8 A Question of Content

KATHLEEN AKINS

PART I

1. Introduction

The topic of this essay is, or is supposed to be, the influence of Daniel Dennett's philosophy of mind on neuroscience. Since I do not have a straightforward answer to the question, "What has Dennett contributed to the neurosciences?," I am going to start off with a *different* question, one that I stand some chance of answering, and then with that answer in hand, I will circle back around to the original question in Part II. In other words, I am going to take (what I hope will be) the scenic route. (Got it?)

The question with which I will begin, then, the 'different question,' is one that almost every philosophy of mind student has asked since the publication of *Consciousness Explained (CE)*. Why does Dennett think that our folk attributions of perceptual states, attributions made from the intentional stance, do not merely *describe* the propositional contents of the subject's conscious perceptions? *CE* makes clear that, for Dennett, there are neural representations 'in there' (after all, Dennett is not *that* kind of behaviourist) and that the contents of these representations somehow determine how we consciously experience the world. You see the world as you do because certain visual 'specialists' within the visual system have determined that the world is a certain way. But if so, it would seem to follow that at least, we must know how the world seems. If so, then surely *we* are capable of making accurate (self-) attributions of our own perceptual beliefs. So why does Dennett think that not even we, as conscious subjects, get it right? This is an extremely good question. Before taking it up, however, let me make the put the question in philosophical context.

2. A Cartesian Stalking-horse

Like Ryle before him, the principal motivation of Dennett's view of mind is a negative thesis about what the mind *could not possibly be*, namely Ryle's

Ghost in the Machine or Dennett's Cartesian theatre. On the strongest version of this view, no doubt a view that not even Descartes espoused, the mind is portrayed as like a stage on which the self-luminous denizens of consciousness parade for an audience of one, the ever present, all-seeing Self. Famously, on this view, the mind has immediate, infallible, and incorrigible access to the passing mental stream. First, the contents of the mind are, as Ryle would have put it, 'phosphorescent.' Insofar as one has a conscious mental state, one knows one has it, without inference, attention or any other quasi-perceptual process of 'illumination.' There are no phenomenal events – could not be any phenomenal events – which dance in the dark unknown and unappreciated by the Theatre's audience. For conscious events, to be is to be perceived.

Second, because the mind's access to the stage is unmediated there are no quasi-perceptual processes that might 'go wrong.' A conscious event cannot be misperceived or unrecognized for what it is. If I see a tree before me, then I know that I do so, and moreover, I know that I seem to see *a tree* (although, of course, I may be wrong about the state of the external world, whether or not there is a tree before me). In the same way that the mind could not fail to notice a conscious event as it passed by, it cannot fail to apprehend the phenomenal events as they really are. Our access to conscious mental events is infallible.

Third, because there is but one Self to a theatre, a person's access to the conscious contents of his or her mind is unique. My access to the events of your mind is mediated necessarily by both our bodies. I can only infer, on the basis of what you say – or more generally what you do – which conscious states lie within and the same is true for your knowledge of me. Hence, no one can know better than the Self the nature of its own conscious experience. Self-knowledge, in this limited sense, is incorrigible.

All of the above, I hope, is a familiar portrait of the 'Cartesian' view or at least familiar enough from its detractors' descriptions. One more piece needs to be added, however. Because Dennett's positive view of consciousness is largely a theory about *perceptual* consciousness – even more specifically, about visual consciousness – something needs to be said about the role of perceptual states in the Cartesian theatre of mind. Unfortunately, insofar as I know, Dennett has not given a sustained description of the 'Cartesian' view of perception that he is targeting. The target arises piecemeal in Dennett's writing. Substance dualism aside, though, Descartes' view of perception (as opposed to a Cartesian view) has many of the features of contemporary materialist theories of perception against which Dennett is reacting. So, for our present expository purposes (and with apologies to Descartes for the historical vilification) it should do.

On Descartes' view of perception, neural events in the brain 'occasion' phenomenal sensory events, sensations that in and of themselves tell the Self almost nothing about the states of the world which caused them. That is, *some* of these states may be pleasant or unpleasant, a 'content' that informs the Self that the cause of the sensation is either harmful to or good for body. Take the case of a scratched toe. Damage to the body results in a phenomenal state, that of pain, a sensation that is intrinsically unpleasant. Hence, it warns the Self not of a particular kind of bodily damage per se, say, puncture, nor of damage to a particular part of the body, say the toe, nor of *damage to the surface* of the body. An unpleasant sensation serves only to warn the Self *that* the cause is, in some unspecified way, harmful. The same holds true for some noxious tastes and odours, unpleasant temperature sensations, and so on. To put this another way, apart from the contents of pleasantness or unpleasantness, the qualitative properties of sensations, *what they are like*, exhausts their 'content,' what they tell you about the world. By comparing sensations, one with another, you can come to know *something* about their causes in virtue of their qualitative similarities. For example, given three sensations, a 'blue' sensation, a 'yellow' sensation, and a 'red' sensation, by comparing these sensations to each other and to the multitude of other sensory sensations before you, you could come to realize, that these sensations are of the same type ('visual' or 'colour' sensations). You could also infer that the *causes* of the yellow sensation and red sensation are more similar to each other than to the cause of the blue sensation because yellow is more similar to red than to blue qua phenomenal property. But all of this is inferred knowledge, made on the basis of qualitative similarity. More generally, on Descartes' theory of perception, on the basis of its innate understanding of the general properties of the physical world and the particular facts of the sensory array before it, the mind infers the presence of the properties and events that might inhabit the world beyond it. The mind must *infer*, for example, that it is my toe that seems to be hurt, that my right leg seems to be crossed over the left, that there seems to be a wooden table before me, and so on. And then, with these inferential conclusions before it, the mind *judges* whether or not these states of affairs actually obtain, whether the content of the inferential conclusions is true *of the world*. So judged, we come to have conscious perceptions of the properties and events of the world external to the mind.

One feature of Descartes' view of perception is worth special note in this context, namely the role of 'pure sensations' – or rather the two distinct roles that *qualia* are made to play. As I said above, sensations are the phenomenal *inputs* to the mind, the raw materials of (perceptual) cognition. They are the purely qualitative (and sometime evaluative) states, *qualia*, which the physical senses deliver up to the mind for the scrutiny of its rational capacities – for the mind to categorize, make inferences about and from, and perhaps to like or dislike according to the idiosyncratic nature of the subject. (Some likes and dislikes are merely a matter of personal preference. For example, I may love broccoli and George Bush Sr. may hate it, but the 'awfulness' of broccoli-for-Bush is not a property it has in virtue of God's goodness.) That said, our sensory perceptions of the world, of tables and chairs and the bodies we possess, have purely qualitative aspects to them as well. For example, while the mind can infer from the sensory array, the size, shape, and position of objects within the visual field, our sensations of colour tell us nothing about the particular surface properties of objects. In Descartes' terms, God has constructed our sensory processes such that we can tell from certain sensations only *that* a particular property of the world has caused them, not *what* the particular cause is. To infer that the daffodils in the vase *are* yellow is to mistake a phenomenal property for a physical one. Strictly speaking, daffodils are *not* yellow (because yellow is a phenomenal property), even though the fact that they appear so does indicate something about the surface properties of daffodils, we know not what. Thus, our fully formed, intentional perceptions also have qualitative 'aspects,' the second role of *qualia* within the theory.

In Descartes' theory of perception, then, *qualia* play a dual role: They are both the *inputs* to higher-level perceptual processes and the qualitative aspects of the *output* of perceptual processes, of intentional perceptions per se. Dennett would argue, I expect, that the *qualia* play a dual role in our contemporary understanding of perception. Certainly we speak of, say, the taste of Marmite qua a part of the sensory experience of eating bread and Marmite, a part that we can isolate and 'examine' as 'the exact taste' of Marmite. In the same breath, however, we will speak of the memory of one's very first taste of Marmite, and how *that* taste affects or compares with the experience of Marmite we have at present. We speak as if what is stored is a purely qualitative sensation that can later be used, say to assess one's current perception of Marmite on toast. (Is today's Marmite on toast, for example, up to the standards of that very first bite?) So the tensions implicit in Descartes' 'antiquated' theory can plausibly be said to continue on in contemporary theory.

This, at least in rough outline, is the Cartesian stalking-horse.

3. The First Move

As I understand it, Dennett's philosophical project is this. Dennett wants to develop a positive, stable theory of the mind by steering clear of the reefs and shoals of the Cartesian view, while incorporating the 'landings' supplied by the cognitive sciences within his charted route. To put this another way, he sees quite clearly where most of the Cartesian hazards lie, what he has to avoid. He has also some (but fewer) positive landmarks to visit, 'stops' from neurophysiology, social psychology, and on to evolutionary theories about the brain that he should include. Still, those constraints leave a lot of open water. What is far less clear is which route, of the indefinite number of possible routes the landmarks leave open, he should take. This is why I suspect (and please excuse both the metaphor and the pun) so many people have found Dennett's theory of mind a 'moving target.' It seems that way because it *is* that way. From Dennett's perspective, there is nothing wrong with deviating from one's stated route given a good reason – that is, an unexpected (Cartesian) hazard or an unforeseen need for a (empirical) pit stop. After all, just how else would a person navigate? ("Officer, I ran into the pedestrian because he was there and because I had not planned to avoid him.")

Famously, both Ryle and Dennett make the same crucial – and irrevocable – first move. The problem, as both Dennett and Ryle understand it, is that psychological states, the garden variety beliefs, desires, and emotions that we attribute to ourselves and others, are not (as Ryle so often said) either conscious *or* unconscious states. Strictly speaking, psychological 'states' are not states in the mind/brain at all: They are behavioural dispositions (Ryle 1949) or "patterns of behaviour" (Dennett 1991b). We, as practitioners of folk psychology, may well suppose that our attributions of propositional attitudes serve to pick out or refer to particular mental/neural events, tokens of the appropriate type, in a subject's mental/neural economy. But if we do so, we are in the grip of a systematic misunderstanding of both what we do (Dennett) and what we say (Ryle) through such attributions. Rather, according to Dennett, both folk psychology and its more sophisticated cousin, the Intentional Stance, are "best viewed as a rationalistic calculus of interpretation and prediction – an idealizing, abstract, instrumentalistic interpretation method that has evolved because it works and works because we have evolved" (Dennett 1987b). We begin our predictions by looking at the organism's behaviour and ascribing to it the beliefs it ought to have given its particular needs for information and its perceptual history. We ascribe to it the desires it *ought to have* – its primary needs being those for procreation and survival, its secondary needs being whatever conditions are required to survive and procreate. Finally, we deduce what the organism *ought to do* (or ought to have done) to satisfy its desires relative to those beliefs given the constraints of rationality. Thus we use intentional attributions to predict and "retrospect" (Ryle 1949, pp. 166–7) the actions of intentional systems without recourse to any knowledge of the (astoundingly) complex causal processes within the mind/brain. This is *in fact* how we make intentional attributions and the reason why we do so – and we should not suppose in light of this predictive success that we have gained any great insight into the concrete workings, the neural processes, that result in human action.

Let me stop for a moment to highlight just how radical this first move is. Unlike tables or chairs or the average kumquat, we – you and I and most of the rest of the species homo sapiens – *are* conscious entities, beings with full and rich phenomenal lives. This is not something that either Dennett or Ryle denies. What is radical about the first move, then, is not that the theory portrays us as zombies, poor old automatons without two conscious states to rub together. What is radical about the move is how little it leaves us vis-à-vis what we thought we understood about ourselves and about our perceptual relations to the world. To put this another way, what the Cartesian view of mind makes possible is a very neat and orderly epistemic portrayal of what goes on 'both within and without.' One begins with the realist assumption of a 'set' ontology, of the world with all of its objects, events, and properties neatly delineated, at the ready for us the fair perceivers. Now our perceptual systems given to us by God (or evolution) allow us to infer our way from sensory input to what lies beyond the mind/brain. If everything is working correctly (and as Descartes often reminds us, it will work correctly only if we too are working hard), we will have reliable information about how the world lies. So, for example, if you are standing looking at a plate of kumquats the sensory input from your eyes will eventuate in a conscious perception of that fruit, perhaps with a content <Lo! Kumquats!>. With the perception of kumquats before you, you can now think about those kumquats. You can wonder how much the kumquats cost, who bought them, or what they really are (i.e., are they really just very tiny oranges?). Moreover, given the "phosphorescent" nature of all conscious events, you know what you think and you can *express* your perceptual belief to others should you be so moved. You as the speaker of the English language can express it with the sentence, "Lo! Kumquats!" Of course, if a thought can be *expressed* by means of an utterance, then it can also be *described* in the same way. Using the appropriate sentence, I can refer to or pick out your perceptual thought by means of a belief attribution with the appropriate propositional content. All of which, as I said above, yields a very orderly view: (a) our perceptions reflect the properties of world; (b) our utterances express the propositional contents of our conscious perceptions; and (c) our attributions of psychological states refer to those perceptual states by means of their propositional contents. The mind may be necessarily private on the Cartesian view, but fortunately through the grace of God (or evolution), we gain a window on the world and the minds of others via perception and language.

The classic behaviourist move attempts to annihilate this tidy picture. Our intentional practices give us no reason to think that our attributions pick out, by means of their content, specific conscious (or unconscious) states. Nor do we have any reason to think that our utterances are *caused by* these conscious states, and hence that our utterances express their contents (see Brook, this volume). Nor can we adopt, without critical examination, the Cartesian picture of perception. According to Dennett, when we attribute beliefs to an intentional system, including perceptual beliefs, we do so according to what beliefs the system *ought* to have, relative to its informational needs and perceptual history. When a fly buzzes by a frog, the frog ought to believe that a fly, as opposed to a lead pellet, is present. But to attribute a belief about a *fly* leaves open a rather wide range of possibilities of how – or even *if* – the property of being a fly is represented by the frog's brain/visual system. The intentional stance works just insofar as the frog behaves *as if* it sees the fly. The same point holds, *mutatis mutandi*, for the attribution of *any* intentional state to any intentional system. Given a 'correct' attribution of an intentional state (i.e., one that works) is an entirely open question 'what goes on within,' whether the neural capacities of the species are genuinely representational (whatever that might mean) or merely 'contentful' or 'informational' in some other sense. Moreover, it is an entirely open question what kind of causal relations the organism's conscious states bear to the world beyond it, to its perceptual capacities in general, and to its other behaviours. To put this another way, there is no reason to think, on the basis of an attribution of a perceptual state, that we are referring to or picking out particular conscious states, the 'end products' of a process that begins when we open our eyes and stare out at the world – that is, that we refer to an event of this kind when we say, for example, "She sees a plate of kumquats." Seen this way, the behaviourist strategy is a truly radical first move.

We are now almost back to the question with which this paper began, namely, "Why does Dennett think that the contents of conscious perceptual states will not 'line up' with the content of the attributions of perceptual beliefs that we make from the Intentional Stance?" Note that the behaviourists' 'first move' does *not* give us the strong negative thesis that Dennett wants: It does not prove that the two kinds of contents will not align. What he gains from the first move is really only an open question, the very possibility of nonalignment as it were: 'Will they or won't they?' To show that the two kinds of content will not align, a different sort of argument needs to be made.[1] In particular, for a strong argument, what Dennett needs is a positive view about the nature of neural content, the nature of perceptual processes as well as a theory about the relation between conscious perceptual states and perceptual processing. Of course, it is these views that are going to be primarily of interest to the neuroscience community, and so I will turn to them now.

4. Neural Content

Where Ryle and Dennett part company is over the issue of 'what goes on inside?' Ryle was disinclined to sanction anything that looked vaguely Cartesian, of course, including all talk of representations, conscious perceptions, and inferential processes (although, touchingly, Ryle often admits that he does not know what to say, *really*, about perceptual phenomenology). In stark contrast, Dennett begins with the assumption that 'subpersonal' neural states can also be said to have 'content', but of course not the (bad) propositional content of folk beliefs and desires. He says in "Styles of Mental Representation" (1987d), that subpersonal psychology "will be 'cognitive' in that it will describe processes of information-transformation among content-laden items – mental representations – but their style will not look or behave like sentences manipulated in a language of thought" (p. 235).

As a subpersonal psychologist, Dennett is a 'pure laine' computationalist: Brain-states are computational states. To determine the content of a subpersonal functional/computational state one must engage in a process of holistic interpretation: One must look at the role of the state in the overall cognitive economy of the organism, its contribution to overall function. Says Dennett: "It is only the globally defined role of such a state (the role that is characterized in terms of rules of operation the whole system 'follows' when it goes into that state) that fixes its information or external semantic properties" (1987d, p. 224). Most important, brains, like any computational devise are mere *syntactic engines*. Says Dennett:

> the brain, as physiology or plain commonsense shows us, is just a syntactic engine; all it can do is discriminate its inputs by their structural, temporal and physical features and let its entirely mechanical activities be governed by these 'syntactic' features of its inputs. That's all brains *can do*.... If you want to get semantics out of syntax... in the end all one can hope to produce (all natural selection can have produced) are systems that *seem* to discriminate meanings by actually discriminating things (tokens of no doubt wildly disjunctive types) that co-vary reliably with meanings. (1987e, p. 61)

In virtue of evolution, certain computational states have come to function within the computational system as if they had meanings of a certain kind. Hence, most of the time (insofar as both evolution and the organism's environment have settled on a mutual solution for the moment) we will be able to assign neural content by appeal to function. But this is all there is to a neural state's having meaning to or for a system. At bottom, there are no facts of the matter about neural contents in exactly the same way that there are no facts of the matter about the semantics of linguistic utterances (à la Quine) or facts of the matter about the intentional states of whole persons (à la Dennett). There are only better or worse holistic *interpretations*. Neural content, *like all content*, is mere *as if* content (see 1978b, 1978c).

Most important, we should not mistake our privileged positions as external interpreters of function for "the organism's being in a position to know or to recognize or to intuit or to introspect that fact from the inside" (1987f, p. 304). To put this another way, says Dennett: "Brains... do not assign content to their own events in the way observers might: brains fix the content of their internal events in the act of reacting as they do" (1987e, p. 63).

5. Visual Processing: Shoals and Rip Tides

In 1978e, Dennett presented a thought experiment of the following kind. You wake up to find yourself locked in a windowless room, all alone but for a desk, the chair on which you sit, and, covering one large wall, something that looks for the all the world like a control panel (think here of your first experience of the cockpit of a 747). Lo and behold, it *is* a control panel. A note, left on the desk, tells you that you are trapped in the control room of a giant robot, on whose activities your life now depends. 'Your' robot inhabits a world of physical events and objects, some nasty and some nice, or so you are told, and your job is to 'steer' the robot safely through it all (whatever 'all' is) using the control panel in front of you. There are thousands of 'input' lights and 'output' switches on the control panel. The problem is that they have *no labels*.

Dennett argues from this thought experiment that if an unlabelled control panel were your *only* access to the 'your' environment and motor capacities, you would never learn to perform the Cartesian deduction (or inferential) task. Says Dennett (private correspondence) "you don't have the leverage without the labels... even if you have 'OUCH' and 'YUM,'" that is, some kind of evaluative labels affixed through the beneficence of God or whomever has so kindly placed you there. Were this the predicament of the human neonate, neither motor control nor perceptual understanding would ever get off the ground. Dennett goes on, "The only way you could survive is if the system is to some degree pre-labelled or, more realistically, some of the input-output linkages are already made for you." Here Dennett voices a familiar enough philosophical complaint: The inputs to perception could not be purely qualitative states.

As might be guessed, Dennett does not believe in *qualia* qua the *outputs* or products of perceptual processes either. There are no *qualia* served up by the senses on the stage of consciousness, to be liked or disliked, judged, examined, stored in memory, or used for other cognitive purposes by the Self. However it may seem to you from the first-person point of view, there is no qualitative state '*the* taste of my first bite of Marmite.' There is no taste that you immediately dislike, that you judge to be suis generis in your gustatory experience, that you can examine at leisure for its various qualitative parts ("what, exactly, *was* that?") or that you can store in memory in anticipation of future Marmite encounters. More subtly, by Dennett's lights, our intentional perceptual experiences do not have 'qualitative aspects' at all. Because the inputs to perceptual processes are not phenomenal states (à la the Empiricists), there are – can not be – phenomenal 'survivors' of the perceptual process (there being none to begin with). Hence, there are no qualitative states in the phenomenology of my intentional perceptual experience which, through a suitable act of concentration, I can access by abstraction. Explaining this will take a little work.

Let me begin with a classic misunderstanding about vision, one that inevitably arises in undergraduate classes, in either the neurophysiology of vision or computational approaches to vision. Take Marr's (1982) computational theory of vision. Recall that according to Marr's theory there are four levels of representation in image processing. First, an image represents light intensity at points in the retinal image. Second, the Primal Sketch makes explicit intensity changes and their spatial organization. Third, the 2.5-D sketch represents edges, depth, and orientation information. Fourth, the 3-D Model represents shapes and their spatial organization in terms of stick figures and volumetric primitives (e.g., generalized cylinders). As soon as one begins to explain the Primal Sketch, however, a very large number of introductory students will look deeply puzzled. "Why do we have do *anything* to the digital photograph?" some students immediately ask, "Why do we have to make the contrast information *explicit* when it is already clearly visible?"

Other students will find Marr's methodology more or less comprehensible, at least until the very last step. As a model for scene segmentation or discerning depth information, that is, Marr's theory may look reasonably plausible. Where most students balk is later, at the suggestion that something like Marr's model is – or even could possibly be – a good model for human visual perception. That is, as a model of what our visual systems eventually do – namely produce conscious visual phenomenology – Marr's representational scheme seems strange at best. When you look out the window to the forest before you, you do not *see* little dots and arrows affixed to 'cartoon' outlines of the trees (the primitives of the 2.5-D sketch) nor do you *see* stick figures or generalized cylinders (the primitives of shape recognition). Our visual experience insofar as it is 'like' anything within the model seems more 'like' the photographic image, a kind of 'pixel-by-pixel' depiction of the scene. Yet, in Marr's model, pixel-by-pixel intensity information is eliminated by the second stage of processing, the Primal Sketch. Starting with the photographic image, the more levels of representation completed – each one serving to make explicit more information – the less 'like' our visual phenomenology the representation seems to become. So how could a visual perception be produced by *that*? In both cases – for the student who refuses to leave behind the digital image and the one who is unwilling to take the final step – a number of confusions are at play, each of which will emerge in the explication of Dennett's positive theses below.

Dennett's central claim about visual perception is just this: *It is NOT the function of perceptual systems to produce our conscious perceptual experiences.* For example, given the visual experience you are having right now, as you read from this printed page, it may *seem* to you that *this* – your visual phenomenology of the page in all its complexity – is just what vision is for. With this image 'in mind,' you can *see* the letters on the page, see the structure of each individual letter. Perhaps you will use that information to identify the font or typeface. You can also use it to *read* and *comprehend* the text. You can form beliefs about what typeface has been used, which words appear on the page and what the text says. And, of course, in virtue of seeing the page, you can flip it over and lay it face down on your desk. On the basis of your conscious perception, you can 'act with respect to' the page. That is, as the end product of (no doubt) many complex unconscious processes, a conscious visual perception can be used for an indefinite number of cognitive and motor activities – to turn one's attention toward a salient part of the visual landscape, to visually guide one's actions, or to form inferences about the nature of the visual scene. Or so it seems. Dennett's claim is that this common understanding of vision is wrong. It is merely an entrenched *illusion* engendered by your first-person point of view. As such, it embodies any number of confusions.

First of all, it is a mistake to think that visual processes produce a single static representation, one that the mind's eye then uses to discern, as needed, the relevant details of the visual world. Just as one serially examines, say, an architect's plans for a building, looking first at the site placement, then at the general layout, followed by the dimensions of individual rooms, placement of stairways and so on, the mind's eye scans the static visual image for relevant information. This is the first and most obvious mistake of the introductory student, the student who refuses to move past the digital/retinal image. The student mistakes the visual phenomenology that arises out of looking at a static digital image for that which is presented to the mind's eye: a static phenomenal 'visual field.' Hence, there is no need to *extract* information from the digital image, thinks the student, because the digital/retinal image itself just *is* the kind of static 'presentation' required by the mind's eye. Of course, you cannot interpret an architect's plans without some prior knowledge of the conventions of drafting, but with that in hand, you can infer what you need to know. In this, the student shares much with Descartes. The mind's eye must come equipped with some kind of knowledge about the nature of the external world in order to interpret the static phenomenal array.[2] But you do not need, first, to transpose each element of the plan into a different format. Why would one *re-represent* that which is already represented?

Second, what makes the visual image so tempting qua phenomenological presentation to the mind's eye is that the image is, well, an *image*. Because what we see when we look at a photograph of a kumquat is somehow similar to what we see when we look at the kumquat itself, this tempts the unwary into the belief that vision, and hence the representations that support vision, must somehow be 'imagistic,' just as photographs are. Of course, explicitly, very few people would assent to the most naïve view of visual processing or neural representation in general. No one thinks that the only way to represent an oval-shaped, orange kumquat is with an oval-shaped, orange kumquat – or even with an oval-shaped, orange *plastic* kumquat (Akins 1996). And of course, given the very limited means that brains have to encode information, by frequency encoding, spatial encoding, and population encoding of various kinds, visual representations are unlikely to be 'picture-like' in any straightforward sense.

Still, if you start from the other end, if you think about neural representations beginning with what we *see* – our visual world in all its complexity – as opposed to concentrating on the representational relations between neural states and what they represent, the 'naïve' view is far more compelling. That is, if you think of vision as producing a static phenomenal image, then you need not be terribly naïve to think that what underlies that experience is, well, a static *neural* image (at least in some sense of the phrase). Certainly this is one reason why the student who willingly follows Marr through each level of representation may still balk at accepting the whole as a model of human vision. That is, most students understand immediately that information about the visual world must be extracted from the intensity information of the digital image before it can be used. They start to worry, however, about what the whole process actually buys us. Why is it – *how could it be?* – that anything like Marr's model supports visual experience? For all that Marr's 3-D model may help us to *identify* the shape of, say, a horse, it seems hard to imagine that a 3-D model of a horse has anything to do with my present conscious perception of the horse before me – that is, my perception of a noncylindrical genuinely horsey horse. The underlying structure must be *something like* the perception of the horse I actually have. But unfortunately, this way of thinking takes one right back to where few of us have wanted to go – namely to a naïve theory of perception, the view that properties of our visual phenomenology can result from only 'like' properties of the representational vehicles. This is the second mistake of the introductory student, the naïve view of perception.

Dennett's response to this intuition is distinctly 'hardline.' Indeed it is Berkeley's line *tout court*. The brain does not represent states of the world with neural 'pigment' nor are phenomenal states composed of qualitative 'figment.' For example, an orange kumquat need not be represented in the brain by an orange neural state, and seeing an orange kumquat is not a case of having a kumquat appear 'orange-ishly,' of having in mind some 'orange figment.' As Berkeley was at pains to teach us (for rather different reasons!), phenomenal properties are not 'like' physical properties if by 'like' one means 'qualitatively similar.' The very idea is incoherent.[3] Only an idea could be 'like' or qualitatively similar to another idea, said Berkeley (and Descartes agreed). Add, now, the further proviso that 'qualitative similarity' does not involve *qualia*, and Dennett would also agree. Entirely.

Finally, there is a third intuition at play in the students' reluctance – the intuition that visual perceptions qua static images are 'complete.' I will not pretend that I can give anything but a very rough sense of what this means, but perhaps it will suffice to say something like this. A perception must be 'complete' in the sense that it should contain all of the visual detail that the mind's eye comes to infer or notice or act on. That is, insofar as the static image is the *input* to the mind's eye, the fuel for the inferential engine as it were, then it must be capable of sustaining our actual cognitive/motor activities. The image must contain enough 'visual detail' within itself to guide action, or sustain inferences about the identities of visually present objects, or determine the typeface on a page, and so on. So, if you see the horse's glossy chestnut coat, that information about the texture and colour of the horse's coat must be *there for the taking*, as Dennett would say. This intuition as well as the one above, that genuine perceptions must be underwritten by imagistic vehicles, is what confounds the more sophisticated student. Although he or she realizes that the digital/retinal image could not guide activity in and of itself, what makes for conscious experience is yet another question. And Marr's 'sketchy' 3-D model, while both static and imagistic (in some sense), looks like a rather poor possibility for this role. Hence the sophisticated student's reluctance to take the final step.

So, the function of vision is not to produce a conscious state, a static, 'complete' visual image. But if not that, then what? Ah, Dennett would say, but that is the crux of the matter. *There is no 'what,' no single product of vision, much less a single conscious product.* What we must take on board – *really* take on board – is that for any species with a lens and retina, the *only* image in vision as a whole is the retinal image, an ever-changing complex mosaic of light. And like any proximal stimulus, visual or otherwise, it is of no use to the organism in and of itself. Insofar as evolution manages to exploit the various properties of a retinal image – its changing patterns of light wavelength and intensity – within the behavioural economy of a particular species, can it *become* useful. To put this another way, we have to remember that the march of evolution is a march *away* from the retinal image. If the behavioural/cognitive economy of a species is large and diverse, so, too, will be the ways in which the visual system exploits the retinal image. But we make a (large) mistake when we equate this fact with a march toward a *reconstruction* of the image or toward a representational format suitable for producing 'figment.'

6. Visual Processing: First Sightings

So what really happens in vision? Insofar as I can discern, once the negative debris is cleared away, Dennett's theory of vision has very few premises. What the human visual system is attempting to do – like any sensory system at all – is to make use of the properties of the proximal stimulus, the retinal image. It does so through using a series of 'visual specialists,' parallel, multi-track processes, each of which serves to 'fix a content,' 'make a discrimination,' or 'perform a micro-taking.' Starting with the receptors themselves (which act as the first filters of the retinal image), down through the three types of neurons of the human retina and ending with the ganglion cells, the retina itself serves as the first visual 'specialist.' It does not 'render' the retinal image into a digital code, with a description of the wavelength and intensity of light at each point on the retina, as the metaphor of the image suggests. Rather, it sends the brain an intricate mosaic of luminance and wavelength contrast information, across a wide range of spatial resolutions. Then, in series of cascading events, second-tier cortical and subcortical 'specialists' (those downstream from the ganglion cells) make discriminations using the ganglion cell signals. In turn, these second-tier discriminations are passed along to other specialists, which make further discriminations, and pass along their fixations to yet other specialists ... and so on in a cascade of content fixations. (Note that by using the term 'cascade,' I do not mean to imply that the effects of micro-taking only go in one direction, i.e., away from the sensory periphery. The effects can flow upstream and downstream, as well as laterally.)

As Dennett will admit, exactly what happens next – what happens to the outputs of the visual specialists – is a very good question. In simple visual systems, such as in the visual system of the frog, one class of retinal ganglion cells responds to small dark spots moving across the retina. It is the spatial and temporal pattern of this population of 'fly' detectors that signals the fly's direction of movement across the frog's retina. Together, the ganglion cells encode, in egocentric space, spatial properties of the distal world, and it is this population response that initiates and guides the appropriate behavioural response: The tongue swipes toward the location of the fly encoded in egocentric tongue-swiping space. For many visual systems, the loop between input and output is really this short. One can virtually read off from the pattern discriminations made by the ganglion

cells, the motor uses of those signals. Of course, even a frog tongue-swiper might be a bit more complicated than this, depending on how 'fail-safe' the system actually is. For example, it may be that the frog can correct his tongue motion in mid-swipe in reaction to a change in the flight trajectory of the fly – a move that would require a translation of fly flying co-ordinates into tongue position co-ordinates and vice versa. But such 'tricks' still make for a reasonably simple input-output loop – a short hop from 'discriminations' to behaviour.

The human visual system is not a 'short loop' system of course. We, unlike the frogs, see a stable world of objects, properties, and events, and our visual systems must support this perceived ontology. We use our representations of the world to predict future events, reminisce about the past, draw complex inferences, or just muse on the passing scene. As yet, we do not know how this occurs. What Dennett is betting is that two positive empirical theses will turn out to be true. First, all visual systems, our own and the frog's included, will be shown to have the following property in common: All have evolved for the visual guidance of behaviour. Whatever 'discriminations' are made by the human visual system, that is, they must still subserve our repertoire of behaviours, both motor and cognitive. And however vast and complex that repertoire might seem to be, like the frog, it is still tightly constrained, both in the present, by our bodies and the environment, and historically, through evolution.

To put this another way, the type of specialists on board any visual system – the kind of 'micro-takings' made – is the result of a complicated, historical process, driven at each step by a process of mutual constraint satisfaction. Each specialist is the product of the co-evolution of the species' behavioural repertoire/motor systems and its sensory systems taken together, and a particular (co-evolving) environment. For example, flies being what they are, the pattern of light produced as a fly goes by, is a small convex moving edge – a pattern of light that also makes possible the computation of the fly's position and trajectory. But if during the evolution of frogs, the most abundant of flies had large, slowly oscillating wings, the ganglion cells of the frog might now respond to a different pattern, say a small flicker of movement at a certain frequency. Now, while a discrimination of visual flutter might have worked well to identify or pinpoint the fly, it might not have lent itself very easily to 'on the fly' (sorry) predation. In that event, a very different kind of motion computation might have come into existence, either as a separate class of ganglion cells or further downstream in the frog's (limited) visual system – and so on throughout the remainder (limited) of the frog's cascade of vision specialists.

This same lesson, believes Dennett, continues to hold true of our own visual system, only more so given what profoundly visual creatures we are. Even though we now represent a world of stable and changing objects, properties, and events, our visual system evolved to guide behaviour and continues to do so. Hence, it is likely that the cascade of human 'visual specialists' and the 'micro-takings' they make, continues to reflect its history of evolution. What the neurosciences will show is that our army of visual specialists has not been co-opted into the production of static 'master' representations, the kind of 'complete,' 'detailed,' 'unified' images of the visual scene that our first-person experience makes plausible. Rather, we, as a species, have found a way to represent a stable, coherent world of objects and properties and yet our visual system has nonetheless retained its army of specialists.

Second, neuroscience will show that *primate vision works on a need-to-know basis, to a large extent constructing its representations as required.* In primate vision, the structure of the retina represents a necessary informational trade-off, between spatial resolution and high sensitivity. Without the spatial resolution, we could not see the fine-grained spatial features of objects, such as object texture, recognize faces or make fine-grained judgments about object location. Without the high luminance sensitivity, we would loose the ability to see under less than ideal light conditions, to see objects in the light of dawn and dusk, in shadows and so on. Hence the organization of the primate retina, into a central, very small area of densely packed, 'high resolution' cells – the fovea – and the large surrounding retinal area, which, through cell convergence, sums the spatial results of many different cells. In this way, the parafoveal region increases its sensitivity to light but loses spatial resolution. As a result of this evolutionary compromise, primates make eye movements. We (or, for the most part, our visual system) rapidly move the fovea from point to point in visual space, foveating first on one small area of the visual world (3° of visual angle or the width of your thumb held at arm's length), then on another. To do object recognition or texture resolution, to see fine-grained depth or motion information the eye must fixate the appropriate portion of the world with the centre of each eye's visual field. Hence, as a result of this arrangement, the visual system must construct much of its understanding of the visual scene serially, from a sequence of very small (relative to the visual field as whole) high-resolution encodings.

What we do know about the eye movements of primates is that our eyes do not 'scan' the scene in the same way a computer refreshes a screen, for example, by starting at the top left corner, back and forth, back and forth, back and forth, until the entire screen has been covered. Rather, our eyes move in specific patterns of saccadic movement, directed by the particular features of the visual scene itself. For example, when looking at a human face, we make eye movements largely in the triangular region between the eyes, the nose and the mouth, while giving scant attention to the jaw-line, hair, and so on. Our eyes are 'entrained' by objects/properties of interest. Perhaps then primate vision does not even try to construct a representation of all the events and properties before it. That is, given that for the most part, the visual world is stable and unchanging, there is no reason why the visual system must store within itself all of the detail of the external world. It can operate on a 'piecemeal' basis, moving from point to point according to current interest – according to what specific visual specialists want to know, or what seems worth exploring on first glance, or what other neural specialists might be asking. All the system would need is some very rough information about the general lay of the land, and some cue(s) that inform the system if anything within the visual field moves/changes. If so, then in a blink of the eye (actually, much faster than a blink of the eye), it can turn its attention where needed.

In light of these two empirical assumptions, the central question Dennett tries to answer is this: How is the illusion sustained? Why do we seem to have a 'complete,' 'unified,' static 'image' of the scene before us?

First and most centrally, says Dennett, we should not forget *that vision looks out not in*. What we (humans) see is *the world as represented* – with the properties we represent it as having. Insofar as the phenomenology of vision gives us any insight into the processes of vision, it is not through the examination of the phenomenological properties of visual experience per se. We do not inspect *qualia* or the qualitative aspects of visual experience (for there are none). Rather we can only do what we always do: We can look at *the world* and say how the world seems to us under various visual conditions. We can say, for example, whether this illuminated matte surface looks the same colour as the sample directly beside it. We can say whether an illusory contour appears in front of or behind the point of fixation. In other words, we can do no more than use the standard methods of psychophysics to test what we can and cannot see.

Second, given the actual nature of our visual experience, the human visual system must be able to individuate events, order them in time, assign a coherent set of properties to a single object and understand an object's/event's/property's coherent and orderly progress through space and time. To do so surely requires an enormous integration of information about objects qua objects, events qua events, and properties qua properties. But this does not mean that the representational vehicles that make possible this kind of perception must be 'like' what they represent – the vehicles qua vehicles need not be physically unified, or make progress through neural space and time in the appropriately ordered fashioned, and so on. To believe as much is to cling to the naïve theory. Insofar as we perceive orderly events, coherent objects, or stable properties, there must be an explanation of how our perceptions come to have that content. But being aware of this kind of content is consistent with any number of representational formats or systems.

Third, insofar as the specialists of vision 'fix' external properties of the world via properties of the retinal image, their states do not discriminate particulars per se (even though, through clever contrivance, such discriminations may allow the visual system to track or follow individuals). That is, even at the lowest levels of visual processing, the system both encodes and makes inferences about the general properties of the visual scene, which it can then use to direct visual attention. For example, a specialist might signal "there is a large-ish dark something about *there*," "there was a small movement in peripheral vision," or "something of high spectral contrast is over there," or even "there goes a familiar but not identifiable biological motion," and so on. If low level vision does carry general information, it follows that the visual system can represent the properties of objects *as determinate, without representing those objects as having any particular determinate properties*. Take the example above, of a visual state that carried the information "there is something of high spectral contrast over there." Here, it need not be the case that the precise hue or saturation of the object has been determined anywhere within the colour system. Perhaps *all* that exists is the information that there is something of a bright determinate colour or other. Alternatively, the system might represent the object as being a very bright, determinate shade of red, without computing, anywhere in the system, exactly what shade of red that might be. What the visual system can safely assume about the world, in other words, is that 'everything is what it is and not another thing.' The world (at least on the space/time scale on human vision) is fully determinate. So whatever generalizations the visual system fastens on, the system can rest assured that those generalizations are grounded in determinate properties of the world. One has only to look.

The illusion of 'taking in the world at a glance' – of seeming to see, as fully determinate, the properties of the visual scene before you – is maintained, for one, because we simply see the determinate properties of the world when we look for them. A brightly coloured object in the undergrowth catches my attention and now I see it as a discarded shirt, bundled

into a particular shape. I pick up the newspaper and glance at the front page and now clearly see the front photo, a portrait of an ex-president who never came to like broccoli. For the most part, when particular properties of the visual scene are required, they are retrieved by a series of unconscious saccadic eye movements, *movements made, for the most part, without any conscious recognition that the requisite information has been missing*. Moreover, if the content of visual information is abstract, then, as I said above, we need not represent the particular properties of the world in order to represent the world as containing determinate properties. I see the shirt as bright and as determinately coloured, but not as red; alternatively I see the shirt as a bright determinate red, but without actually seeing what determinate shade of red that might be. Again, if the question arises, "But I wonder if that is the deep scarlet shirt that blew off my clothesline?," it can be quickly answered – by foveating the shirt and calculating its colour. The abstract nature of visual content leaves open the possibility that, at first glance, the shirt was not seen as *red* at all. As Dennett puts it (ever so nicely):

> It seemed to him as if . . . his mind – his visual field – were filled with intricate details of gold-green buds and wiggling branches, but although this is how it seemed, this was an illusion. No such 'plenum' ever came into his mind; the plenum remained out in the world where it didn't have to be *represented*, but could just *be*. When we marvel in those moments of heightened self-consciousness, at the glorious richness of our conscious experience, the richness we marvel at is actually the richness of the world outside, in all its ravishing detail. It does not 'enter' our conscious minds, but is simply available. (1991a, pp. 407–8)

7. Answering The 'Different' Question

We are now almost back to a point where we can address the students' question with which this essay began. So far I have outlined Dennett's theory of neural content in general, as well as his views about both the nature of mammalian visual processing and the kinds of neural contents involved in those processes. I have said very little about Dennett's theory of consciousness per se. Indeed, the only thing I have done was to hint, implicitly, at one of Dennett's central tenets about consciousness, namely that it is the contents of neural states that inform consciousness, not the physical properties of their representational vehicles. Still, even when made explicit, that statement is not very helpful. What is Dennett's view about how consciousness comes to be? What is Dennett's view about why we have the conscious visual events we do and not some others – that is, why is some neural content conscious? Again, Dennett's answer is dead simple. Information becomes conscious when it is shared (in some as-yet-unspecified way) with other neural specialists, specialists from other sensory modalities, motor specialists, 'cognitive' specialists – indeed, whatever kinds of neural specialists there turn out to be. Perhaps one might put it this way: In the very act of information becoming available, conscious content comes into being. It is therefore incorrect to say that consciousness is what makes possible the use of certain information – as if a large spotlight picks out, shines on, a previously dark representation in the brain, illuminating its content for all the other specialists to see and hence send for. Consciousness does not have a function in this sense. Dennett has also stressed (1991a) that at least in principle, any kind of neural contents could enter into our stream of consciousness. That is, because Dennett does not want to posit a conceptual/perceptual divide, or a distinction between 'conceptual' and 'subconceptual' neural content, there can be no divide, along these lines, between contents that can and cannot become part of the stream of consciousness. All neural contents are potentially conscious, at least insofar as they can be shared in the appropriate way. Exactly what more Dennett wants to say about consciousness is not clear to me but fortunately, for our purposes, even these few words suffice.

We can now provide for Dennett an answer – or two answers – to the question with which we began, namely: Why does Dennett think that the contents of conscious experience will not map nicely onto the contents of our intentional ascriptions? Why do the perceptual beliefs made from the Intentional Stance not describe the contents of our perceptual phenomenology? After all, it would not be surprising if other people were not able to describe the contents of your conscious experiences very accurately. But among the attributions of intentional states are self-attributions. Surely I know what I see! For one, as I said above, our visual system sustains the illusion that, at any given moment, we have within us a representation of the visual world in all its detail. This illusion is made possible by three things: the directionality of vision (we look out not in); the general and abstract nature of visual contents; and the visual system's ability to answer questions (posed by specialists) usually before we realize that the questions have been asked. If the Cartesian picture of the mind were true, that is, you would be able to examine the phenomenal contents of visual experience, and the illusion could not be sustained. If you wanted to know something about the contents of your visual perceptions, for example, "Do I really see the determinate colour of that shirt before me?," the answer would be available by introspection. You could just 'see' whether the shirt was represented as 'some bright colour of high contrast with its background' or whether the shirt was represented as 'bright fuschia.' But given that vision looks out, you cannot examine your phenomenology itself (whatever that would be) only the properties of the external world as they are represented. Aha! Then why not just look out at the shirt? The problem is this. Whenever you ask the question, "What kind of visual detail is being represented?," you will see just the detail you thought you might. You will see the fuschia shirt as fuschia. The visual detail is produced on demand. So looking out at the world is not of any use, at least not if you want to know what it was you just saw a moment ago. Finally, you cannot even notice in passing that this or that complex detail was available to you or missing – or not with any certainty. You may be fooled, in retrospect, by the general and/or abstract contents of perception. Certainly you saw *something* – more specifically, something to do with the colour of the shirt – but what? You will never be certain whether you really saw the shirt's determinate colour a moment ago, or whether you just now moved your eyes in order to determine its colour. For the most part, in the ordinary run of events, we do not have a very good handle on the contents of our visual representations in virtue of first person experience. So there is no real reason to think that even our own self-attributions of propositional attitudes somehow 'get at' the true content of our perceptual states.

There is another more important reason, hinted at above. Recall from the Introduction that the Intentional Stance is an 'idealized rationalistic calculus,' the purpose of which is the prediction and explanation of action. It is an interpretative strategy, based on what an organism *ought to* believe and desire, all things being equal, that allows us to ascribe propositional attitudes to the whole organism – neural content be damned.

As I said above, when a fly buzzes by a frog, the frog ought to believe that a fly, as opposed to a lead pellet, is present. But to attribute a belief about a *fly* leaves open a rather wide range of possibilities of how – or even *if* – the property of being a fly is represented by the frog's brain/visual system. The intentional stance works just insofar as the frog behaves *as if* it sees the fly. As we also saw above, the frog's visual system 'fixes' flies in virtue of one class of ganglion cells that respond to dark, convex, moving edges. We, too, 'fix' flies (not to mention Doritos and salsa), but not through a gang of specialist ganglion cells. What lies between the eyes and tongues of the two species is very different indeed and so too ought the contents of those neural events be distinct. To put the point more generally, each visual system embodies its own particular cascade of specialized visual 'discriminations'; conscious visual perception occurs whenever these discriminations are shared among specialists in just the right way. But if so then the conscious visual experience of two species (say of two mammals) ought to differ in nontrivial ways. Again, we can say that the frog sees the fly just insofar as he behaves *as if* he sees the fly – and that a frog surely does. But so do we.

These, then, are two positive reasons, rooted in Dennett's empirical theories of vision, which predict that there shall not be an 'assignation of contents.'[4] The Intentional Stance will not meet up with neural contents within the stream of consciousness.

PART II

1. Introduction

The conclusions in Part I may seem to be very small ones indeed, certainly ones that Dennett himself could safely acknowledge and incorporate into his repertoire – and I think they would be, were it not for one thing: Dennett's suggestions about vision are turning out to be entirely too prescient. Our most recent neurophysiology of vision suggests that our visual system engages in a drastic division of computational labour, the wholesale subcontracting of visual tasks. This is both good news and, well, 'interesting' news for Dennett. The good news is that Dennett's negative views about the mind will probably turn out to be largely correct: The Cartesian picture of mind is about to come up against hard empirical evidence. The 'interesting' news for Dennett, however, is that this new way of looking at visual function will bring with it many new questions about the functions of the visual specialists, the 'discriminations' they make, and the kinds of informational exchanges that occur between them. My guess is that characterizing what every specialist does as a 'content discrimination' or 'fixation' will turn out to be too theoretically impoverished for our theory of visual function, and hence too impoverished to play the required role in Dennett's (or anyone's) theory of consciousness. While Dennett may just turn out to be right about the nature of visual processing, we will come to see that most of the interesting questions about neural content have yet to be answered. In what follows, I will try to say why I think this is so.

2. Coming Full Circle

Let me begin by talking about sensory 'subcontracting' and why it is so ubiquitous. First of all any sensory system must negotiate the purely formal

considerations around the possibility and feasibility of computational functions in a real time system. When, in the course of day, you see a car drive by, a pedestrian cross the street, a beetle scurry across the floor, or your own hand as it grasps the door handle, what you see, consciously, are a variety of objects – a car, a person, a beetle, a hand – instantiating a common property, namely, *motion*. These motions are, however, relevantly different: pedestrians stroll, beetles scuttle, hands rotate, and cars flash by. Hence, an algorithm that works to discern one kind of motion may not work for any of the others. Consider rigid motion (a moving car) versus nonrigid motion (a hand as it grasps the door knob). Given an image of a car moving relative to a stationary background, it can be assumed that for any point A in the car's image, if A moves from one point to another, then all of its nearest neighbours will be translated in space *mutatis mutandis* (as you go, so go your neighbours). In nonrigid motion, the nearest neighbours of a given point, A, need not retain their positions relative to A across time (think here of how your open hand contracts as it rotates and grasps the door knob). Scene segmentation involving nonrigid or 'biological' motion is an inherently more difficult computational task, one that cannot be performed using the simple algorithms of rigid motion. Of course, it may be possible to devise algorithms that work for both rigid and nonrigid motion but in a real time system, where speed is of the essence, such dual-purpose algorithms may not be feasible. A computational division of labour may well make sense: that is, use the faster, more economical algorithm when possible; when not, employ the slower, more computationally demanding strategy. Hence a variety of computational strategies for a 'single' property, motion.

Dividing (what we see as unified) events into various subcomponents is also entirely common in the visual system. For example, when a marble rolls from A to B in a linear motion, we see, in some sense or other, a seamless event – the motion of the marble through space from A to B. But again, qua event, the processes of early vision subdivide this task: There are separate processes, each using different aspects of the visual stimulus, which compute the onset of motion, direction of motion, its speed or velocity, and its trajectory. These are, then, dissociable mechanisms – it is possible, in theory if not always in practice, to provide a stimulus that will drive one mechanism without affecting the others. Thus, in colour and motion experiments, one can use stimuli that will trigger the 'onset of motion' specialist but not the mechanisms for the calculation of velocity or trajectory. Such stimuli will cause the subject to perceive that a coloured grating has moved, even though the subject is unable to say in what direction or how quickly it moved. Similarly, using another display, one will see the grating as moving in a particular direction, for example, to the right, but not see it as moving at any steady or discernible velocity (sometimes the grating appears to 'hop'). Here, we can see the utility of representing, separately, the various aspects of an object's motion in terms of the different kinds of behaviours that the different components of the event might guide or trigger. For example, a very fast motion detector can be used to capture visual attention or initiate a saccadic eye movement to that point of interest, to where the motion of the marble has begun. A slower 'direction' mechanism might initiate a tracking motion of the eyes (e.g., start the eyes moving to the left so that they can foveate on the marble itself). Finally, a much more complicated and hence slower 'trajectory' mechanism might provide the 'endpoint' for a visual saccade, that is, it might tell the eyes where they ought to land in order to see the properties of the now stationary marble. If there are many different reactions that could be initiated by a single event – a marble moving through space from A to B – there is a clear utility in subdividing the task, representing different properties of that event in response to particular motor needs.

Another kind of visual subcontracting is nicely illustrated by the multiple systems of depth perception. Depth information is essential to vision as a whole given the nature of the task, of determining from a two-dimensional image what lies in a three-dimensional world. We use depth information to discern figure from ground in scene segmentation and to discern other properties of objects, such as location, shape, and texture. We use depth information to differentiate (through depth-ordering cues) motion in the retinal image from the motion of surfaces in the world. We use depth cues to see motion toward us or away from us, motion we would not otherwise be able to see. Now, as any classic text in psychology will tell you, the primate visual system uses a number of different strategies for computing depth: Depth can be discerned from pictorial cues, occlusion cues, stereopsis, stereodepth-in-motion, motion parallax, and relative size, among other cues. These are computations that, for the most part, make use of different properties of the visual scene/image, thereby extending the range of conditions under which depth can be discerned. For example, the computation of depth by stereopsis is possible because of the small distance between our eyes. If you look at one point in visual space, the images of most other points will be slightly displaced on the two retinae, relative to one another. So, given two fixed points in the external world at different depths, the relative displacement of their left and right retinal images can be used to compute their relative depths. Depth from stereo-motion, by contrast, depends on the relative motion of the left and right images. When an object, directly in front of you, moves toward you, the left and right retinal images move outward at the same velocity and the image expands in size; when the object moves away, the left and right images move inward and they contract. More generally, the direction of motion in depth can be computed by two different functions: by the ratio of the velocities of the left and right retinal images and by the rate of change of image size, relative to the speed of translational motion. Yet a fourth computation for depth perception depends on parallax motion. When you move your head back and forth parallel to the plane of fixation, the retinal images of any objects behind the fixation plane will move in the same direction as your head; the images of objects in the front of that plane will be displaced in the opposite direction. For example, suppose you are looking at a vase of flowers on a table. In front of the vase (nearest you) is a water glass. Sitting opposite you at the table, behind the vase, is another person. Move your head to the right. Here the image of glass will move left; the image of your table companion will move right.

Note that the four different computations for depth probably use at least four different kinds of image properties: relative displacement of the left and right images, the ratio of velocities for the left and right image, the change of image size relative to translation motion, and direction of image motion relative to fixed point. As such, each will be possible when and only when the retinal images have these properties. For example, stereopsis will not work when the left and right images are not adequately displaced, that is, when the two fixed points are quite far away. You can not use stereopsis to determine which building in the distant skyline is in front of another (although you can use occlusion cues). Similarly, if you are blind in one eye, you cannot compute stereo-depth from motion, but given your single field of vision, you could use motion parallax (take note all one-eyed philosophers). Classic textbooks on psychology usually cite this reason as the evolutionary rationalization of the brain's multiple depth mechanisms: No matter what the nature of the visual scene, the visual system will almost always be able to compute whatever depth information is needed, using one depth system or another. Multiple representations offer the visual system a 'fail safe' system for a very important property of the environment, depth.

While the above is no doubt true, what interests me here is not merely that depth is used in so many computational/visual tasks but that depth information is used to perform an indefinite number of motor tasks – to put on our socks, tie our shoelaces, walk over uneven ground, kick a soccer ball, skim past the goal posts, and on and on. Importantly, different types of action require different representations of depth information.

Let me give two examples, that of depth from stereopsis and depth from stereomotion. What the psychophysics of stereopsis have shown is that depth perception of this type is maximally sensitive at the fixation plane and falls off rapidly both before and after. This is a system, in other words, which provides depth information about objects, not relative to the viewer, but relative to a fixed plane – depth information that rapidly diminishes in accuracy the further in front of or behind the fixed plane an object moves. This is a system suited to a particular kind of motor task, of manual movements relative to a fixed point, say, to threading a needle, plucking a Cheerio off the floor, picking grapes from the vine, and so on.

One hypothesis about how the brain performs 'fine stereopsis' or makes fine-grained static disparity judgments is Richards's Pool Hypothesis. According to this theory, stereo-depth is computed by three separate pools of neurons: one pool sensitive to objects in front of the plane of fixation (the plane on which the eyes are focused), one to objects behind the plane, and the third to near objects around the plane itself. On this view, the 'near' and 'far' pools act in an antagonistic fashion, such that the perceived depth results from the relative activity of the two pools; the third pool serves to enhance depth acuity at or near the fixation plane. In 1981, Poggio and Talbot reported finding two classes of neurons in both V1 and V2 of macaque monkeys, just as predicted by Richards's Pool Hypothesis: a class of neurons sharply tuned for near-zero disparity (i.e., for points very near the plane of fixation) and a class of reciprocally activated 'near' and 'far' cells. Impressively, the tuning characteristics of the two populations were reasonably well matched to our judgments of depth – or to put this another way, what the cells do; as a population, is surprisingly similar to what we do when making fine disparity judgments. Despite this correlation, however, Poggio's strong conclusion (1984; Poggio and Talbot 1981) that the depth cells of V1 alone are responsible for perceived depth by stereopsis was probably premature. Recent experiments (Cumming and Parker 2000) suggest that the depth cells of V1 signal only local disparity while perceived depth also depends on more spatially distant cues. In a sense, though, this was to be expected. That is, insofar as there are distinct 'depth specialists' in cortex, it would be surprising to find one of them in V1, the first cortical visual area after the retina. More realistically, the depth neurons in V1 form a preliminary stage to perceived depth, a process completed in extrastriate cortex.

To catch a ball or bat a badminton birdie, however, the visual system must judge the depth of a moving object, the ball relative to oneself or

the birdie relative to the racket. As explained above, the received view, at least until the mid-1990s, was that depth from motion was computed by two independent systems, one that uses change in image size relative to the speed of translational motion and the other of which uses the relative velocities of the two images. Exactly how motion in depth is actually discerned is now a matter of some debate but there is at least some agreement about the psychophysical facts about motion-in-depth. We can model our capacity to perceive motion-in-depth in terms of four channels: two channels for objects behind the plane of fixation, one for rightward motion and one for leftward motion; and another two channels for objects in front of the fixation plane, again for leftward and rightward motion. On this hypothesis here is a nonlinear polar representation of visual space, of depth relative to the viewer's face, with the lion's share of representational space devoted to the cone of directions that intersect the head. Again, the rationale for this representational scheme is intuitively obvious – we are most 'interested' in objects that are quickly approaching our heads. Intuitively, that is, it is the baseball that's going to hit you smack in the face that is interesting, not the ball about to clear your left shoulder – and the representational system reflects this fact. Suggestively, the same class of 'far' and 'near' neurons in V1 (mentioned above) also subdivide into two more classes of neurons, those tuned to leftward motion and those tuned to rightward motion. Again, as in the case of static stereopsis, the direction-sensitive neurons are unlikely candidates for the specialists that compute motion-in-depth computations, although they may well be the first specialists to contribute to that (or those) computation(s) performed further downstream.

What the above two examples illustrate is just why different actions will require different encodings of the same objective property. What those examples do not give is any indication of the extraordinary heterogeneity of the depth systems that will be required for motor control. To illustrate this point, I can think of no better way than merely to present the abstract of a recent review article on the depth perception systems involved in hand control in parietal cortex alone. What matters here is not whether the reader understands where the various parietal regions are, or can keep straight which is which (I cannot), or understands all of the terms in the nomenclature of depth perception. It does not even matter whether the research conclusions summarized in this review are correct (for I suspect that some of these conclusions exceed the experimental evidence as I understand it). Still, for a sense of the sheer complexity of the situation, the following ought to prove illuminating. Without further ado, then, the abstract of "The parietal association cortex in depth perception and visual control of hand action":

> Recent neurophysiological studies in alert monkeys have revealed that the parietal association cortex plays a crucial role in depth perception and visually guided hand movement. The following five classes of parietal neurons covering various aspects of these functions have been identified: (1) depth-selective visual-fixation (VF) neurons of the inferior parietal lobule (IPL), representing egocentric distance; (2) depth-movement sensitive (DMS) neurons of V5A and the ventral intraparietal (VIP) area representing direction of linear movement in 3-D space; (3) depth-rotation-sensitive (RS) neurons of V5A and the posterior parietal (PP) area representing direction of rotary movement in space; (4) visually responsive manipulation-related neurons (visual-dominant or visual-and-motor type) of the anterior intraparietal (AIP) area, representing 3-D shape or orientation (or both) of objects for manipulation; and (5) axis-orientation-selective (AOS) and surface-orientation-selective (SOS) neurons in the caudal intraparietal sulcus (cIPS) sensitive to binocular disparity and representing the 3-D orientation of the longitudinal axes and flat surfaces, respectively. Some AOS and SOS neurons are selective in both orientation and shape. Thus the dorsal visual pathway is divided into at least two subsystems, V5A, PP and VIP areas for motion vision and V6, LIP and cIPS areas for coding position and 3-D features. The cIPS sends the signals of 3-D features of objects to the AIP area, which is reciprocally connected to the ventral premotor (F5) area and plays an essential role in matching hand orientation and shaping with 3-D objects for manipulation. (Sakata, Tiara, Kusunoki, and Tanaka 1998, p. 350)

I take it that my central point about depth perception is now obvious: Given the multiple processes involved in depth perception, it is misleading to say that there is 'a' depth system doing 'depth detection' or 'depth discrimination.' There is no single depth mechanism the function of which is to represent depth information. Nor is there likely to be any 'canonical' form of depth representation, one that mirrors, faithfully, the objective spatial relations between objects in the world and hence that anchors our conception of objective depth relations. Rather, by using a variety of mechanisms for discerning depth and a variety of representational formats, the depth systems, taken together, make possible the efficient control of diverse visually guided motor behaviours as well as whatever cognitive 'behaviours' that make use of visual depth information. *And somehow*, in all this, our conception/perception of the objective spatial relations that hold between us and the objects before us is anchored as well.

From the above examples alone, there are at least three very deep puzzles about how the contents of the 'specialists' are unified (or at least seem to be) in conscious experience that Dennett must explain before any theory of consciousness is possible. (And as Dennett well knows, if you are not going to explain them, you must at least explain them away.)

(i) When you see a car drive by, a pedestrian cross the street, a beetle scurry across the floor, or your hand as it grasps the door handle, you see a variety of different motions, no doubt each 'underwritten' by a different computational mechanism. Yet nonetheless you see each motion as *motion*, as an event of the same type. You see an object make a change in space over time. But how, in Kantian terminology, does 'subsumption under a concept' work? What does 'recognition as an instance of a type' amount to in modern neural terms?

(ii) When you see the marble, as it rolls from A to B, you perceive its movement as a single, seamless event despite the multiple mechanisms which must sustain this perception. In all likelihood, the onset of motion, its direction, velocity, and trajectory are all encoded separately, by mechanisms with different times of onset, temporal duration, and sensitivities. From your viewpoint, as a conscious subject, all this is news to you. You just see the marble roll from A to B.

This would seem to be a quite different type of 'unification' than the one above, in (i), however. For one, while it may be possible, by means of clever psychophysical experiments, to tease apart the independent mechanisms at work for motion onset, velocity, and so on, it may not be possible for the observer likewise to 'tease apart' the 'informational contribution' of these mechanisms in conscious experience. You will readily admit, when a car zooms by, that *that* was motion; you will readily admit, when you see your cat chased up a tree, that *that* was motion. And if you are a shrewd psychophysicist, one day you may even be able to sort these different perceptions of motion into distinct categories, categories defined by the distinct motion specialists that underlie them. But there is something rather more 'subpersonal' about the mechanisms that underlie the perception of each 'movement event.' You acknowledge, of course, that the movement of the marble has a beginning and an end, and that the marble moves at a certain velocity throughout, first accelerating and then decelerating to a stop. But there could be (and indeed almost surely are) many more specialists at work in your perception of that single event. If so, you would simply have to take the psychophysicist at his or her word. That you cannot necessarily recognize the effects of these specialists in consciousness experience *as* particular effects of a certain kind does not impugn the psychophysicists' results. The psychophysicist need only show that without the mechanism at issue, you cannot perform the appropriate judgments about the movement of the marble, and that, moreover, in an 'A or B' forced choice test, you can reliably indicate the difference between how the motion seems with and without the mechanisms at issue. The same conclusion holds, I suspect, for much of what goes on in, say, shape perception and object recognition.

(iii) The example of depth perception furnishes us with yet another puzzle about the unity of perceptual content. Over the course of a few moments, during the course of any visually guided task, say threading a needle, numerous different mechanisms of depth and spatial perception will be involved, each using different representational formats. You see the depth of the thread relative to the needle as the thread goes in and out of focus, using depth from stereopsis. You also see the needle and thread as but two objects in the room, as occupying fixed positions relative to other objects. For example, behind the needle and thread, across the room and out of focus, is the television; in front of the needle and thread are your equally 'fuzzy' moving arms. These facts you know from motion parallax. (Or to put this in a less misleading way, if you concentrate first on the foreground and then on the background while holding steady the needle as your fixation point, this is what you know from motion parallax.) No doubt, occlusion cues tell you that the wicker chair stands behind your hands and the needle and the television across the room. And so on, for the many different depth processes. What seems striking to me is that, despite the plethora of depth mechanisms, each with a distinct associated phenomenology, you nonetheless maintain some conception of the needle and thread as being, at a given moment, in a single objective position. I see the thread weave in and out of focus as I move it about, but I do not conceive of the thread as having moved so far away I can no longer see it, as having backed up into the television. We do not mistake the contributions, the form and spatial resolution, of, say, depth from stereopsis for that of, say, depth from parallax motion. We seem to move effortlessly between them, hardly ever noticing where one leaves off and another begins. Instead, what we tend to notice are only the rare occasions when such transitions fail (e.g., when one looks through a screen door and the standard cues for stereopsis are degraded). So what makes the depth systems co-operate so seamlessly? What makes it possible for us to move back and forth between distinct – and some people might say incommensurate – representational systems?

It is also interesting that we do not focus on the contributions of one depth mechanism as providing *the* objective spatial properties of the world. From your point of view, the information that you get about depth as you thread the needle might be roughly characterized as coming in the following 'gradations': 'way out – nowhere near the needle,' 'ballpark,' 'homing in,' or 'on target!' This does not shake your sense of spatial objectivity, however. The fine spatial resolution around the fixation point of static depth from stereopsis does not 'trick' you into thinking that somehow objective space itself is more 'fine-grained' around the needle. In some sense or other, we 'see' the deliverances of those mechanisms as being *in alignment with* the objective spatial relations of the visual scene as we understand them. Depth from stereopsis informs our spatial perception (as, no doubt, do a myriad of other spatial mechanisms), but does not impugn its objectivity. So how exactly does *this* work? What maintains our sense of spatial unity and objectivity in the face of representational heterogeneity – hence, in the face of a heterogeneity of both content and phenomenology? Good question.

Or rather one good question among many – the many one would need to solve in order to connect content with consciousness. What this new understanding of vision results in, I suspect, is a plethora of questions about content – about how the 'subpersonal' processes of the specialists eventuates in 'our' visual experience.

3. About *Qualia*

And what, one may ask, happened to *qualia*? Recall, from Part I, that Dennett's view is that there are no *qualia*: There are no 'purely qualitative' inputs to perceptual processes and no 'purely qualitative aspects' of perceptions either. Assuming that Dennett's example of the robot has been convincing (or that most people no longer believe that the inputs to perception could be qualitative states), certainly it seems to many of us that our conscious visual perceptions have qualitative aspects to them. I see the hyacinths on my desk as a particular buttery yellow and I smell that almost overpowering, sweet, slightly spicy fragrance that marks that species of flower. If those are not 'qualitative' aspects of visual experience, then what are they? I want to take up this question to show how the questions about neural content posed above may well offer Dennett some new avenues for explaining (away) *qualia*.

Dennett's standard answer to the question "what are *qualia*?" is this:

> Sensory qualities are nothing other than the dispositional properties of cerebral states to produce certain further effects in the very observers whose states they are. (1998, p. 146)

As Andrew Brook points out (this volume), this answer works reasonably well for some cases but not very well at all for others. Take, for example, Dennett's explanation of what happens to a subject's phenomenology when he or she is fitted with inverting goggles, goggles that reverse the retinal image from left to right, and invert the retinal image to 'right side up' (oddly enough, the retinal image is normally 'upside down'). At first, subjects report just what you would expect, that they see the world as upside down and reversed left to right. After adaptation, subjects report that the world has once again regained its 'right side up-ness.' In between, pre-adaptation, subjects do not know quite what to say about their experience. As Dennett points out, if you conceive of the output of visual perception as a static image, then you will ask this: "*Have they adapted by turning their experiential world back right side up, or by getting used to their experiential world upside down?*" (1991a, p. 397). The problem is that neither option explains the subject's pre-adaptive experience. It is hard to say what could be happening to a single visual image such that the world appears neither right side up nor upside down, or why the question would start to seem senseless (as it does) by the mid-point of adaptation. On the view of vision as governed by multiple specialists, however, this is just what one would expect: one by one, the multitude of visual/motor specialists adjust and co-ordinate their outputs until finally the goggle wearer is able to get around the world much as before. If there is no single 'product' of vision, but rather individual answers and changing coalitions that control behaviour, then behaviour ought to follow the adaptation of individual specialists and coalitions, that is, behaviour ought to improve over time, in a series of leaps and plateaus.

What Dennett does not explain about the example is the subject's pre-adaptive phenomenology, or more accurately, why the subject cannot seem to get a grip on the phenomenology of his or her partly adapted visual system in order to explain it. The suggestion made at the end of Part I about why self-attributions of conscious content are not to be trusted given both the abstract and 'need-to-know' nature of visual perception does not apply here. The problem is not so much that the subject's descriptions are inaccurate (although they may be) but that the subject cannot settle on *any* description that seems reasonable. Perhaps Dennett might say in response that the subject is too much in the grip of the Cartesian metaphor (of the static image) and hence for this reason does not have the resources to describe a world that is neither right side up nor upside down. This is one possibility. Then again, Dennett might say that there is no reason to think that, in the midst of adaptation, the responses of the specialists will necessarily cohere or be consistent with one another. In Dennettian terms, when the question is asked, "Are the kumquats on the plate?," there could be a chorus of answers: "No, they're *under* the plate!," "No, they're *on* the plate!!," "No, *under* the plate!!!," "Which kumquats are you talking about – the ones under the table?," and "How should I know?" Hence, it would not be surprising if the subject were confused about just how the world seemed. This is another explanation. Still, Dennett needs to solve this puzzle in some way or other. Why is it that the subject, by concentrating on the properties of the *world*, cannot accurately describe which dispositions have/have not been adjusted?

Dennett's answer to the problem of qualia looks even less plausible vis-à-vis the classic philosophical cases of 'pure' *qualia* – for tastes, sounds, colours, and so on. If my experience of 'buttery yellow' is merely 'a complex of dispositions,' then exactly what dispositions make up this complex? Here, the problem is this: The only dispositions that readily come to mind are *reactions to* a particular qualitative experience – what we would say or do in response to a visual experience of the appropriate qualitative kind. Thus, on seeing the buttery yellow hyacinths, you might exclaim about their colour, wonder why you ever started buying margarine, recall a dress you had as a child, or suddenly remember that you have forgotten the birthday of your mother (who loves yellow hyacinths). This is what comes about when your experience of the hyacinths has a 'buttery yellow' aspect. But how could the dispositions to behave in response to *qualia* literally *be* the qualitative experiences that cause them? Nor does the example give us any reason to think that qualitative experiences, in and of themselves, could not play the causal roles that folk psychology suggests. For 'classic *qualia*,' Dennett's theory is really quite puzzling.

What Dennett can gain from the wholesale 'subcontracting' of visual specialists is a way to address these two puzzles – of why we lack access to various parts of our visual experience (the case of the inverted goggles) and how it might be that 'real' *qualia* can be explained in terms of dispositions (the case of buttery yellow). Once one acknowledges that the ontology of the perceived world is not mirrored by the ontology of visual processing – that what appear to us as singular properties, objects, and events in the world are unlikely to be discerned by single specialists – Dennett gains a very useful distinction.

On the one hand, there is that which makes certain information conscious – call it 'shared' information. In a parallel computational system such as vision that works (at least in part) through mutual constraint satisfaction, the 'answers' of any modules must be 'share-able' with at least some of the other modules. For example, if an object actually has moved from left to right, and if the luminance contrast of the object with its background has changed, then the spectral contrast ought to have changed as well (because objects that differ in luminance almost always differ in spectral reflectance). So spectral contrast information and luminance contrast information ought to be available, at the beck and call of different specialists including those that perform motion analysis of various types.

On the other hand, given both the disjoint and heterogeneous outputs of the various specialists, plus the fact that we manage to see a world of orderly and unified objects, properties, and events, there must be some way – some ways – in which this is accomplished. Somehow out of the labour of many specialists, we come to have a perceptual ontology, one that is distinct from the outputs of the specialists per se. We come to see unified objects with a multitude of properties, to see coherent events in orderly succession, to understand the properties of world as objective or as independent of our particular perceptual presentations. Now, no doubt these various processes of 'unification' require that information be exchanged between modules. If so, the means by which we come to gain our perceptual ontology will result in conscious experience. Still, there is no reason to think that all information exchanged between modules is for the purposes of unification. For example, when the linear motion module requests luminance contrast information, it may do so in order to compute linear motion, not in order to unify the luminance properties of the moving object with its property of motion. So a distinction seems in the offing: There is (very likely) a difference between the general processes in virtue of which information is shared between specialists and whatever specific processes result in the perceived world's ontology.

What the distinction gains Dennett is a way to explain why various aspects of conscious visual experience seem to be 'purely' qualitative and why we have such a difficult time explaining just those (and pathological) aspects of our own experience. Start again with Dennett's view that we look out not in. Insofar as we have any access to our phenomenology, it is through looking at *the world* and saying how the world seems to us under various visual conditions. I move my head back and forth, this way and that way, I rotate the object so that the light strikes it at various angles, in order to see what is the case. Suppose then that there is distinction between those processes that make information a part of conscious experience and those which

serve to 'carve' the world into ontologically sound divisions. When we look out at the world, the content of conscious experience is determined by the multitude of on-going processes through which information is shared. Yet when we try to see *what* is in the world, when we ask ourselves *what* we are looking at, we can see the world only in terms of our perceptual ontology, only by using whatever capacities make it possible to 'carve up' the world into objects, properties, and events. What arises, then, is a profoundly non-Kantian possibility: *There may be a difference between that which informs our visual phenomenology of the world and that which we can grasp qua properties of the world.* There may be a distinction between the (conscious) discriminations that visual specialists make and the properties *we* can see in the world. If so, we will be able to understand our visual phenomenology of the external world only insofar as we can corral the (conscious) discriminations of the specialists into a stable ontology – here, an ontology of the external world, not of phenomenal experience. Our access to our own phenomenology is available to us only *through* our intentional perceptions of the world, through the properties we see as existing in the external visual world. And here, at last, is the punch line: There is nothing to guarantee that we can (or can learn to) corral *all* of the discriminations made by the visual system that affect our visual phenomenology into perceived properties of the world.

Let me go back to the examples of 'subcontracting,' in particular the example of depth processing from motion parallax. When told how motion parallax works, most people find the story quite surprising (certainly I did). It is just not self-evident to most people what will happen to the visual image when fixation is maintained and the head moved – that images in the background will move with the head, and that images in the foreground will move in the opposite direction. What makes motion parallax a nice example, is that the subject can come to understand how the 'subpersonal' visual process works through demonstration. If you fixate on the window frame, and squint the right way, the distant mountains will 'move' with your head, while the flowers on the window sill will 'move' in the opposite direction. In some sense, you see how motion parallax works. Note however that, through squinting, the subject is not gaining access to the retinal image or indeed to any image at all. Rather, normally, when the visual system discerns depth through parallax motion, the image motion, insofar as it can be co-ordinated with head movement, is discounted. Stationary objects in the visual scene are seen as stationary. When one 'squints' in order to understand motion parallax, your visual system manages to do what you might have done by drawing a picture – you are able to understand the optics of the situation.[5] While you can still feel your head moving, and while you know that, say, the mountains are *not* moving in objective space, nonetheless you are able to see the mountains 'as appearing to slide back and forth.' That the mountains are not actually sliding – that the motion is not veridical – is a part of your phenomenology. But then, so too are the mountains qua mountains. You have not, through some miracle, managed to attend to your retinal image qua a retinal image. You are still looking out not in. Nor, in all likelihood, have you managed to access the inner workings of whatever 'subpersonal' specialist performs depth from motion parallax. It is entirely possible, that is, that a person who, through some deficit, did not have depth from motion parallax might nonetheless be able to 'see' what optical facts make motion parallax possible in the ordinary case. What you have managed to do is to see the mountains 'move,' by accessing information about image movement and using it in whatever processes normally result in intentional visual perception.

For the subject who wears inverted goggles, the situation is quite different. When the goggles are first put on, there is a handy way to re-conceive of his or her visual experience – "It is just like the world was upside down." As the sensory/motor adaptations – realignments – come into being, one by one, however, the subject will have no method in general with which to re-conceive of his or her visual experience in terms of perceived properties of the world. No doubt if we knew more about how the processes of adaptation, and, in addition, more about the normal processes of vision, the psychophysicist might be able to make some helpful suggestions (as in the case of parallax motion). But again, there is no reason to think that the adaptive processes can be made to neatly mimic any aspects of the normal phenomenology of seeing the world. (Very similar points could be made, I think, about attempting to recognize the multiple processes that underlie the normal perception of a marble moving from A to B.)

There is obviously an enormous amount more that needs to be said about this suggestion but perhaps enough has been said to say a little about the 'hard case,' the classic cases of *qualia*. First, as I said above, when we try to imagine how *qualia* could be complexes of dispositions, all that comes to mind are our reactions to a qualitative experience – for example, the experience of 'buttery yellow' reminds me of how much I dislike margarine. What the discussion of subcontracting makes clear is that properties we perceive as unified need not be 'underwritten' by a single kind of information, by the output of a single specialist. We know, that is, that the perception of a marble rolling from A to B is 'underwritten' by the results of numerous specialists – of movement onset, velocity through luminance processing, velocity through spectral processing, of trajectory, and so on. We also know that these individual components of the perception probably have quite distinct causal effects or uses. A very fast motion detector could be used to

capture visual attention or initiate a saccadic eye movement to that point of interest; a slower 'direction' mechanism might initiate a tracking motion of the eyes and so on. But if we could find this out about motion perception, so, too, could we find this out about any of the other *qualia* that have historically plagued philosophical discussion. There is no reason why colours, pains, and so on could not be similarly 'decomposed' into perceptual parts (see Akins and Hahn, in progress) and hence into complexes of causal dispositions, each without mention of a 'purely qualitative aspect' as cause or effect at all. Such a decomposition is just what would show that felt *qualia* are, in fact, exactly analogous to the feeling of upside down-ness, to the 'qualitative experiences' induced by inverting goggles. If the phenomenology of 'upside down-ness' is no barrier to explaining consciousness then neither should be the phenomenology of pain or colour.[6]

Second, the theory would explain just why *qualia* seem both complex and ineffable, why we have so little access to, say, colours, other than to speak in terms of 'the very blue' of the teapot before me now. We know, for example, that it is very easy for a normal trichromatic observer to discern that one colour is different than another colour sample, when both samples are of identical hue and saturation. It is impossible to tell, however, which of the two samples is brighter. Similarly, given two colour samples that differ in either hue or saturation, it is extremely difficult – in some cases impossible – to tell along which dimension the difference occurs. But if the samples are put side by side, that there is a difference is entirely obvious. In other words, in colour perception, there seem to be any number of factors that make a difference to how the colour of an object or surface is seen. Yet, it does not follow that we can differentiate or categorize the phenomenal differences, or at least not in virtue of merely looking at the colour sample. This inability to delineate, say, brightness in vision as an independent property of a surface under normal viewing conditions is what makes colours seem both complex and 'ungraspable.' We can detect differences in surface brightness. Our phenomenology is affected by that discrimination. But under normal conditions, we cannot see brightness qua a distinct stable property of coloured opaque surfaces.

Certainly it seems to me that this is one route that Dennett could go to further dismantle the Cartesian theatre. Whether or not he will go that route is an open question.

4. Conclusion: Dennett and Neuroscience

Were I to don a fortuneteller's garb, I would say that Dennett's negative views, his arguments against the Cartesian theatre, will come to be acknowledged as largely correct. Indeed, already so many of Dennett's criticisms have been so thoroughly incorporated that no one seems to remember having ever held the criticized views, for example, the naïve view of perception. Of his positive views, my guess is this: In retrospect, they will be seen to have been profoundly prescient, at least in outline form. This is not a view I held fifteen years ago. In fact, it is a view I have come to only gradually over the last decade, largely in virtue of the last six years or so of neuroscience research. What explicit effect Dennett will have on the actual course of neuroscientific progress – what role he will play as an acknowledged intellectual force in the community – is another question. At this point, what Dennett lacks – what we all lack, philosophers and neuroscientists alike – is a substantive theory of neural content/representation. At least as I understand much of the recent neuroscience, *most* of the interesting questions about neural content have yet to be asked, much less answered. So, the question of Dennett's influence on the neurosciences is not unrelated to the question of the influence of the neurosciences on *him*. Should Dennett join the party, and I certainly hope that he will, his influence may well be profound.

Notes

1. Throughout his work (e.g., "Language and Understanding" in *Content and Consciousness* [1969], "How to Change Your Mind" in *Brainstorms* [1978a]), Dennett has stressed the role of verbal expression in how we understand (or misunderstand) our 'thoughts.' Very briefly, Dennett suggests that when we 'express our thoughts' with sentences of a public language, this fosters the illusion that the *causes* of our self-attributions are both sentence-like in syntax (i.e., compositional) and sentence-like in meaning (i.e., endowed with whatever 'firm edges' of semantic content that a public language provides). Says Dennett:

> [O]ne point here is paramount: while it is certainly true that there is in general no better way to tell what someone (oneself included) thinks than seeing what he says, if one views the clues one gets thereby on the model of say, the publication of a poem or the release (by the Vatican library) of a heretofore sequestered volume, one may well be making a mistake along the lines of supposing that a head cold is composed of a large set of internal sneezes, some of which escape. The self-questioning process that individuates belief expressions so crisply need not be revealing any psychologically important underlying individuation (or beliefs, presumably) but be just an artifact of the environmental demand for a peculiar sort of act. (1987c, p. 115)

The request for a statement of belief from the subject merely *precipitates assent to a public sentence*. We choose a sentence to which, on the whole, we can assent, now that we stop to consider the linguistic options. But this choice says little about its causes.

Following Churchland, Dennett asserts that "the process that produces the data of folk psychology, we claim, is one in which the multidimensional complexities of the underlying processes are projected *through linguistic behaviour*, which creates an appearance of definiteness and precision" (1991b). What folk practice reveals, however, is that in some sense we do not take belief ascriptions 'at their word.' We presuppose that the content specified is merely 'rough and ready' content, that revisions, emendations, or even 'ad hoc' restrictions made by the subject are all fair game, a standard part of self attribution. Whatever the merits of this line of argument (and I myself think it has merit), the evidence is 'circumstantial' at best. These are arguments that show only what could be the case or what neural causes are consistent with our actual practices of belief attribution.

2. Dennett's point above is that these kinds of knowledge are not analogical. That is, what the architect knows is which icons represent which features of the world. But on Descartes' view, the innate knowledge of the mind is confined to the essence of material objects, not what properties of objects are correlated with what qualitative aspects of the phenomenal array.

3. Dennett does not deny that neural representations could be 'picture-like.' Finding an orange plastic kumquat in visual cortex is a logical possibility. But Berkeley's point, and Dennett's, is about the relation between phenomenal events and neural representations, as opposed to neural representations and their objects.

4. This goes back to an old joke between Dennett and myself. When I worked for Dan at Tufts, it became very clear to him that most of my vocabulary had been learned not from conversation but from reading. Hence, much of what I 'knew' about the English language was not quite right. One day, Dan discovered that I had made it to the age of twenty-seven with the unshaken belief that the word 'assignation' meant roughly 'an assignment to or attribution of.' Thus, much of my dissertation had been devoted to what I called 'the assignation of contents' – you know, what was involved in an assignation, under what conditions they occurred, what 'the experts' had to contribute, and other such standard philosophy of mind fare. Needless to say, Dan being Dan, found this exceedingly amusing – and no less so when I demanded that he look it up 'assignation' in the Oxford English Dictionary. What can I say? There are no assignations of content, not even in the stream of consciousness.

5. I owe this point to Martin Hahn.

6. This is, in fact, what Dennett attempted to do with some success in an early paper, "Why You Can't Make a Computer That Feels Pain" (1978d).

References

Akins, K. 1996. Lost the Plot? Reconstructing Daniel Dennett's Multiple Drafts Theory of Consciousness. *Mind and Language* II:1–43.

Akins, K., and M. Hahn. More Than Mere Colouring. In progress.

Cumming and Parker. 2000. Local Disparity Not Perceived Depth Is Signalled by Binocular Neurons in Cortical Area V1 of the Macaque. *The Journal of Neuroscience*, June 15, 20(12):4758–67.

Dennett, D. 1969. *Content and Consciousness*. London: Routledge and Kegan Paul.
Dennett, D. 1978a. *Brainstorms*. Montgomery, VT: Bradford Books.
Dennett, D. 1978b. Brain Writing and Mind Reading. In Dennett 1978a.
Dennett, D. 1978c. The Abilities of Men and Machines. In Dennett 1978a.
Dennett, D. 1978d. Why You Can't Make a Computer that Feels Pain. In Dennett 1978a.
Dennett, D. 1978e . Current Issues in Philosophy of Mind. *American Philosophical Quarterly* 5:249–61.
Dennett, D. 1987a. *The Intentional Stance*. Cambridge, MA: MIT Press /A Bradford Book.
Dennett, D. 1987b. True Believers. In Dennett 1987a.
Dennett, D. 1987c. When Frogs (and Others) Make Mistakes. In Dennett 1987a.
Dennett, D. 1987d. Styles of Mental Representation. In Dennett 1987a.
Dennett, D. 1987e. Three Kinds of Intentional Psychology. In Dennett 1987a.
Dennett, D. 1987f. Evolution, Error and Intentionality. In Dennett 1987a.
Dennett, D. 1991a. *Consciousness Explained*. Boston: Little Brown & Co.
Dennett, D. 1991b. Real Patterns. *Journal of Philosophy* 88:27–51.
Dennett, D. 1998. Real Consciousness. In his *Brainchildren*. Cambridge, MA: MIT Press/A Bradford Book, 131–9.
Marr, D. 1982. *Vision*. San Francisco: Freeman Pub.
Poggio, G. F., and W. H. Talbot. 1981. Mechanism of static and dynamic stereopsis in foveal cortex of the rhesus monkey. *Journal of Physiology* 315:469–92.
Poggio, G. F., 1984. Processing of stereoscopic information in monkey visual cortex. In G. M. Edelman, W. E. Gall, and W. M. Cowan, eds., *Dynamic Aspects of Neocortical Function*. New York: Wiley, 613–35.
Regan, D., J. P. Frisby, G. Poggio, C. M. Schor, and C. W. Tyler. The Perception of Stereodepth and Stereo-motion: Cortical Mechanisms. In L. Spillman and J. S. Werner, eds., *Visual Perception: The Neurophysiological Foundations*. San Diego: Academic Press, 317–47.
Prince, S., A. Pointon, B. Cumming, and A. Parker. The Precision of Single Neuron Responses in Cortical Area V1 during Stereoscopic Depth Judgements. *Journal of Neuroscience* 20(9):3387–400.
Ryle, G. 1949. *The Concept of Mind*. London: Hutchinson.
Sakata, H., M. Taira, M. Kusunoki, A. Murata, and Y. Tanaka. 1998. The Parietal Association Cortex in Depth Perception and Visual Control of Hand Action. *Trends in Neuroscience* 21(2):350–7.

ARTIFICIAL INTELLIGENCE AND EVOLUTIONARY THEORY

We have explored the nature of Dennett's contribution to consciousness studies and the explanation of child, animal, and adult social behaviour, and we have looked at two deep running concerns with his contributions in both areas. In this final section, we will explore two more areas of Dennett's influence. One is artificial intelligence research. The other is theory of evolution.

Dennett has been writing on artificial intelligence for thirty years and has generally been sympathetic to the idea that genuine intelligence and even consciousness can and probably will be achieved in systems built very differently from the human organism, systems built out of silicon chips or neural network nodes for example. Yorick Wilks, chair of a major school of computer science and an internationally known artificial intelligence researcher, casts a rather jaundiced eye over Dennett's work in this area. Wilks's view is that many of Dennett's claims are not taken seriously by AI researchers, a leading example being his well-known distinction between what the brain can process (physically expressed differences sometimes called syntax, in a vastly extended sense of that term) and what it cannot (meaning, informational content). Furthermore, says Wilks, Dennett is far more optimistic than researchers actually doing the work. Wilks's diagnosis of these problems is that they reflect a deep-running gap in cognitive science. Cognitive science lacks clear canons for determining when claims made in it are justified and when they are not.

Dennett's role in the progress of evolutionary theory has taken a remarkable turn in the past few years. Few major scientific debates can ever have been conducted in the popular press as much as this one, or with as bad a ratio of light to emotional heat. Much of the dispute has centred on Dennett's views on the evolutionary foundations of consciousness. Don Ross surveys the battlefield and urges that both Dennett's opponents and popular commentators on the dispute have seriously misunderstood him. In particular, they fail to understand how his position grows out of his prior theory of intentionality (the term is defined in the Introduction). One commentator who *has* understand what is at stake is Jerry Fodor. Fodor