I think they can be space-time regions, surfaces, proto-objects, and objects: all of the above.) The color bearers are, I hope, as they were. Two other issues raised in this passage are more difficult. The first one it mentions--the redness in the experience itself--has already been acknowledged to be beyond the reach of feature-placing alone. That is, if our target is the *experience* of something red, then presumably the subject of the experience is not only sensing red but is also conscious of it, and the latter is not achieved in preattentive processing. The opening query is also open to a higher-order interpretation, suggested by Rey, in which we ask how the experience of seeing something red presents itself to one who is conscious of seeing something red. I agree with Rey and Levine that we need an account that answers this question, and that feature-placing is not it.

III.C. Same and different

The question raised in the second sentence in the passage above is the most difficult of the lot: it reiterates some more specific criticisms of the relational account made elsewhere both by Levine and by Matthen. To put it bluntly, the reduction of chromatic phenomenal properties to relational definite descriptions seems to leave out the colors themselves. There are two cases: what do we sense when we sense a chromatic similarity, and what do we sense when we sense a chromatic difference?   In neither case does the bare relational description seem to capture the color itself.

First, chromatic similarity. Levine says that in such a case, our visual system

> doesn't just say to us--"same old same old"--it says how it is the same, it's red, or red31. It looks a certain particular way in its sameness, and different from the way that other patch looks in its sameness. (Levine 2004, [section 3, 3rd para from end of that section, ms p 15])

Matthen raises a similar criticism. Whittle (2002) showed how a patch in a circle of eight from the red-orange part of color space could be made to look the same as a patch from a circle of eight in the blue-violet region.

Whittle concentrates on the fact that they look the same, clearly an important discovery. But notice that they don't just look the same: they look *green*. If contrast is all that is being coded, why should this additional information be provided? ... Why is there a green area? (Matthen 2004, [section VI, third to last paragraph of ms, on ms p 29])

Why is it not only *similar to that* but also, specifically, green? Second, chromatic difference. Again the mere relational description seems to leave something out. Here is how Levine puts it:

> it isn't pure difference alone that is represented by border regions. We don't just say, this is different from that, but we describe the difference in terms of a quality - our visual system says, "redder here than there". But what's being redder? (Levine 2004, [section 3, ms p 14])

And Matthen tightens the screws on the question:

> since the same *difference* can occur anywhere in the colour map, this leaves it mysterious why a certain pattern should look brown-and-orange rather than turquoise-and-olive (assuming that the differences within these pairs are the same) (Matthen 2004, [section VI, second to last paragraph, ms p 29])

The relational account seems to treat chromatic differences as colorless, but there are different kinds of difference. "Redder than" is different from "browner than" even though the "size" of the differences might be the same. Putting the problem in a nutshell, Levine says:

> sensory qualities have an "absolute value" as well as a relative one; a determinate identity over and above their set of similarity relations. (Levine 2004, [section 2, penultimate para, ms p 11])

Excellent criticism!

A premise essential to my reply is that the human color space is in fact asymmetric and so non-invertible. (The counterfactual case of a symmetric and invertible color space will be considered in the next section.) Asymmetry implies that if we extend the relational description far enough, and include enough relata, the description becomes a definite description; it is satisfied uniquely. Given the structural definite description, there is just one color (here, one phenomenal property) that stands in exactly that place in the structure of relations among the colors.

A second premise is mentioned in passing by Rey (citing Clark 2000, 13): there are aspects of the structure of relations among chromatic phenomenal properties that are not readily accessible to introspection.[4] Some of them might be completely invisible. The fact

---

4   This was meant not to express a reservation about the preceding argument, but rather to drive home the point that the study of the structure

that the actual human color solid is non-invertible is among them. Introspectively, our color solid appears indistinguishable from one that is symmetric. In particular, the fact that for each color there is a structural definite description--one which *only* that color satisfies--is not introspectively obvious.

Suppose we describe the qualitative structure of colors in a purely relational way, using variables as place holders for all the color names. In accord with the second premise, to us it will seem entirely arbitrary how to assign colors to that structure. No point is labeled at all; red might go here, but it could just as well go over there. The actual assignments of particular chromatic qualities to particular places in the relational structure seem wholly a matter of "arbitrary Determination", as Locke put it; he attributed them to "the arbitrary Determination of that All-wise Agent, who has made them to be, and to operate as they do, in a way wholly above our weak Understandings to conceive" (Locke 1975, Book IV, ch. III, para. 13). If this assignment is indeed arbitrary, then given any relational description, it is still an open question of what color to assign to it.

But the first premise implies that the relational structure descriptions are in fact definite descriptions; there is nothing arbitrary in the assignment of colors to structure descriptions at all. So given a (sufficiently full) structure description, there does not remain an open question of what color to assign to it; that color has already been named. The structure description named it.

This point is not intuitively obvious, which is why the objections of Levine and Matthen seem so powerful. There appears to be considerable arbitrariness possible in the assignment, just as spectrum inversion appears to be nomologically possible. This appearance is what I labeled (in Clark 2000) an illusion: the intellect is bewitched. It requires careful treatment. Suppose you are given a completely unlabeled structure description--one which consists solely of relations like "x matches y" and "u is more similar to v than to w". It will then indeed seem possible arbitrarily to assign the chromatic quality red to any of the placeholders u, v, w, x, y. But notice how this appearance of arbitrariness goes away once one or two colors are assigned. Once you know which point presents red, it is pretty much determined where green has to be. Given a fix on both red and yellow, there is only one place where orange could go.

The next step is to suggest that the apparent arbitrariness of

---

of phenomenal properties is distinct from introspective phenomenology, even though it too can be called "phenomenology".

assignment in fact goes away once the structure description is complete; there is never a real choice possible, even in assigning the very first color name to the structure. Red provides the prime example. There is a unitary red, and among unitary hues, red can be more saturated than any other. That a hue is unitary will show up in the structure descriptions (see Clark 2000, 25, 256); distance from the achromatic core of the color solid will as well. So in fact there is no choice in the matter: if it really is the color red, the favorite example of philosophers, then it has already been identified by that structure description.

Consider then the case of sensing chromatic difference. Levine and Matthen are quite right: one does not simply sense a "pure" difference. Suppose (to oversimplify) the color solid has three dimensions: red to green, blue to yellow, and white to black. The difference between any two colors would then be a vector (a oriented line of a particular length) with three components: degrees of difference in red to green, in blue to yellow, and in white to black. It may seem as if it is an open question what kind of difference this is--is this one redder than that one? Is this the difference of turquoise compared to olive, or is it brown compared to orange? But if the argument above is correct, then the structure description in fact, and contrary to appearances, already settles these questions. Given that difference, we already know that it is of the sort "redder than". This one has to be turquoise v. olive, and cannot be brown v. orange.

A similar moral applies to similarity judgments. The visual system does not just say "same old, same old"; at the very least it is similarity in various respects (the ones that can be discriminated) and to certain degrees (from indiscernibility, to matching, to resemblance). One specifies a point in quality space and distances along each axis within which candidates could allowably fall. But once you have specified that, you have already specified that they are both red31 or (given a bigger ambit across hues) that they are both red. If you are handed such a structure description, it is not an open question what kind of similarity you have in hand. The kind of similarity has already been identified.

III.D. Relationalism in symmetric spaces

The previous reply relied on the brute fact that human color space is not invertible, and this reliance exposes the account to a final, particularly difficult objection. After all, as Levine points out, it seems perfectly conceivable that humans have a symmetric color space. This

poses a problem for the relational view:

> If we have a perfectly symmetrical space, our problem is to say what it is that distinguishes orange from turquoise in the first place. If you say, well, orange is between red and yellow, then we just push the problem back to these hues. What distinguishes red from green and blue from yellow? We know we have two distinct regions in the space, but no principle for identifying one as the red-orange-yellow arc and the other as the green-turquoise-blue arc. Relationalism leaves the actual identities of the color qualia undetermined (or underdetermined). (Levine 2004, [section 2, ms p. 7])

Similarly, Matthen's question "why is there a green area at all?" is particularly effective when applied to a symmetric quality space. Why is this end of the axis green and not red? Why is the spectrum colored at all?

Suppose the quality space is symmetric with respect to the red-green axis, so that the "red" end of the axis satisfies exactly the same relational structural description as the "green" end, even when one takes into account relations to every other point in the quality space. In such a case, standard spectrum inversion would be possible: red would have all the structural properties of green, and a structural definite description could not uniquely identify either one of them. The problem generalizes: if inversion is possible then most of the colors in the quality space would, in this way, have an "inverse". Nevertheless, even in this case the relational view does not imply that the two endpoints are qualitatively identical to one another; it does not imply that red = green. Instead the structure of the quality space still yields two distinct endpoints of the axis, both of which satisfy structural descriptions of exactly the same form. The failure of the relational account is rather more limited: relational description would fail to identify which one of the two is red. It still implies that the axis has two endpoints, and that they are distinct from one another.

In such a world there *is* a qualitative difference between the sensation of a color and the sensation of its inverse, and the account implies that there is a difference, but it is at a loss to describe exactly what that difference is. It cannot describe that difference with precision sufficient to identify which one is which. Any manner of difference which the relational account might identify to distinguish the red end from the green end also applies equally well in the converse direction. So it can tell you that there are two distinct qualities, and that they are distinct, but it cannot say which one is red.

Is this consequence so bad? The interesting point is that the same would apply to any person inhabiting that world. One's grasp of the intrinsic property in question goes indeterminate in the same way. Put it this way: there is of course a difference between red and green, and it might seem that you can pick out the intrinsic property GREEN by simply focusing your attention on it. GREEN is the one that looks like *this*. Red looks like that. We suppose that the demonstrative and the focused attention uniquely identifies the intrinsic property in question.

But the problem is that for all you know this ceremony, in that world, picks out the intrinsic property that everyone else picks out when they pick out RED. Remember in that world inversion is possible. So even though the two colors are obviously distinct from one another, your grasp of the intrinsic property GREEN is not as firm as you might think, since for all you know you could be holding onto RED instead. You could establish that there is a difference between red and green, but you could not in any way latch onto that difference with sufficient exactitude to pick out exactly one of them. The same is true of the relational account: it too would imply that there are two distinct endpoints of the red-green axis, but it could not provide a relational description true of exactly one of them. Unless there is more to it, I do not see how this counterfactual example provides a decisive refutation to the relational view. Failure to provide a uniquely identifying relational description in this counterfactual case is not clearly a failure of the account. Instead, perhaps its failure is ours too.

IV. Some particular replies

That concludes the common themes; what remains is a dauntingly long list of comments and objections each of which is found in only one commentary. Since it is impossible to address all of them, I will limit replies to the most important ones for which I have some sort of answer.

IV.A. Rey and intentionalism

I was distressed to be read as being opposed to "Intentionalism", for I certainly do think that sensory states are intentional states, that they have a content, that they are more or less veridical, that they can be evaluated semantically; indeed that they have something like a subject-predicate structure, though they are not sentential and do not manifest most of the hallmarks of compositionality. What is the relation between qualitative properties of sensation and phenomenal properties (characteristics of how things appear)? The simplest answer is that the former represent the latter. Phenomenal properties are part of the

intentional content of sensory states. The point of the book is to make the case that even sensing is a kind of representing, and (more importantly) to detail what kind it is. It is a kind that differs in interesting ways from sentences, and from object-based perception, and these differences help explain some of the classical distinctions between sensing, perceiving, and thinking.

One way in which I seem to be opposed to "Intentionalism" is that I think we can do without what I call the "third variety" of visual field. The third variety treats the visual field as a sum of the intentional objects of visual representation: the world as it is represented visually, including all the things and features one merely *seems* to see. In many contexts we need to talk about the intentional content and the intentional objects of visual representation; the terminology is often useful and is, in any case, inevitable. The problem with the third variety of visual field is, therefore, not the terminology it employs but rather the ontological status it grants to the objects that terminology purportedly names. In the one passage that Rey quotes to demonstrate my opposition to "Intentionalism" I say that we do not need to quantify over sense data, Meinongian objects, phenomenal individuals, or merely intentional objects in those cases in which sensory reference goes awry. That is, when visual representation is less than veridical, we do not need to enter existence claims for its intentional objects. A disagreement, if there is one, arises only in these cases--only when the representations are false--for when they are true the "objects" they are about are just the "objects" that in fact do exist in the field of view-- e.g., for feature-placing, the space-time regions in or about the body of the sentient organism.

So a disagreement, if there is one, can be confined to a very narrow compass: how to treat talk about the intentional objects of representations when those representations are less than veridical. Particularly interesting are those cases in which a representation refers to nothing, employs a vacuous name, or names an impossible object. (I think these cases provide the central focus for what Rey means by "intentionalism"--it's not a general claim equivalent to the representational theory of mind.) Rey argues that even those representations are, in ordinary parlance, about *something*; "there are things that don't exist" can be a true sentence in standard English usage, and perhaps we are talking about one of those things. Nevertheless it is not the case that there is an x such that we are in that case talking about x. We are talking about something but there is not something we are talking about. My rejection of the third variety of visual field was based on nothing more than standard Quinean strictures, which frown on such (apparent) frivolity. Perhaps Rey has a new and better way of talking about this unsettling domain.

It would help, since so much talk about what we sense, or perceive, or think about is talk about entities in this domain. Focus your attention, please, on this Gabor patch, not that one. This Gabor patch is oriented horizontally, not vertically. Such talk is heavily invested in the esoteric market of intentional objects. "This" Gabor patch is merely something pictured on the screen; it is not literally "on" the screen. If you doubt this, think for a second about what it means for this Gabor patch not to be on top of that one, but instead to be transparent. (Focus your attention, please, on all the transparent things on your screen.) Likewise, whatever "horizontal" means in this context, it surely does not mean "aligned with the horizon". Probably it means "aligned with the bottom edge of the screen"; and notice that the thing that is aligned with that edge is only *pictured* to be aligned with that edge. So here we are, giving instructions to our "naive" experimental subjects, using a language in which the edges of an object as represented--the object pictured--have geometrical relations to the edge of the piece of plastic which houses the cathode ray tube which we use to produce the stimuli of the experiment. If deflated intentionalism can help make sense of this sort of talk, then many, many experimentalists would be friends of deflated intentionalism.

So I look forward to developments in Rey (forthcoming); as far as I can tell, there is nothing in feature-placing that is unfriendly to it. The account already allows that sensory states have a content even in cases in which they have no referent (Clark 2000, section 5.5.4), and in the end I think Rey too blanches at the prospect of quantifying over *merely* intentional objects--the ones named by vacuous names. Instead "REP" is systematically ambiguous, and "is" sometimes means "subsists".

IV.B. Matthen and retinotopic v. retinocentric

Matthen's detailed commentary is extraordinarily useful; many sections make notable advances to the clarity of our understanding of sensory processes. The parts on visual field places, directions, and uniqueness of place are outstanding. It is fun to have a Kantian argument for the distinctiveness of visual location. His discussion of audition and somesthesis highlights a critical ambiguity in the term "feature map". In one sense they are "maps" because the fibers streaming into them are organized to respect the topographical organization of the receptor surfaces; in a different sense some of them

are also "maps" of space in or around the organism. Some of the former "maps" do not have the latter characteristic. My description of auditory feature maps was focused on ones found in the barn owl and brown bat, which do have both characteristics.

One line of Matthen's argument requires some further discussion. After describing feature maps in audition and somesthesis, Matthen introduces a distinction between the content of a map and the information implicit in it, arguing that

> the content of a particular map should not be identified with information that is implicit in that map. Content is rather information that has been extracted and explicitly coded in a form that the organism can use with no further processing. (Matthen 2004, [section IV, next to last paragraph, ms p 20])

Visual maps in V1 extract information from the receptoral array. Matthen argues that the only explicit coding of information in V1 is coding about receptoral locations; information about distal locations is, at that stage, still merely implicit. So maps in V1 are not about distal locations; their content (the information they code explicitly) instead concerns locations in the retina. He says:

> Notice the difference of the terminology used by Clark and myself. I use the term 'retinotopic' in this context, meaning 'pertaining to places on the retina'. Clark, by contrast, uses 'retinocentric', meaning 'represented with the retina at (0, 0, 0)'. ... Clark's terminology suggests different coordinate schemes related to one another by a mathematical transform; mine suggests that the places themselves are different, retinal vs distal. This difference of conception goes to the heart of the issue. Do the feature maps of early vision convey the information that certain qualities are present at certain places on the retina, or do they convey the more indirect message that certain qualities are present at the distal places that correspond to certain places on the retina? (Matthen 2004, [section IV, para 3, ms p 16])

This is a wonderfully clarifying distinction, and Matthen has here identified a disagreement that is not confined to Clark *v*. Matthen, but reverberates throughout large parts of neuroscience. The dispute cannot be settled here, but I will indicate a few reasons why, after reflection, I think I will keep all of my chips piled on the "retinocentric" side.

First, notice two problems with Matthen's development of the "retinotopic" alternative. The distinction between "implicit" and "explicit" coding of information is clear in his example, in which we have both statements and logical derivations from them. But once we shift into contexts in which we have neither statements nor inference

rules, the distinction is not at all easy to make. The coding is explicit if it is "extracted by some computational process" and "put in a form the organism can use with no further processing." (Matthen 2004, [section IV, penultimate para, ms pp. 19, 20]) How does this apply to maps in V1?

> information about distal places may be implicit in the retina. But to extract it, a statement of the form "*F* is present at retinal place *R*" has to be transformed by the application of a background theory into a statement of the form "*F* is present in object *O* at distal place *P*". (Matthen 2004, [section V, 3rd para from end, ms p. 22])

Read literally, this says that explicit coding requires *statements* of particular logical *forms*, transformed by the application of a background *theory*. Surely Matthen does not think that V1 is in the business of making statements; but if it is not in that business, it is not clear how to determine whether information coding in it is implicit or explicit. And that distinction is determinative for the retinotopic/retinocentric dispute. (This issue is not unique to Matthen; Rey too says that the content of perceptual states reduces to a collection of statements.) I think the system of representation in play here is much more map-like than sentence-like. And what information on a map is explicit? Try to apply the distinction to a road-map: if you need the key, or the scale, the information must be implicit; but what is left?

Second, even by Matthen's lights I don't think it follows that maps in V1 are about places in the receptoral array. After all, this skips the intermediary processing in the lateral geniculate nucleus (LGN). This has six layers, three for each eye, with corresponding points from retinas of the two eyes in correspondence across all six layers. Many LGN cells behave in a center-surround fashion; the first chromatic opponent cells were found here long ago (see De Valois & De Valois 1975); and double-opponent cells also (De Valois & De Valois 1993). It would seem to follow from Matthen's reasoning that maps in V1 are not about the states of the retina, but are instead about states of the LGN. How could such a map be about a retinal place, when that information is merely implicit in the signal it receives from the LGN? (And note in any case it could not be about the *receptoral* array, since the receptors are several synapses away from the neurons that make up the optic nerve, which synapse only on bipolar and amacrine cells. Information about the actual receptors is only implicit in those!) We seem faced with a rather uncomfortable regress: signals in the optic nerve are about the states of the bipolar and amacrine cells in the

retina; signals in the cells of the LGN are about states of optic nerve; maps in V1 are about the cells of the LGN; maps in V2 are about the maps in V1, and so on.

Whereas if we can construe all these different visual areas to have as their subject matter distal places, differently coordinatized, then no such regress threatens us. The sorts of transformations between areas are readily conceived to include transformations and elaborations of coordinate schemes, often folding in information from other maps as well. Such transformations are not trivial: on this line maps in V1 cannot guide motor behavior, not because they are about the retina, but rather because they are as yet incapable of representing what happens when (for example) the eyes or the head move. Yet all along the subject matter of all those different visual areas remains the same: those visible goings-on in those distal places of interest. What changes is how those places are coordinatized.

A simple example might help: think of "orientation tuning" of columns in V1. Cells in some cortical columns fire optimally for inputs oriented in a particular direction; cells in a neighboring column fire optimally to a slightly different orientation. But what are these orientations orientations of? Are they orientations of activated cells across the back of the eyeball? Or are they the (quite distinct) orientations of bars of light arrayed in front of the animal, providing it visual stimulation? The retinotopic team plunks for orientations across the back of the eyeball; the retinocentric team reads them as orientations of bars of light. My hunch is that most of the visual science literature is at least implicitly retinocentric, not retinotopic. (So argumentum ad populum favors retinocentrism!)

Admittedly, this is not so much an argument as a statement of the prejudices inclining many people, including myself, to endorse the "retinocentric" way of thinking about the different visual areas. We owe Matthen a debt for clearly identifying the two alternatives, forcefully arguing for one of them, and thereby transforming this previously implicit dispute into an explicit one. This is progress indeed!

IV.C. Cohen and cross-modal

Cohen has ingenious interpretations of many of the experiments that were taken to favor location-based models of the allocation of attention. But I find his treatment of the cross-modal cases less than compelling. Since a major motivation of *Theory of Sentience* was to understand cross-modal identifications--to give some account that applies to multiple modalities--these cases are worth discussing, at least briefly.

A sound can cue a subject so that visual stimuli in the vicinity are discriminated more quickly and accurately. Likewise for pairings in the other direction, and for pairings between touch and either vision or audition (see Driver and Spence 1998). Suppose we try, with Cohen, to explain this in a purely object-based way (Cohen 2004, section 5.1 [ms p. 17 f.]). The sound, he says, allows the subject to identify an "auditory object": the source of the sound. A telephone might be a sound source; when it rings "it is natural to say here that the subject binds attention to the sound source" (Cohen 2004, [section 5.1, ms p. 17]). The cross-modal case is then described as one in which sound draws a subject's attention to an auditory object, located in a particular region. Next, the sound stops: the auditory object "ceases to bear auditory features". Nevertheless the auditory object at that region persists, and the subject continues to attend to it in that interval. When a visual object is presented nearby, it can be discriminated more readily because "attention is moved between objects as a function of their propinquity" (Cohen 2004, [section 5.1, ms p. 18]).

I think the cross-modal case is more difficult than this; to use his cost-benefit analogy, there are some additional costs hidden here on the "object" side of the ledger, and a potent windfall on the "location" side. At the crux of the difficulties lies the notion of an auditory object that persists in a given region even though it no longer bears any auditory features. These currently inaudible auditory objects can continue to attract attention. What species of object is this?

One natural way for philosophers to understand "auditory object" would take it to mean something *essentially* auditory: something as represented by audition alone, or (to use an older language) a "proper object" of audition alone. It has only auditory features. The advantage of such an object is that audition could on its own identify such a thing. But it can't be what Cohen means, since under that interpretation an "auditory object" would cease to exist when it loses all auditory features. (You might remember that there *was* a sound there, but that is not enough to give us a persisting object.)

The other alternative is that an "auditory object" is just any ordinary object that happens sometimes to have auditory features. Something that sometimes makes noise, in other words. The telephone on the desk is an example; it sports qualities perceptible in multiple modalities. You can sometimes see its shape under your papers; the buttons feel sticky; occasionally it rings. Perhaps your office is neater than this. In any case the ambiguity just spotted also

applies to the phrase "source of the sound". In a purely auditory sense this applies to a location identified auditorily: the one from which the sounds appear to emanate. It can also mean the physical object that happens to cause a sound. Cohen uses this latter sense, and notes that the piece of metal and plastic on the desk can qualify.

What's the difficulty under this alternative? The problem is that the perception of this sort of object presumes successful cross-modal identification. The "auditory object" under this interpretation is not just the sound of the telephone: it is the telephone itself. What's the difference? It has to be non-auditory. Other modalities oblige: the telephone has shape, color, texture, etc. If a currently inaudible auditory object is perceived to persist, the perception thereof must be non-auditory; and to justify the claim that we are perceiving *that* auditory object, one must already presume the truth of some cross-modal identifications. That partially occluded shape is, for example, the shape of the very thing that sometimes rings. Unless we avail ourselves of one of those other modalities, I don't understand what might retain one's attention when one remains attentive to a currently inaudible auditory object.

The common sense notion of an object presumes that cross-modal identifications have already been successful. This is a problem if our goal is to explain how those identifications are possible. In contrast, the account proffered by feature-placing is at least non-circular. To contribute to an identification of the form "*a* = *b*", where the two terms are provided by different modalities, each modality must operate with some term that could be placed in an identity relation with some term available in other modalities. A ringing noise is not visible and a telephone shape is not audible, so those features and others like them-- the ones specific to one modality--are of no help in either confirming or disconfirming an identity. If both vision and audition are to contribute to the identification, we need something identifiable in vision that could be identical to something identifiable in audition. Location and other spatial relations are the classic candidates. Without them it is hard to see how we would ever get to the notion of one thing that has both a telephone shape and a propensity to emit those annoying noises. With them, these and all the other cross-modal identifications are relatively easy to explain. Spatial discrimination is a lingua franca that can unify the senses.

So I am afraid we must make a last-minute amendment to our previously filed accounting statement. An unbiased estimate of the hidden costs on the object-based side: think Parmalat. Deduct twelve billion euros to cover the costs of illicit purchases of auditory objects,

somesthetic objects, kinaesthetic objects, *e cosi via*. Whereas accountants on the location-based side have discovered some old stock options in the bottom of the closet that have matured a bit. Anyone heard of Microsoft, 1986, 1000 shares?

IV.D. Levine and qualitative similarity

Levine is quite right that similarity relations alone will not answer the question of how it is that brain states count as instantiating the relevant quality space. A causal model will be needed to answer that question; he is right that causal relations are not being entirely replaced by qualitative similarity in functional definitions. Nevertheless, the relation of qualitative similarity is the one that is basic to individuating the various qualities. As Levine notes, this hypothesis is threatened by different kinds of spectrum inversion than the sort that threatens standard functionalist accounts. If we need to include causal relations to explain how brain states count as instantiating the relevant quality space, the result gets even more complicated. Perhaps (as in Shoemaker 1975) we employ one set of resources (e.g. that causal model) to answer absent qualia, and a different set to address qualia inversion.

It's not quite true that the data on the basis of which the relations of qualitative similarity are derived consist entirely of *judgments* of similarity (as Levine suggests in his section 2 [7th para, ms p. 4]). A better word would be "discriminations"--as in all the psychophysical discrimination tasks one can conceive--or, even more broadly, anything that might yield data on relative similarities. The phenomenon of stimulus generalization can yield such data, even though it shows up in learning, not just discrimination tasks. I mention this because both Rey (with his "sensational sentences") and Matthen (with his account of how to make content explicit) seem also to endorse the notion that statements provide a touchstone for sensory content. Feature-placing is an alternative to the idea that sensory individuals are objects all the way down; it is likewise an alternative to the idea that sensory content is statements all the way down.

V. Conclusion

There is still more to be done, but here no more time or space in which to do it. I thank my commentators for all their work, and the editors for providing the occasion.

References

Blaser, Erik, Pylyshyn, Zenon W., & Holcombe, Alex 0. (2000). Tracking an object through feature space. *Nature 408*: 196-199.

Buser, Pierre and Imbert, Michel (1992). *Audition*. Tr. by R. H. Kay. Cambridge, Massachusetts: MIT Press.

Clark, Austen (forthcoming) "Perception Preattentive and Phenomenal", in Paul Thagard (ed.), *Handbook of Philosophy of Psychology and Cognitive Science,* in the multi-volume *Handbook of Philosophy of Science,* eds. Dov Gabbay, Paul Thagard, and John Woods, to be published by Elsevier.

Clark, Austen (2004). Feature-placing and proto-objects. *Philosophical Psychology,* this issue.

Clark, Austen (2000). *A Theory of Sentience*. Oxford: Oxford University Press.

Cohen, Jonathan (2004). Objects, places, and perception. *Philosophical Psychology,* this issue.

Davis, Greg & Driver, Jon (1998). Kanizsa subjective figures can act as occluding surfaces at parallel stages of visual search. *Journal of Experimental Psychology: Human Perception and Performance 24*(1): 169-184.

Davis, Greg and Driver, Jon (1994). Parallel detection of Kanizsa subjective figures in the human visual system. *Nature 371* (27 October) 791-793.

De Valois, Russell L. & De Valois, Karen K. (1993). A multi-stage color model. *Vision Research* 8: 1053-65.

De Valois, Russell. L. & De Valois, Karen K. (1975). Neural coding of color. In E. C. Carterette & M. P. Friedman (eds.). *Handbook of Perception Vol. 5: Seeing*. New York: Academic Press, 117-166.

Driver, Jon & Spence, Charles (1998). Cross-modal links in spatial attention. *Philosophical Transactions of the Royal Society of London B 353*: 1319-1331.

Heydt, R. von der & Peterhans, E. (1989). Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. *Journal of Neuroscience 9*: 1731-1748.

Kahneman, D., Treisman, A., and Gibbs, B. J. (1992). The reviewing of object files: object-specific integration of information. *Cognitive Psychology* 24 (2): 175-219.

Land, E. H. (1985). Recent advances in retinex theory. In D. Ottoson & S. Zeki, (eds.), *Central and Peripheral Mechanisms of Colour Vision.* London: MacMillan, 5-17.

Levine, Joseph (2004). Thoughts on sensory representation: A commentary on Austen Clark's *A Theory of Sentience. Philosophical Psychology,* this issue.

Locke, John (1975). *An Essay Concerning Human Understanding*. Edited by John Nidditch. Oxford: Clarendon Press.

Matthen, Mohan (2004). Features, places, and things: Reflections on Austen Clark's *Theory of Sentience*. *Philosophical Psychology,* this issue.

Moore, G. E. (1918). Some judgments of perception. *Proceedings of the Aristotelian Society x*: 1-29. Reprinted as Moore 1965.

Moore, G. E. (1965). Some judgments of perception. In Swartz, Robert J. (ed.) *Perceiving, Sensing, and Knowing*. Garden City, New York: Anchor Books, 1-28. Originally published as Moore 1918.

Moore, G. E. (1993). Proof of the external world. In Thomas Baldwin (ed). *G. E. Moore: Selected Writings*. London: Routledge, 147-170.

Nakayama, K. & Silverman, G. H. (1986a). Serial and parallel encoding of visual feature conjunctions. *Investigative Opthamology and Visual Science* 27 (Suppl 182).

Nakayama, K. & Silverman, G. H. (1986b). Serial and parallel processing of visual feature conjunctions. *Nature 320:* 264-265.

Peterhans, E. & Heydt, R. von der (1989). Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. *Journal of Neuroscience 9*: 1749-1763.

Rey, Georges (2004). A deflated intentionalist alternative to Clark's unexplanatory metaphysics. *Philosophical Psychology,* this issue.

Rey, Georges (forthcoming). *Representing Nothing: Intentional Inexistents in Cognitive Science*. Oxford University Press.

Shoemaker, Sydney. (1975). Functionalism and Qualia. *Philosophical Studies* 27: 291-315.

Strawson, Galen (1989). Red and 'red'. *Synthese 78*: 193-232.

Treisman, Anne (1993). The perception of features and objects. In *Attention: Selection, Awareness, and Control: A Tribute to Donald Broadbent*. Edited by A. Baddeley and L. Weiskrantz. Oxford: Clarendon Press, 5-35.

Treisman, Anne (1988). Features and objects: The fourteenth annual Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology* [A] *40*: 201-37.

Whittle, Paul (2002). Contrast colours. In D. Heyer and R. Mausfeld (eds), *Perception and the Physical World: Psychological and Philosophical Issues in Perception*. Chichester, England: Wiley.

Zeki, Semir (1993). *A Vision of the Brain*. Oxford: Blackwell.