Comments on Bill Lycan, "More Layers of Perceptual Content"
Austen Clark, University of Connecticut

Presented at "Naturalized Philosophy of Mind and Language: A Conference in Honor of Ruth Garrett Millikan", UConn, 4 October 2008

(Draft of 6 October; please do not quote without permission)

I'm very happy here to be sandwiched between Lycan and Millikan, two of the living philosophers from whom I've probably learned the most, and to whom I am the most grateful. Plus the intermediary position is appropriate for someone commenting on intermediary representations in vision.

There's much to like in Bill's account of "layering" in visual representation. For one, it makes explicit and publicizes the notion that there are multiple layers of representation involved even in the seemingly simple achievement of *seeing a tree* or *spotting a duck* (or a rabbit, as the case may be). Representations at one level might represent colored expanses, whose shapes are (perhaps) just two dimensional silhouettes; another (much informed by vision, if not within "vision alone") represents everyday three dimensional physical objects. So, as Lycan puts it, visual representata are not always "physical objects of the everyday sort" (see Lycan 1996, 152). Here, here.

The representations at different levels can engender differing ontological commitments. A possible visible world might satisfy the content of all the visual representations of colored expanses, but contain no physical objects of the everyday sort. Contrariwise, we might have--and according to Bill we do have--a visible world that satisfies the content of visual representations as of everyday physical objects, but that does not at the same time always satisfy the representations at the "colored expanse" level. In such cases the intentional objects of representations at the colored expanse level are represented, but non-actual. So, to use the original example from Peacocke: One sees two trees of the same size, one of which is further down the road. Hypothesis: One represents a bigger tree-shaped expanse, and a smaller tree-shaped expanse, and by representing those one represents two equally sized trees, one further away. There are trees of those sizes out there, but the tree-shaped expanses, with their differing sizes, are not actual (Lycan 1996, 157-8).

Vision scientists certainly do credit the visual system with a variety of representational schemes of differing orders of complexity, from early retinotopic maps of "discontinuities of intensity" up through some sort of visual representation of three dimensional objects or voluminous shapes. It is difficult to gainsay the claim that there are asymmetric priority relations between various of these. One represents the visible edges of objects by representing differences in luminous intensity along loci out there in the ambient optic array. Typically the priority is understood to be causal and computational: in order to come to represent the borders of objects one must start by representing discontinuities of intensity, determine which form

continuous loci, sort them into luminance or reflectance edges, then extract the overall structure. If one couldn't represent the initial discontinuities of intensity, then one could not come to represent the edges of objects. So we get an asymmetric priority relation between the contents at one level of visual representation and those at others.

But it turns out that the theses just stated are not sufficient to get us to *Bill Lycan's* view of layering, and that the differences are important. I'll start with the many things that are right and good about layering, how it is generally in accord with contemporary vision science, and how (I think) recent developments in that field can make it even better. But then I will focus on some of the particularities of Lycan's position, and how they might create some obstructions if the goal is to find an interpretation that maximizes agreement with vision science.

I.

The latter is not explicitly Bill's project, but I think it gives a refreshing and useful perspective on the issues. So I propose we doff the philosophy of mind hat and replace it for the remainder with a philosophy of psychology hat.

Something very interesting has been happening in the last fifteen or twenty years to the "theoretical constructs" describing these "intermediary representations" within vision. They have been getting smarter; more and more of the principles of perceptual organization and gestalt goodness that had been thought the exclusive province of central cognitive processing have been reassigned to these relatively early, preattentive, autonomous underlings. I think some of this recent work can substantively strengthen Lycan's layering thesis.

Consider in particular the work of Ken Nakayama and his collaborators on "surface representation" within vision. It was Nakayama and Silverman (1986a, b) who found one of the first counterexamples to Anne Treisman's theory that search is always slower for targets defined by conjunctions of features (red & square) than it is for uniquely instantiated features--eg if there is exactly one red thing in visual perimetry, it can be picked out very quickly. They found that if the targets were presented stereoscopically, so as to appear on different planes, or at different depths, then subjects could (in effect) confine their search to just those items appearing at a given depth. So even if there were multiple red squares on the display as a whole, as long as there was just one red item appearing on the depth plane to be searched, it would quite efficiently "pop out". Notice this is a conjunction of color + apparent depth. They found a similar effect with apparent motion. A display might present green things moving horizontally and vertically, and red things moving horizontally and vertically. If there was only one red item moving vertically, however, it would pop out. The suggestion again was that direction of motion defines an apparent surface--the display looks like two transparent surfaces sliding across one another. Subjects can confine their attention to just those items on one of the two apparent surfaces.

Anne Treisman modified her model, allowing both stereoscopic depth and direction of apparent motion to serve as two of several "surface defining" features (see Treisman 1993, 9). They provide parameters that can confine the scope within which selective attention must search. Since a vast number of items at a different apparent depth (or direction, etc) can all be ignored simultaneously, the representation as of an item at a given depth must be achieved preattentively, autonomously, and in parallel for many items at once. Yet it is a potent principle of perceptual organization; it defines limits within which attention itself can be confined. (This is what I mean by saying that these early bottom-up preattentive processes have been getting smarter and more organized.)

Since those early years Nakayama, He, and Shimojo (1995) have developed a systematic model of the intermediary "surface level" representation in vision. They claim it is essential to the eventual recognition of objects.

> Before the process of object recognition can begin, an object must be separated from the rest of the image and made available to the mechanisms of pattern recognition. ... we cannot think of object recognition as proceeding from image properties such as those captured by early cortical receptive fields; there needs to be an explicit parsing of the image into surfaces. (Nakayama, He, and Shimojo 1995, 14-15)

> Surface properties rather than image properties are decisive. It appears that all higher visual processes must have, as a data format, a surface representation. (Nakayama, He, and Shimojo 1995, 45)

To keep terminology straight I'll call the intentional objects of these representations "Nakayama surfaces" (as opposed to "Lycan expanses").

What are the essential properties of Nakayama surfaces? Nakayama, He, and Shimojo (1995) provide three rules--three commandments, if you will--that these entities must obey.

> 1. An occlusion edge between two surfaces is owned by only one of them.

One surface is seen as having that edge as a border; the other is not. A major task of early vision is to determine which surfaces own which borders.

> 2. If the surfaces are opaque, the one that owns the border is in front of the one that does not.

and

> 3. A surface region that does not own the border is unbounded, and with others can form a larger surface "completing behind".

Those various regions are seen *as* (represented to be) parts of one surface, occluded *by* the other surface.

It's simplest to illustrate these principles with a figure (see figure 1). Even though they aren't recognizable objects, one cannot help but see *x* as a surface in front of the rest. In part because of the T junctions, x "owns" the border between itself and y: that border is part of the boundary of x, not of y. So x seems in front of y. The same holds for the border with z. Both y and z are

unbounded surfaces which "complete behind": they are seen as two portions of one surface, which is occluded by x. The "occluded" portion of the y-z surface has perceptual reality, even though there is no stimulus for it. Its "completion" is hence "amodal".

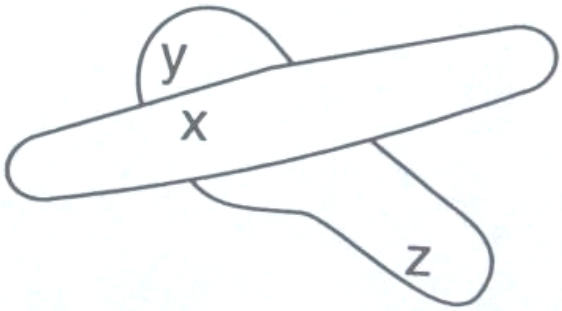


Fig. 1. Region x owns the borders and lies "in front"; y and z "complete behind".

Completion behind distinguishes Nakayama surfaces from Marr's 2.5d sketch, as well as from the famous two dimensional color patches floating serenely across the visual field. The system is representing occluded portions of a given surface *as* portions of that surface. It is imposing a certain minimal organization (or "segmentation") on image data, which unlike the representation of the image includes representation of some currently invisible portions. The surfaces also stand to one another in some depth-defining relations: they can be "in front" or behind, and they can occlude or be occluded. So more than one surface can be found in the same visible direction: one in front, which occludes; and others behind it, which are invisible but not inexistent. (It is therefore more than the 2.5 d sketch can do. But as in the 2.5 d sketch these surfaces also have a "surface normal" which gives their orientations.)

So how do surfaces help Lycan's layering thesis? I think they can provide a respectable home for close relatives of his "colored shapes" and "tree-shaped expanses"--entities which otherwise might seem unsavory, disreputable, misbegotten, good for nothing. It turns out they are respectable citizens, but of a different species.

First, the model vindicates the idea that we are representing entities like "plushy red expanse at 2 o'clock" in a robust way which nevertheless is *not* yet representing it to be an everyday physical object. It is not an everyday physical object because it not represented as filling a volume; the spatial relations between surfaces are much sparser. Surfaces can be opaque, but vision represents nothing about what is inside them; they lack stuffing, or any kind of solidity besides the visual sort. Lycan suggests that his expanses and shapes are physical in the sense that they are not immaterial, and they seem to be located in space, yet nevertheless they are not physical objects in the everyday sense. Nakayama surfaces have all these features too.

Nakayama et al (1995) argue in detail that representation of surfaces proceeds independently of and prior to the representation of objects. It "reflects a more basic, autonomously driven mechanism that is relatively free from top-down, object-level knowledge" (Nakayama, He, and Shimojo 1995, 9). Furthermore, they argue that in the perception of "impossible objects" and perhaps in certain kinds of agnosia, the two sorts can be dissociated: the representation as of surfaces is intact, but perception as of objects is disrupted or destroyed.

Second, I think Nakayama surfaces can help with all the new examples Bill considers in his paper. Specifically, the surface level provides a simple account of a representational content that changes in the various examples of ambiguous figures, attentional grouping, aspect shift, and so on. To take these in turn.

The face/vase figure was first published in 1921 by Edgar Rubin, who in fact used it to state an early version of the "border ownership" idea. Only one surface can own the border between them, and in the ambiguous figures, ownership is up for grabs. It changes when perception shifts between faces and vase. Which surface is in front, and which is background, changes as well. As Rubin noted, this works even with asymmetric, unrecognizable, novel, and abstract shapes. First the concave black shape appears in front, appearing to occlude a white background. Then the convex white shape owns the border, appears in front, and appears to occlude a black background.

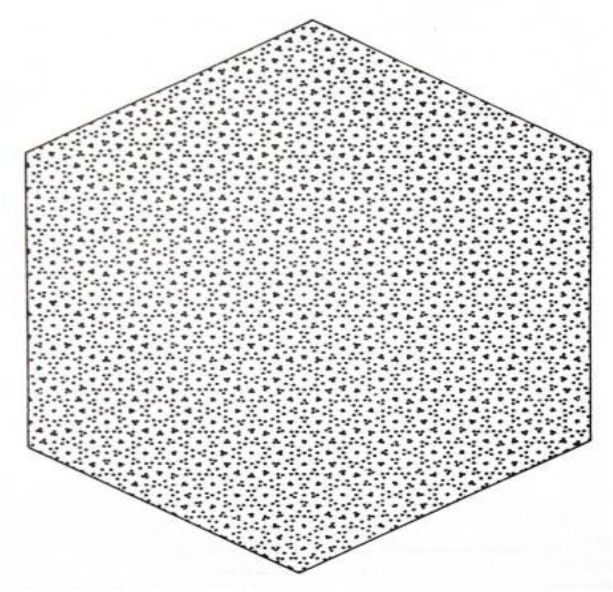
In the Necker cube and similar examples, we see multiple simultaneous changes in the appearances of ownership, occlusion, and front/back relations. Which surface seems on top, which on the ground, and which closest to the observer all change; the apparent orientations of the surfaces relative to the ground also change. Interestingly, the point from which one seems to be looking at the Necker cube moves ninety degrees, from looking at it from the side to looking at it from above.

It is useful to think of representation of surfaces as one way to organize or segment image-level data. The layout includes portions that are currently invisible; like a hypothesis it has implications beyond the current data set (here, image data). Also like a hypothesis, it has to be generated, actively. The processes that generate it are often called "grouping processes" or "gestalt processes". Bill calls them "attentional phenomena". (This isn't necessarily selective attention, but what vision scientists call "visual spatial attention"--the sort that manages binding and spatial grouping.)

As he notes, one of the responses to Nickel's (2007) "tiling" example cites some version of figure-ground. The border-ownership issue makes this explicit. The various groupings and patterns of squares seen as "prominent" are the ones seen as in front. They own the relevant borders. The rest of the figure is seen as a background, which completes behind.

An especially vivid example of "active grouping processes" is found in Marr (1982, 50) who found it in turn in a MIT master's thesis by J.L. Marroquin. See figure 2. Peacocke's "dot diagram" generates similar visual

phenomena. Both diagrams can be organized in many different ways; many different groupings satisfy all the principles of gestalt goodness. Which portions of the diagram are seen as a organized "figure", in front, and which form the background completing behind, change from moment to moment. As Marr notes, the diagram "apparently seethes with activity as the rival organizations seem to complete with one another" (Marr 1982, 50).



Finally, in many of the examples of aspect-shift, one thing that shifts is the apparent organization, and orientations of, the constituent surfaces. The "rat man" and E. G. Boring's "old woman/young woman" show this vividly. Even the duck-rabbit has some apparent surfaces whose orientation relative to the observer appears to change with the aspect-shift. The duck isn't looking in the same direction as the rabbit; the two bills of the duck are in plane and equidistant, while one ear of the rabbit is slightly behind the other; and so on.

But, as Bill also notes, other examples of aspect change can't be handled in the same way. Even in the duck-rabbit the surface-level changes seem paltry; this sort of aspect-shift might also require *recognized* part-whole relations among the parts of a recognized object. Likewise, seeing a biscuit as a ship of the line is quite beyond the capabilities of the surface representation system.

II.

Now on to some problems, or at least issues, with the specific form of layered representationalism that Lycan proposes for vision. They fall under two main headings: (a) the visual representation of non-actual objects, and (b) two different kinds of layering: Lycan's "same-token" kind--that one and the same representation has multiple contents--and a kind I will call "sequential" layering. I'll address these in turn.

Non-actual objects, intentional objects that do not exist, make their

appearance in two prominent places in Bill's development of the layering thesis. The first is at the origin point: how to respond to Christopher Peacocke's example of becoming aware of two tree-shaped expanses, one larger than the other, when one sees two trees of the same size, one further away. The other is the green pointy-shaped after-image, about which Bertie cannot tell a lie (see Lycan 1996, 71). Bill answers with a thorough-going and bracing representationalism. If F is a "sensory quality"--one of the distinctive introspectible qualitative features of sensory experience--and one is aware of the visual appearance of an object x bearing feature F, then one is visually representing there to be an object x having property F. In both size constancy and afterimages, there is no actual object x having property F--no larger and smaller tree shaped expanses, no green pointy thing before the eyes--but there is a layer of visual representation that has those contents. Indeed, according to Bill in size constancy it is by falsely representing these non-existent expanses that one comes veridically to represent the actual sizes and distances of the everyday physical objects known as trees.

Now size constancy is in play in practically every visual scene. As long as the subject can see different objects at different depths, a divergence can arise between the size of the visually apparent "object shaped expanses" and the size of the objects themselves. In every such case, if the subject achieves a veridical perception of the object's size, it is on this account mediated by a non-veridical representation of differently sized expanses. In perceiving them we are perceiving individuals that exist in other possible worlds not our own. Of course no cost accrues for non-actual objects per se, but the representing of them does take time and energy, neurons and electric potential, glucose and oxygen. A believer in biosemantics will wonder about the pay-back. Someone who wants to hold some form of direct realism will likewise balk at the idea. Ruth takes a hard line on non-actual objects:

> Teleological theories have in common that they deny that there *is* any object at all that is being represented when one is thinking emptily (say, when seeming to think about "phlogiston" or "the ether") or that there *is* any state of affairs or occurrence being represented when one thinks falsely. Similarly, there *is* no object, not even an inner one, being seen when one has an hallucination. Teleological theories all deny this for the same reason. They take it that mistaken representations, rather than representing peculiar objects, things called "contents", are merely representations that are *failing* to represent. (Millikan 2004, 64)

There is though a way to avoid many if not all of the non-actual objects in Bill's account of size constancy: move from Lycan expanses to Nakayama surfaces. In particular, the fact that Nakayama surfaces are represented as residing at different depths yields a simple way to resolve the conflict between apparent sizes. Size is already a representandum, and once depth is represented as well, surfaces in a scene so represented are thereby also represented as occupying larger or smaller visual solid angles (angles defined from the eyeball's point of view, of that surface at that depth.) In seeing the closer member of a pair of equally sized objects, one is also seeing the member that

occupies a larger visual solid angle. Perhaps the "larger expanse" is just the one that fills a larger visual angle. The person who trains himself or herself to become aware of such angles is aware of an actual property of the scene.

(Nakayama surface level representation does not *explicitly* represent visual angles, and it is worth noting that its explicit representation of depth is somewhat sparse: "occluding" *v.* "is occluded by", and "in front of" *v.* "behind". Nevertheless, in representing a surface of a given size at a given depth, one is representing a surrface that fills a given visual solid angle. Any visually represented scene that satisfies the former must also satisfy the latter.)

With this reinterpretation the statements about larger and smaller expanses come out mostly true. We don't need non-actual objects. Surfaces at varying depths will do as well. This is also nice for the biosemanticist, for most of the surfaces represented at this level turn out to be the surfaces of actual physical objects, standing more or less where they are represented to be.

In working through this I discovered a few passages in which Bill countenances the possibility that his "shapes" or "tree shaped expanses" *are* represented as potentially occluding one another, or as standing at varying distances from the observer. For example

> We do visually represent the trees, and represent them as being of the same size, etc., but we do this *by* representing colored shapes and relations between them. Some of the shapes--in particular those corresponding to the trees--are represented as being larger shapes than others, **as occluding others**, and so forth. As with all intentional objects, it does not follow that there are any actual things that have such relational properties. (Lycan 1996 151-52, emphasis added)

So perhaps Lycan expanses are in this respect not far away from Nakayama surfaces. Once we add occlusion, or varying depth, the talk of "bigger expanses" can mostly be mapped to the truth value True. It need not be interpreted as referring to entities in a possible world other than our own.

III

Bill's other prime example of non-actual objects is the pointy green after-image.

> Suppose Bertie experiences a pointy green after-image. There is a green spot in his visual field; if he denies it, he is a liar. (Lycan 1996, 71)

Actually I have long had difficulty with this example. It makes me want to lie. More precisely, I don't think having a green afterimage is an episode of representing something to be green.

Recall Sellars' distinction, reiterated by the disjunctivists, between three cases:

> (i) seeing an x to be green,
> (ii) seeing an x that seems to be green, and
> (iii) seeming to see an x that is green.

Arguably afterimages belong in category (iii), along with visual hallucinations, dreams, and other mental episodes that are not perceptual, but may

experientially resemble perceptual episodes. I'd like to suggest a new theoretically possible niche for the description of such episodes. An afterimage of the kind caused by photographers' flashbulbs results when a swath of retinal receptors is overloaded by too many photons; almost all the photopigments in the receptor are isomerized, and for a while that receptor is out of commission. (This is not true of every kind of afterimage; some, such as the McCulloch after-effect, don't even have a retinal genesis.) While it lasts you cannot see *anything* in the region of the afterimage; no new inputs can be registered. If the function of this "producer mechanism" is to produce a token that maps onto the world--a token upon which some consumer can rely--then here it is *failing* to do its job. Like a sensor that has gone haywire, the outputs do not track anything. So an episode of having an afterimage is an episode of visual malfunction--brief and reversible to be sure, but one that renders the optic apparatus more or less entirely insensitive to new inputs in that region of the retina. But as the receptors recover you have an experience as if you were seeing something. You seem to see something that is green.

I suggest that having an afterimage, like having a dream or other episodes in category (iii), is *not* an instance of visual representation. Experientially it might *seem* like seeing something that is green, but that fails to show that one *is* seeing something green. The internally produced token might *seem* to be representing something to be green, but that fails to show that it is representing something to be green.

Suppose you have a computer program that monitors fan speeds in your laptop. On the screen there is an array of pixels representing a numeral representing the fan speed. Most of the time it is accurate. But the program has a bug in it, and sometimes after a particular sequence of system events x, y, z the fan speed reads 65535, then 0, then 65535 ... etc. One alternative: the program is falsely representing the fan speed to be 65535. Another: the screen shows a numeral, and in that sense the program seems to be representing something. However, the software isn't very good, we have buffer-overflow, and now random values are being pushed into the memory location storing the fan speed. 2^16-1 is a bad number to see in a system monitor. It means the monitor has gone haywire. The program is malfunctioning. It seems to be representing the fan speed to be 65535, but it is not. It is not representing anything.

(How so? The value is not being produced in a Normal way. *No* semantic mapping function can be relied upon for guidance by a consumer. The producer is malfunctioning, even though it sends something to the consumer, and the consumer is functioning properly. The consumer responds as if to a token of a representation, though that token is not representing anything.)

Bill hews to a crisply defined form of representationalism: that all "sensory qualities"--all of the distinctive introspectible qualities of visual experience--are *represented* features of visual intentional objects. As he puts it:

> Sensory qualities are only intentional contents or objects of mental states, represented properties of representata... (Lycan forthcoming, ms. p 7)

Now any creature that has representations is also endowed with a representational system that produces them. That system has its own internal states and operating parameters; like any biological product it is prey to a range of infirmities and malfunctions. Some of its states are tokenings of representations, which qua "vehicle" have their own properties, distinct from the properties represented. All such states and parameters come for free: we can't have an instance of representing without having all that other stuff as well.

So here's the idea: perhaps some of the similarities and differences in the introspectible qualities of perceptual experiences can be explained by appeal to some of these *other* variables; to properties besides *represented* properties. I credit this idea to our own Franklin Scott (forthcoming a, b). Scott argues that some of the similarities and differences among the appearances of the colors can be explained by appeal to properties of the *vehicles* that in us and other primates happen to represent the colors (e.g., that they happen to be neurally opponent processes). I think of this as Franklin Scott's "vehicular" hypothesis.

Similarly, perhaps some of the experiential features of these episodes are to be explained by appeal to those other variables--to properties other than the represented ones. If so, then we can admit that Bertie has an experience as of seeing something green, while denying that Bertie is visually representing something to be green. So we can tell the story without having to tell a lie.

(It is still true that there are no differences in the introspectible sensory qualities of visual experiences of kinds (i), (ii), or (iii) without differences in what one might call the "state of the visual representation system". So in a sense we still agree with the maxim "No sensory-quality difference without a representational difference!" (Lycan forthcoming, ms. p 7.) But it differs from Bill's strong form of representationalism in that not all the differences in the state of representational systems are differences in *represented* properties. In this kind of afterimage we have a malfunctioning system, which (I suggest) is not visually representing anything. Nevertheless Lycan's form of representationalism can remain fully in force for episodes in categories (i) and (ii): seeing something to be green, and seeing something that *seems* to be green.)

The conclusion is close to what Ruth says about "seeing pink elephants" and other examples of visual hallucination:

> When delirious, you say that you "see pink elephants" even though you are surely not succeeding in seeing pink elephants, because there are not elephants there to see. ... Seeming to see is confused with actually seeing... (Millikan 2004, 65).

I don't think that having an afterimage is hallucinating, but it is an instance of "seeming to see" something, belonging in category (iii). It experientially resembles seeing, but fails visually to represent anything.

IV.

Now on to (b): the distinction between kinds of layering, "same token" versus "sequential". Bill's view is *not* adequately characterized by the claim that there are different levels of representation within the visual system. It does *not*

suffice to add that the representations are of different kinds or carry different ontological commitments. What is distinctive about his view is that *one and the same representation* has multiple layers of content. The same token or icon within the system has multiple and non-equivalent truth conditions. I've suppressed discussion of this feature until now, mostly because it is harder to reconcile with contemporary vision science.

Consider the perception of an ambiguous figure or an aspect-shift. As Bill puts it

> There is here a Wittgensteinian mystery about visual content: The attending is a visual phenomenon, or *at least* it affects visual phenomenology. But in some sense what is seen does not change. (Lycan forthcoming, ms. p 8)

One sees that the lines and vertices printed on the page do not change. But how the figure looks does change.

Here allow me to register an interpretive assumption that might be incorrect. Bill suggests, but does not overtly state, that layering applies to the perception of ambiguous figures and some instances of aspect-shift. That is: (a) there is some visual representation of that stimulus which itself has multiple layers of content; (b) on one of the layers the stimulus is seen not to change; and (iii) on another layer of content of the same token, the stimulus appears to change. If Nakayama surfaces help this account, they help only with the layer that is presumed to represent different contents at different times. I've suggested that the changing appearances all involve changes in ownership of edges and occlusion relations that a Nakayama surface representation makes explicit. But if my interpretive assumption is correct, then for the *same* visual representation there is a lower level of content according to which the stimulus in front of the eyes is represented as unchanging. As Bill puts it re the Necker cube:

> ...the two cube experiences share some shapes, edges, and lines; and, according to me, all those items are visually represented. So there is after all a representational sameness underlying the aspectually different seeing-as experiences... (Lycan 1996, 155-156).

Here the "shapes, edges, and lines" would indeed be representata of something other than everyday physical objects. But more importantly they would also be representata distinct from Nakayama surfaces. Since this lower level just represents the unchanging lines and vertices, it is closer to something like the "primal sketch", representing just image-level data, perhaps augmented with descriptors. But then, according to Bill's "same-token" layering, the same representational token has a content representing image-level data, and it has a content representing surface level features.

The problem: That's a big jump to make between layers of content of one and the same representation. It is a big jump *computationally.*

For example, one way to make Lycan's layering consistent with Ruth's consumer semantics is to suppose that there is more than one consumer mechanism for one and the same representation. Those consumers can have different functions, and the normal conditions for successful performance of

those functions might require different semantic mapping functions for the sign. So the very same sign has different contents. But if this is the way the ambiguous figure cases work, then we have to imagine one consumer relying on a normal mapping function defined in terms of image-level data, and a second consumer of the same token relying on a normal mapping function defined over surface-level features. For the second consumer to be able to do that, I would wager that it would have to perform computations on the original token of such complexity that it would forced to form its own intermediary representations. Furthermore, to add to the wager, those representations would be representations of Nakayama surfaces. The computational problems Marr described have to be solved somewhere, and this just pushes the ones found "from images to surfaces" into the innards of one of the consumers.

(One way to make this problem go away is to deny my interpretive assumption: that in the perception of ambiguous figures there exists a visual representation with multiple layers of content, one layer of which reflects the unchanging aspects of the perception and another layer of which includes the changing aspects. Or one might deny the consumer semantics approach to the constitution of distinct layers of content.)

What is the alternative? Vision science is quite happy with layers, but typically the layers are distinct representations. At least within early vision we find a bucket brigade of producers and consumers, where the consumer of one sign has as its function the production of another sign, which in turn is consumed, and so on. (At some point of course the buckets have to travel outside of vision per se, and serve some more potent biofunctions, if the whole business is to be worthwhile at all.) I will call this a "sequential" layering view. In it is still true that for some X and Y, a subject perceives Y by perceiving X; there are "asymmetric priority relations" among the distinct contents. But the priority relation is just causal or computational: in the sequence a necessary condition for constructing a representation of the content Y is first to construct a representation of the content X.

There are at least potential empirical differences between the two kinds of layering. In the sequential view, one could not have a double dissociation between the layers. Certainly one could retain the first layer in a sequence but lose the second. So for example a person can retain the representation of image level data but lose the capacity to represent the organization of surfaces, surface shapes, and so on. But on the sequential view, if one loses the first layer, one must also lose the second. If the subject cannot represent image details then surface representations will be lost as well. Whereas on the same-token view, if the consumers are distinct, there is no a priori reason to deny the possibility that a percipient might lose exactly one of the two, and that it might be either one. So there could be subjects that represent image details but not surfaces, and there could be subjects that represent surfaces but not image-level details. I don't think the latter has ever been observed, and perhaps its psychological implausibility is what inclines vision scientists to sequential layering rather than the synchronic kind.

Could Ruth hold a layering view of the sequential variety? Like Bill, I find no passage that commits her to it, though there are passages that are consistent with the possibility. For example, Millikan (2004) describes visual processing as a sequence of natural signs that are also intentional representations, in which one might represent first a collection of edges, then a shape, then identify a particular 3d individual. Language processing is

> a process in which intentional representations...are formed, passing through a number of layered stages of intentional representation in the process of translating public language signs into inner representations of world affairs. Similarly, neurological evidence suggests that ordinary visual perception involves the translation of gradients of luminance across the retina as signs of various rudiments of visual forms, such as lines or edges with a particular orientation ... which are then interpreted as signs of such constancies as shape... (Millikan 2004, 116-117)

She describes the "transitions from sign to sign that take place during perceptual processing" (2004, 119) and calls them "translation". "Consider the move from seeing the shape that is like that of Johnny's face to mentally representing the presence of Johnny." (2004, 119). Much of this could be combined with the sequential layering view.

The problem though is straightforward: the asymmetrical priority relations that obtain between distinct representations in a sequential layering view are just causal or computational. Perhaps that's not enough to sustain what Bill would consider an instance of perceiving Y *by* perceiving X. Ruth acknowledges the causal and computational priorities, but goes on to argue that even representations late in the "series of signs" still qualify for the moniker "direct perception". While the sense of "direct" employed in those passages does not rule out all kinds of asymmetrical priority relation, it does exclude some traditional exemplars--perceiving a tree by being acquainted with a color patch in the visual field, for example. For Ruth the intentional objects of the mediating perceptions are resolutely distal. Although she never says it herself, it would be consistent with those texts to say that one perceives the tree by perceiving its Nakayama surface, or one perceives the presence of Johnny by perceiving the shape of Johnny's face (see Millikan 2004, 119). But then Bill might be left unhappy: this sense of "perceiving by" yields at best a very weak notion of mediated perception. With that I must stop: let me step out of this intermediary position and let them duke it out directly.

References

He, Z. J. and Nakayama, K. (1992). Surfaces versus features in visual search. *Nature 359:* 231-233.

He, Z. J. and Nakayama, K. (1994). Surface shape not features determines apparent motion correspondence. *Vision Research 34*: 2125-2136.

Lycan, William G. (1996). *Consciousness and Experience*. Cambridge, MA: MIT Press.

Lycan, William G. (1998). In defense of the representational theory of qualia (Replies to Neander, Rey and Tye). In J. Tomberlin (ed.), *Philosophical Perspectives, Vol. 12: Language, Mind and Ontology.* Atascadero: Ridgeview Publishing, 479-487.

Lycan, Willian G. (forthcoming). More layers of perceptual content. Presented at Naturalized Philosophy of Mind and Language: A conference in Honor of Ruth Garrett Millikan. University of Connecticut, 3-4 October 2008.

Marr, David. (1982) *Vision.* San Francisco: W. H. Freeman and Company.

Millikan, Ruth (2000). *On Clear and Confused Ideas: An Essay About Substance Concepts*. Cambridge: Cambridge University Press.

Millikan, Ruth (2004). *Varieties of Meaning*. Cambridge, MA: MIT Press.

Nakayama, K. (1990). The iconic bottleneck and the tenuous link between early visual processing and perception. In C. Blakemore (ed) *Vision: Coding and Efficiency.* Cambridge: Cambridge University Pres.

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.

Nakayama, K., Shimojo, S., and Ramachandran, V. S. (1990). Transparency: relation to depth, subjective contour, and color spreading. *Perception 19*: 497-513.

Nakayama, Ken, Zijiang, J. He, & Shimojo, Shinsuke (1995). Visual surface representation: A critical link between lower-level and higher-level vision. In Stephen M. Kosslyn & Daniel N. Osherson (eds). *Visual Cognition*. (*Invitation to Cognitive Science*, 2nd. ed., vol. 2. General editor Daniel N. Osherson.) Cambridge MA: MIT Press, 1-70.

Nickel, B. (2007). Against intentionalism. *Philosophical Studies* 136: 279-304.

Scott, Franklin (forthcoming a). More than meets the eye? Submitted to *Philosophical Studies.*

Scott, Franklin (forthcoming b). A Theory of Color Perception and Semantics. Ph.D. Dissertation, Department of Philosophy, University of Connecticut.

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.