

## IN REPLY

Nicholas Humphrey

It is very difficult, now that everybody is so accustomed to everything, to give an idea of the kind of uneasiness felt when one first looked at all these pictures on these walls. . . Now I was confused and I looked and I looked and I was confused.

Gertrude Stein, on her reaction to first seeing Cubist paintings. *The Autobiography of Alice B Toklas*.

I heard Daniel Dennett, last year, winding up his Royal Institute of Philosophy Millennial Lecture with the prediction: "I anticipate a day when philosophers and scientists and laypeople will chuckle over the fossil traces of our earlier bafflement about consciousness" (Dennett, in press). And I thought at the time: that's not quite right. Future philosophers won't *chuckle* over the difficulties we now have with understanding consciousness - any more than we, today, *chuckle* over the difficulties people had two centuries ago with understanding the evolution of life.

We chuckle when something is comical. But the spectacle of earnest seekers after truth trying their very best to make sense of something that is presently beyond their capacity is, I'd say, more chastening than comical. In the case of evolution, when we look back on the efforts of pre-Darwinian thinkers to understand the origins of biological design (including even such giants of their time as David Hume), their struggles seem both heroic and tragic. It's like watching the earliest attempts of climbers to climb Mount Everest without the help of oxygen.

Of course, ever since Darwin we have had access to the BIG IDEA - the idea of natural selection - which today allows every schoolchild to understand biological design, and which makes those earlier efforts look so footling and incompetent. Nonetheless we should be clear about what is hidden in that phrase "ever since Darwin". Even though the idea of natural selection can now be summarised and explained in a few sentences, it did not in fact spring into existence all at once, nor from one source. The progress of ideas is almost never like that. In reality so-called scientific revolutions are no more punctuated and discontinuous than is biological evolution itself: zoom in on the details and you will always find all sorts of false starts and in-between, half-baked stages on the way to the eventual solution. Darwin himself

struggled over many years with his version of the theory. And even after its publication in 1859, the argument quite properly went on. Several key bits of the theory were not yet in place. The facts were not yet - at least not yet known to be - as they would have to be.

Now, in the case of consciousness, I see us as being at a very comparable stage of creation, confusion and backsliding, such as existed in the case of evolution in the mid-nineteenth century. This is the time when another BIG IDEA - the idea of how to link the conscious mind and the material body - is in the very process of being thought-through, reworked and edited (although, in this case, since the age of Victorian Genius has passed, it is not being done by just one or two great men but by a whole community of us). I am optimistic enough to believe, as Dennett does, that in a hundred years time, the problem of consciousness will indeed have been finally resolved - so that our descendants may even be able to write the answer on a postcard. But (and here Dennett and I part company) I cannot say that we are there yet. Not quite. And meanwhile it's difficult. And worrying. And it hurts. And it sometimes seems we're getting nowhere - though in retrospect it will be obvious that we were.

My target paper in this issue of the Journal, and the Commentaries on it, do, I hope, make a contribution to the job in hand. The fact that every Commentator has a different view of what's right and what's wrong with the paper (with one Commentator's view of what's wrong being precisely another's of what's right) at least means we are in the zone of uncertainty where interesting things happen.

I wanted to keep the target paper short, and I justified this to myself by imagining that a good many readers might already have some familiarity with the 1992 book in which I treated several of the issues discussed here at much greater length (Humphrey, 1992). Or at any rate, I hoped readers might take the trouble to refer back to it. But I can see now that these assumptions were unrealistic, and that by failing to elaborate at several points within the paper I invited too much second-guessing as to what it is I meant (or should have meant): "If I understand him" (when in fact she hasn't), "Humphrey needs to provide more on this in future" (when in fact he has provided considerably more in the not so distant past), and so on.

One solution, now, might be to try to make amends for leaving so much in abeyance by quoting major chunks from *A History of the Mind* in this reply. But, as luck would have it, the need for this has been partly lifted by one of the existing Commentators who, as if foreseeing my difficulty, has seen fit to provide a blow-by-blow account of how he himself read the argument of my earlier book. In fact Christian de Quincey's sensitive and detailed summary does the job so much better than I could have done myself, that I can only recommend to

anyone who finds that my target paper moves to fast or too far for them that they should turn immediately to de Quincey for on-line Help.

Still, I am well aware that some of the Commentators have found faults with the paper that go far beyond any frustration they may have felt with simply not being able to see what I am getting at. Indeed some have clearly seen precisely what I am getting at - and, the more they have seen, the less they have been inclined to agree with it. In what follows I shall address, in turn, four areas of serious disagreement. For reasons of space, I shall pass over the many other areas where Commentators have had helpful and constructive things to say, because in these areas the Commentaries speak for themselves.

I should confess however that, even as I begin this, I have a sinking feeling. The target paper was short because I wanted, for once, to tell the story quickly and simply without being obliged to fend off every well-worn (and not so well-worn) objection. But now, here I am, about to engage in some of the very arguments I hoped I could ride over. In the end, I would rather people remembered the target paper than this reply.

## **1. Sensation and perception.**

I say that when we talk about the central fact of consciousness - namely, the fact that it is *like something* to have certain experiences - we are (if we but knew it) talking always about the subjective experience of sensation. As I summed this up before (and I am, after all, going to quote here from the book, p. 115):

(1) To be conscious is essentially to have sensations: that is, to have affect-laden mental representations of something happening here and now to me.

(2) The subject of consciousness, "I", is an embodied self. In the absence of bodily sensations "I" would cease. *Sentio, ergo sum* - I feel therefore I am.

(3) All sensations are implicitly located at the spatial boundary between me and not-me, and at the temporal boundary between past and future: that is, in the "present".

Hence, if we are to pin down precisely what *kind* of thing phenomenal consciousness, we must pin down what *kind* of thing sensation is. And this means, I claim, that we must take pains from the outset to recognise the essential difference between sensation and perception. Sensation has evolved specifically to represent and evaluate the current state of stimulation at

the surface of the body, whereas perception has evolved to represent the objective facts about the world beyond. If anything is important to understanding consciousness *this distinction* is.

Now, the absolute separation between two things that are usually run together is, I know, not easy to come to terms with - whether at the level of personal phenomenology or at the level of the underlying psychology and neuroscience (if it were easy, it would not have needed the first fifteen chapters of my book to make the case for it). But I argue that this separation is supported - and indeed demanded - by a variety of lines of argument and evidence which together confirm that despite the fact that both sensation and perception rely on the same sense organs, the central channels that process them have developed along separate lines in evolution and still today are functionally and even neuroanatomically discrete.

Let's consider in some detail, since it seems to be the area where I am most often challenged, the case of *colour vision*. "Seeing colour", according to this theory, usually consists in having not one kind of experience but two. From the moment light, reflected from an external object, arrives at the human eye, there are two kinds of analysis that get under way: one track - the sensory one - leads to a representation of the distribution, intensity and quality of light at the eye, whereas the other track - the perceptual one - leads to a representation of the reflectance properties of the object that gave rise to this light. The colour of the light can be designated as colour with a small c, and the surface Colour of the object as Colour with a big C.

What are the grounds, then, for the claim that colour sensation and Colour perception are separate and parallel processes (rather than being, as Robert van Gulick suggests they might equally be, all part of a single serial production)? I think the case can be argued at two levels, the analytical and the empirical. But it should be enough if I cite here some of the empirical evidence that the two kinds of representation can function independently.

The most dramatic evidence comes from cases of brain damage, where there can be a double dissociation between colour sensation and Colour perception. Thus, on the one hand, in some cases of colour agnosia, sensation is unaffected (so that the subject still experiences the full richness of what it's like to have colourful stimulation of his visual fields), while at the same time Colour perception is lost (so that he is no longer able to judge what the Colour of any particular object is) (Oxbury et al., 1969). On the other hand, in cases of blindsight, Colour perception may be spared (so that the subject can still guess the Colour of an external surface), while colour sensation is lost (so that he now says it's not like anything to have stimulation of his visual fields, and indeed that he is no longer conscious of seeing anything at all) (Stoerig and Cowey, 1997).

But even when the brain is working normally, there is plenty to indicate that colour sensation and Colour perception are relatively separate operations. For example the time courses are different, with sensation leading perception by a considerable margin. We have all experienced this when, say, the lights go up in a dark room full of coloured books, curtains, rugs. All at once we have the experience of a field full of colour sensation, but - as reaction time experiments show - it may be all of several seconds before we can bring the results of Colour perception to bear on identifying the Colour of a target object. With psychotropic drugs, this divergence can be amplified, with sensation overwhelming perception altogether. Thus, as Aldous Huxley described it: “At ordinary times the eye concerns itself with such problems as *Where? - How far? - How situated in relation to what?* In the mescaline experience the implied questions to which the eye responds are of another order. Place and distance cease to be of much interest. The mind does its perceiving in terms of intensity of existence...” (Huxley, 1954). But, on the other hand, there are situations where we perceive Colour without sensing colour at all, as for example in the phenomenon that Michotte drew attention to of the “amodal completion of perceptual structure”, when we perceive the Colour of the surface of an object which is temporarily covered by a black mask passing in front of it (Michotte et al. 1964/1991; Natsoulas, 1999).

Further remarkable evidence comes from young children. It is well known that children are surprisingly slow to learn Colour names, so that long after they demonstrably respond attentionally to coloured light and show colour preferences, they still cannot, for example, label a tomato as Red or grass as Green. But even once Colour perception is firmly established, and the Colour words are there, it seems that the child may still not yet have put two and two together and have realised that she generally perceives Colour in the external world at the same time she senses colour at her eyes - with the result that there’s a stage at which she simply does not appreciate that Colour perception is mediated by eyesight. Put a green squishy ball in a three-year-old child’s hand and ask her what Colour it is, and she will look at it and say Green, or ask her whether it is hard or soft, and she will squeeze it and say Soft; but now put it in a bag and ask her what she would *have to do* either to find out what its Colour is or to find out whether it is Hard or Soft - would she have to put her hand in and feel it or would she have to take a look? - and she’ll likely say she does not know the answer (O’Neill et al., 1992). It’s as if the contingent association between perception and sensation (the contingency that as adults we find it so hard not to regard as *necessary*), has for the toddler simply not yet been figured out.

Yet I realise that you may hear all this evidence, and still find it hard to accept the story as I tell it. Among the Commentators, Robert van Gulick and Valerie Hardcastle are the

most overtly suspicious of the validity of the distinction between sensation and perception. But it is clear from the scattered remarks of several others that they too either do not really buy it (even my best ally, Dennett, will only concede that “something like” this distinction needs to be drawn), or at any rate that they have not understood just how strictly I think it needs to be applied.

In particular, people continue to have trouble with my insistence that the right way to describe it is as a distinction between the representation of “what is happening to me” and of “what is happening out there”. Even if they can see how this may be an appropriate description in the case of the proximal senses such as touch or taste, they cannot see how it can be appropriate in the case of the distal senses of vision and audition. And almost everyone seems to balk at the idea - which remains absolutely central to my account - that *visual sensations* really are to do with representing “what’s happening *to me at my eyes*”. Van Gulick, for example, protests that when he looks at a red Coke can on the table, it seems to him he experiences the phenomenal colour “as a feature of the can out there on the table” not as something happening to himself; and Hardcastle baldly states “we don’t feel redly about parts of our visual field . . . we project our visual sensations as something external to us.”

I agree, it usually *seems* like that to me as well. But we have to pursue this example further. Let’s suppose that van Gulick now walks towards the table and brings his face close up to the Coke can, so that the can’s image fills more and more of his visual field, until he is so close that all he sees is a red blur. The sensation is undergoing a major transformation. But while this is happening, is he perceiving any change in the “features of the can out there on the table”? No, the change in sensation represents a change not in the can out there but only in the *image* of the can. And whose image is this, if not *van Gulick’s* image? And where is it located, if not *at van Gulick’s eyes*?

Whenever people come back to me with this objection, and repeat that it just doesn’t *seem* to be the way I’m telling it: that visual sensations don’t *seem* to be located at the eye (or auditory sensations at the ear, or olfactory sensations at the nose, etc), I am inclined to ask them: So, what precisely would it seem like if they *were* located at eye (or the ear or the nose etc)? In the case of the eye, wouldn’t you expect, for example, that the sensation of a coloured patch would in fact seem to grow larger when the eye got closer to the light source? Or that the sensation would shift when you pressed with a finger on the eyeball? Or that it would change colour when you donned dark glasses? and so on. In other words, wouldn’t you expect it to be just like it is, like this!

## 2. “Agentic qualia”

My view, as expressed in this paper, about the limits of sensation (and of consciousness) is a purist one: namely that sensation always has to do with representing what's happening at the boundary between me and not me - which is to say at the body surface as mapped by sense organs. To continue the litany, from where I summarized things in my book (p. 116):

(4) For human beings, most sensations occur in the province of one of the five senses (sight, sound, touch, smell, taste). Hence most human states of consciousness have one or other of these qualities. There are no non-sensory, amodal conscious states.

(5) Mental activities other than those involving direct sensation enter consciousness only in so far as they are accompanied by "reminders" of sensation, such as happens in the case of mental imagery and dreams.

(6) This is no less true of conscious thoughts, ideas, beliefs, [perception]. . . Conscious thoughts are typically "heard" as images of voices in the head - and without this sensory component they would drop away.

However here I am ready to make concessions. I was already becoming unsure of how far I really wanted to insist on this degree of purity even at the time I wrote this paper (and indeed the "*most* sensations . . ." and "*most* human states of consciousness . . .", in (4) above, indicates that I was hedging on it even earlier). And, now, both Ralph Ellis and Natika Newton, in their commentaries, help me see more clearly what is wrong with ruling out the possibility of there being *any* source of phenomenal consciousness other than conventional sensation.

What Ellis and Newton draw attention to is that there may in fact be a phenomenology of *action per se*. My claim in the paper is that, while it is "like something" to have sensations, it is not like anything much to engage in most other bodily activities: "To say the least, our experience of other bodily activities is usually very much shallower. When I wave my hand there may be, perhaps, the ghost of some phenomenal experience. But surely what it's like to wave hardly compares with what it's like to feel pain, or taste salt or sense red." Yet Ellis and Newton both say that this is wrong, because in reality it always does feel like something to be acting (or, as we shall see, even preparing to act). And I would now agree with them that it is at least partly wrong - the ghost is more substantial than I allowed!

I refer readers to the good discussion of the phenomenology of action that these authors (and Naomi Eilan too) provide. But I have another reason for wanting to pursue the issue here: which is so that I can confront - and I hope deal with - a problem that no Commentator has actually brought up on this occasion, but that I have had raised with me by

others (especially John Searle). It's a problem that might otherwise prove to be the Achilles heel of the whole thesis about consciousness and sensations. (Some of what follows is from Humphrey, 2000b).

Suppose it were indeed true, as I maintained in a strong form above, that the entire content of consciousness is made up of bodily sensations, with nothing being contributed by perceptions or thoughts as such. It would follow presumably that a person will not experience any change in consciousness *unless* there is a change in sensation, even if there *is* a change in what is perceived or thought. But then consider what this means in the case, say, of vision. It has to mean that when someone is looking at a scene, he should experience no change in consciousness unless and until the visual stimulus as such changes (so as to create a change in "what's happening to me"), even if he does come to perceive it differently in terms of what it represents as "what's happening out there". And, in that case, a crucial test would be provided by one of those notorious ambiguous pictures, such as the Necker cube or the duck/rabbit, where, even though there is no change in the visual image, the perception of what it represents can indeed radically alter.

At a recent conference Searle challenged me directly about such cases, saying that, for example, it was perfectly obvious to him that the Necker cube seen in one way really is a consciously different phenomenon from the same cube seen the other way — thus proving, in contradiction to my own position, that we can in fact be conscious of what is perceived as well as of what is sensed.

Well, can we? . . . I admit on the evidence of introspective observation the answer must be: Yes. When the cube reverses in depth there is surely *something* that consciously changes. And it certainly is not at the level of sensation of the visual image (which, *qua* represented image, does not even have a depth dimension to reverse).

What is this something, then? When the cube reverses, is there, as Searle would want to say, a change in some aspect of *non-sensory* perceptual consciousness — perhaps the coming and going of "cube qualia"? (Just as, with the duck/rabbit, he might want to postulate the coming or going of "duck qualia" or "rabbit qualia"? I am not joking: some theorists are really prepared to talk this way.) If so, my argument is lost.

Or, is there perhaps another possibility (as would surely be suggested by Eilan, Ellis, and Newton)? When the cube reverses, is there a change not in non-sensory qualia nor in visual qualia but in *sensory qualia of another non-visual kind*? I believe there are in fact two ways that this could be happening.

One way of understanding it would be to take up an idea of the psychoanalyst Mark Solms (which he himself attributes to Freud), and to suggest that conscious experience is comprised not only of the five basic modalities of sensory qualia but also of an additional



dimension of *affect*. “Affective qualia”, Solms writes, “(which are calibrated in degrees of pleasure / unpleasure) are wholly equivalent to the qualia of vision, hearing, smell etc. and are irreducible to them”(Solms, 1997, p. 773). So, whenever we experience a sight or a sound or a taste etc., perhaps the conscious experience is likely to consist *both* of the specific sensory qualia *and* of whatever affective qualia are being activated. But, while the sensory qualia are fixed solely by the sensory stimulus, the affective qualia may be influenced not only by the stimulus but also by what is being perceived. With an ambiguous figure, then, even though the visual sensation remains constant, when the perception changes the affective qualia may change too.

I think this is a nice idea, and in some cases it might be correct. But I am not sure it will do in general. Different affective feelings for ducks and rabbits? Well, why not. But different feelings for the two versions of the Necker cube? Unlikely.

The other way would be to take up the idea that Eilan, Ellis and Newton all hint at in their commentaries (and which I have in fact toyed with myself in earlier writings), and to suggest that what is crucial is not so much affect as *action*. Suppose that whenever we perceive anything (and sometimes even when we merely think of things) we always implicitly formulate a plan of action — for example a plan to reach out and take hold of it. And suppose that such action, even when implicit, always has a small but noticeable qualitative feel to it — either on its own account via somatic sensation or through modeling of the sensory feedback that would be expected. Let’s call this additional dimension the dimension of “agentic qualia” (Humphrey, 2000b). Then, whenever we experience a sight or a sound or a taste, etc., the conscious experience can be expected to consist not only of the sensory qualia appropriate to the particular sensation but also of whatever agentic qualia are being called into being by the perceived “affordance for action” (in J. J. Gibson’s terms).

This solves the problem of the Necker cube. For, now we can postulate that, even while the visual sensation remains constant, there may be a covert change in action plan when the perception of the cube reverses, and so a slight change in the overall sensory qualia. And of course it also solves much else that might otherwise be puzzling.

The admission of a realm of agentic qualia makes the story I have been telling considerably more complicated. But that’s a good price to pay for making it more likely to be right.

One particular area, I might mention, in which it makes it more likely to be right - and indeed makes it possible to tell a story at all - is in relation to the much-disputed phenomenology of blindsight. Suppose that, despite the fact a person with blindsight has no consciousness of visual sensations in the blind field, he nonetheless does experience it as “like something”, consciously, to detect an object in the blind field (and several reports suggest that

in a strange way it *may* be so) - then perhaps the explanation is that what he is experiencing are the agentic qualia associated with his having an incipient plan to grasp the object.

### 3. “Mongrel” phenomena

In Block’s 1994 paper where he attempts to distinguish two concepts of consciousness - “phenomenal” and “access” consciousness - he remarks that the term consciousness as used in common parlance is a “mongrel concept” Block (1995). Yet his choice of the word “mongrel” for this is not quite right. A true mongrel concept would surely be one where two different parent concepts have been combined in one offspring, which as a result essentially has parts of both. But what Block has in mind is something less than this. He wants to say that the term consciousness typically bundles together several different concepts in a way that makes for problems in telling which particular concept is being referred to; but he does not want to say that these different concepts essentially belong together as a union.

Now, as Eilan points out, Block’s distinction between phenomenal and access consciousness has an obvious affinity to my own distinction between sensation and perception (in fact, as I myself have said, it often amounts to being the very same distinction (Humphrey, 1995)). And there is a parallel to it in the way I, like Block, have seen it as the first task of a theory of consciousness to unbundle and lay out for separate view these two relatively autonomous concepts. Yet at various times I admit I have entertained a different thought: namely, that consciousness really might, after all, be a phenomenon that truly is a mongrel - such that you have to have *both* sensation *and* perception united in a single mental state to yield the real thing.

This same thought has clearly occurred to several of the Commentators (in particular Eilan), who variously suggest ways in which sensation and perception may intimately depend on each other. But, in the end I am not persuaded either by my own moves in this direction or theirs. I do believe that sensation can help to make perception a success for the perceiver; sensation can even be what makes us think (wrongly) of perception as having its own conscious phenomenology; but in the final analysis the two processes remain two not one.

Still, let me at least soften my position, and hold a hand out in Eilan’s direction. For I would say she has identified what is the most important way in which sensation can contribute to perception, which is by adding the crucial element of *presentness*. Indeed I (Humphrey, 1992), Anthony Marcel (1988) and Richard Gregory (1996) have all, in different ways,

proposed the same idea: namely, that sensations are required, in Gregory's felicitous wording, "to flag the present".

The basic idea here is that one of the main (and, in evolution, ongoing) roles that sensation plays is, as I put it in the paper, to "police" perception - or to help "keep perception honest" (Humphrey, 2000a).

The reasoning is as follows. Both sensation and perception, as noted above, take sensory stimulation as their starting point: yet, while sensation then proceeds to represent the stimulation more or less as given, perception takes off in a much more complex and risky way. Perception has to combine the evidence of stimulation with contextual information, memory and rules so as to construct a hypothetical model of the external world as it exists independently of the observer. Yet the danger is that, if this kind of construction is allowed simply to run free, without being continually tied into present-tense reality, the perceiver may become lost in a world of hypotheticals and counterfactuals.

So, what the perceiver needs is the capacity to run some kind of on-line reality check, testing his perceptual model for its currency and relevance, and in particular keeping tabs on where he himself now stands. But this, so the argument goes, is in fact precisely where low level, unprocessed, sensation does prove its value. As I summarized it earlier: "Sensation lends a here-ness and a now-ness and a me-ness to the experience of the world, of which pure perception in the absence of sensation is bereft" (Humphrey, 1992, p.73).

Thus, here I am agreeing with Eilan. At the very least sensation is needed to establish the present-tense credentials of perception - and thereby, as it were, to license perception for use as a trustworthy representation of what's really out there *at this moment*. In fact, as Eilan might have hoped, I have long argued for the idea of *presence* as the defining phenomenon of conscious life, even to the extent of my developing the pun implicit in the word: "The very word 'present' comes from the Latin *prae-sens*. *Prae* means 'in front of' and *sens* is the present participle of *sum* ('I am'). But *sens* is also the root of the past participle of *sentio* ('I feel'). Thus *sens* hovers ambiguously between 'being' and 'feeling, and *prae-sens* carries the implication of 'in front of a feeling being'. Correspondingly the subjective present is comprised by what a person feels happening to him; and when he ceases having sensations - as when he enters dreamless sleep or dies - his present ends" (Humphrey 1992, p.99).

Even so, I will not, on this basis, go all the way with Eilan or anyone else to concluding that perception and sensation belong together as a single package. The reason is first, though I say it again, that perception and sensation are *about* different kinds of thing - so that, when it comes to it, the putative mongrel would not be the kind of within-category mongrel (say a spaniel-collie) that at least in principle makes sense, but a cross-category one

(say a spaniel-tulip) that does not. And second, and in the end more telling, the reason is the one already described above, which is that in reality sometimes we can and sometimes we do experience sensation and perception independently. Perception without sensation, as in blindsight, is rare and certainly not normal - indeed it lacks precisely that *presence*, the hereness and nowness and me-ness, that sensation usually lends it. But visual perception in blindsight still works at the level of representing “what’s happening out there”.

If anyone who knows about human cases of blindsight should doubt me about this - perhaps because they are struck by the lack of visual spontaneity that human patients with partial blindsight show - I would refer them to my own study of a monkey, known as Helen, with complete blindsight after surgical ablation of the entire primary visual cortex. Helen’s visually guided behaviour recovered to such a degree that anyone who observed her freely using her eyes to navigate through her environment would have assumed (quite rightly in my view) that her visual perception was almost back to normal (Humphrey, 1975; and see my discussion in 1992, pp. 88-93).

There is something else that the human blindsight patient lacks (and the monkey too, presumably), that is highly relevant to this discussion: when the patient engages in visual perception he does not experience it as *vision*. Instead he typically says that the experience has no modality at all. (Or, he may even become confused. A patient with the analogous syndrome in the tactile sphere - “unfeeling touch” - while having no conscious tactile sensations, could nonetheless identify where he was being touched on his arm: but, when asked how he could do it, he would variously say he smelt or heard the stimulus (Paillard et al. 1983)!)

In the monkey’s case I had direct evidence that she too did not immediately recognise her capacity to use her eyes as a case of *vision*, when I found that early on her way to full recovery she actually muddled up lights with sounds. Thus, for example, after she had been trained to reach to touch a flashing light presented in the dark on the end of a stick, she immediately transferred to reaching for a clicking speaker on a stick, and it was only with difficulty that I could teach her to respond to the light and *not* the sound; what’s more, when given a light and a speaker placed close together she sometimes reached between them (Humphrey 1967; 1995).

This lack of awareness of the modality of a perceptual channel is just what we should expect of perception without sensation. But of course it draws attention to a further role that sensation normally plays in helping make perception a success. *Sensation tells the perceiver in what manner he is doing his perceiving*. In fact I might well have added a further “-ness” to the list above: “Sensation lends a here-ness, a now-ness, a me-ness *and a modalness* to the experience of the world, of which pure perception in the absence of sensation is bereft”.

Yet, I am afraid Andy Clark will not like this. For in his Commentary he addresses this very same issue of how the modality of perceptions gets to be appreciated - and comes up with an entirely different answer. His own proposal is that the “[perceiving] agent has direct unmediated access to distinctive non-phenomenal properties of the act of detection itself. Where such access is available, the agent must judge there to be a difference in what it is like to gather information by sight rather than by e.g. hearing.” So, according to Clark, the modality of perceptions comes, in effect, free with the territory: the unique style of information processing involved in representing sights (as against sounds, say), and/or the types of affordances that sights (as against sounds) offer, intrinsically mark out visual perception as being qualitatively different from auditory perception - and no recourse to the modalness of accompanying sensations is necessary.

Clark says he is putting forward this suggestion in the context of my paper in order to be helpful to me. He thinks my theory is in trouble without it, because he does not see how modal quality - whether in the case of perceptions or sensations - can arise in any other way than the kind of way that he suggests. I am sorry if I seem ungrateful, in saying Thanks but No Thanks for this offer of help. But I reckon I not only can but must manage without it. For I have to say I think his theory, as a theory of how perception gets to be tagged intrinsically as belonging to a particular modality, just cannot be right - for the simple reason that perception is *not* tagged intrinsically with a modality. Indeed, as we have seen, the reality is the reverse: perception is intrinsically *amodal*.

If Clark were right, several of the phenomena of amodal perception discussed above could simply never occur. It would be impossible for a person (or a monkey) with blindsight to perceive the shape or location of an object *without* appreciating the visualness of the experience. It would be impossible for a child to discover the Colour of a ball, *without* realising that she had used her eyes to do so. Or, to give one more remarkable example, it would be impossible for someone to use tactile stimulation on the skin as a basis for “seeing” *without* this touch-driven experience taking on the phenomenal quality of vision. But in the case of such so-called “skin vision”, where someone has an optical image from a head-mounted camera translated into a pattern of tactile stimulation on the skin of their back, and proceeds to learn to use this pattern to “see” with, in fact the sensory experience remains firmly tactile (Bach-y-Rita, 1972).

Actually, I do not want to dismiss Clark’s theory altogether. Because I believe that, with a little negotiation, he and I would find that, after all, we are basically thinking on similar lines. It is just that Clark is using his theory to bark up the wrong tree. The place to which he ought to be applying this idea of there being direct unmediated access to phenomenal qualities, is not perception (which is not intrinsically modal) but sensation (which *is*). And this

is precisely where I myself - although Clark has clearly not read me this way - have tried to apply a very similar idea: namely, that phenomenal quality, and indeed phenomenal consciousness as a whole, is the direct and unmediated outcome of a certain, modality specific, *way of doing things*.

#### 4. Functionalism and Zombie-Free Zones

Van Gulick takes me to task for saying in the paper: “No one it seems has the least idea how to characterize the phenomenal experience of redness in functional terms. . . in fact there are well-known arguments (such as the Inverted Spectrum) that purport to prove that it cannot be done, even in principle.” “Although [Humphrey] does not explicitly endorse these arguments,” van Gulick writes, “he has nothing to say about the functional strategy after this apparent dismissal.” Hardcastle clearly sees me not as a functionalist but as a physicalist (or why else should she go on about whether my theory accords with the *anatomy*?). Meanwhile Clark, trying to see the brighter side, says he thinks that everything I say is in fact “compatible with the claim being pitched at a functional, rather than a brute physical level”.

But I am amazed. I took it for granted that everyone would recognise that my account of sensations was indeed meant to be a functional one through and through - so much so that I actually deleted the following sentences from an earlier draft of the paper, believing them redundant: “Thus [with this account] we are well on our way to doing the very thing it *seemed* we would not be able to do, namely giving the mind term of the identity, the phantasm, a *functional description* — even if a rather unexpected and peculiar one. And, as we have already seen, once we have a functional description we’re home and dry, because the same description can quite well fit a brain state.”

But perhaps I should not be amazed. Functionalism is a wonderfully - even absurdly - bold hypothesis, about which few of us are entirely comfortable. Bertrand Russell famously said: “The method of ‘postulating’ what we want has many advantages; they are the same as the advantages of theft over honest toil” (Russell, 1919). And it’s only natural to wonder whether functionalism is theft, whereas only physicalism represents honest toil.

Even the most ardent functionalists (among whom I would include myself) seem sometimes not to appreciate just what it is their metaphysical position commits them to. This is that, if and when you’ve provided a functional account of everything you’ve been asked to account for - let’s say everything that distinguishes a person who is having a red sensation - you’ve done the job and you can stop. There is no explanatory residue. At that point the

objections are (like Monty Python's parrot) dead, kaput, ex-objections, objections that have gone to another place.

Michael Frayn stated the case for functionalism plainly in his novel, *The Tin Men* (Frayn, 1965). In this story, Macintosh, who is master of computers, is programming his computers to pray — or to engage in “automated devotion”, as he calls it. But a sceptical colleague objects that the difference is that a man who prays would *mean* it. Macintosh replies: “So does the computer. Or at any rate, it would take a damned complicated computer to say the words without meaning them. I mean, what do we mean by ‘mean’? If we want to know whether a man or a computer means ‘O Lord, bless the Queen and her Ministers’, we look to see whether it’s grinning insincerely or ironically as it says the words. We try to find out whether it belongs to the Communist party. We observe whether it simultaneously passes notes about lunch or fornication. If it passes all the tests of this sort, what other tests are there for telling if it means what it says?”

Now, I would agree that the problem of prayer is arguably an “easy problem”, compared to the “hard problem” of seeing red. And I am not claiming that we yet have a functional account of sensation that passes *all* the tests. But, arguably, we do have an account that already passes *most* of the tests and that looks as if it is at least *on course* to pass the rest of them. Yet, among my colleagues commenting on this paper, it seems that the majority instead of declaring at least provisional success, cannot but have fits of self doubt. As Robert Pirsig wrote in *Zen and the Art of Motorcycle Maintenance*: “The truth comes knocking on the door. And you say ‘Go away. I’m looking for the truth’. And so it goes away” (Pirsig, 1974).

It’s true that I have Newton on board with me. And Ellis and de Quincey are at least ready to sail alongside. But Clark is disappointing in this regard. Although, as I’ve indicated, I count him a friend to my approach, I’m dismayed to find him claiming that one of the key ideas - the idea that the subject is the “author” of his own sensations - cannot deliver what I want, while Clark completely mistakes how I am supposing that this works.

“The mere idea of knowledge of authorship,” Clark writes, “fails to illuminate the question of phenomenal feel. . . Why should knowledge of authorship not be as free of sensational depth and character as, say, knowledge that Paris is the capital of France?” But the idea of “knowledge” of authorship plays no role in my theory. Sensing does not involve “knowledge” of sensing, any more than willing involves “knowledge” of willing. In my view, what makes for phenomenal feel is *authorship as such*. To feel just *is* to be the author of a certain kind of response - and that’s how and why the properties are self-disclosing.

I like Clark's notion of a "necessarily zombie free zone, a zone where facts about access imply (but do not assume) a difference in how things are sensorially given." But I hoped he would see that my theory as it stands (without his revisions) already has the potential to lead into this zone.

Let me make a related point about authorship in reply to Hardcastle, who complains that she needs "a way to distinguish, from the brain's point of view, conscious experience from the rest." Her implicit suggestion that the brain *has* a point of view is of course a rather strange and interesting one. But again I hoped she would see that this is something I have already provided for, by making authorship central to the theory. Minds *and* brains can both in principle be authors - and authors do intrinsically have points of view.

I have to admit that, as van Gulick says, with this part of the theory there is a big I.O.U. outstanding. The idea of authorship and all that follows from it still requires technical explication, far beyond what I have yet provided. But I did at least make a start on it in my book, by trying to develop the idea of "instructions" and recursive intentionality (see de Quincey's summary). Ten years later I suspect there are useful insights to be gleaned from the newly invented field of "consumer semantics" according to which "the intentional content of a state depends, at least in part, on what the down-stream consumer systems which can make use of that state are disposed to do with it" (Carruthers, 2000) - which is just my earlier point. And I'm sure the other way to go with it, as Ellis persuasively argues in his Commentary, will be to borrow ideas from dynamical systems theory to explain the very nature of agency.

It will not, now, be me who does it. But some day soon I have no doubt that others with superior technical skills will be able to set this notion of an "author" on a solid analytic footing. It will then take its due place as a key - probably *the* key - "dual currency concept" in the philosophy of mind. Indeed perhaps it will turn out be that oh-so-obvious big idea that future generations will be able to write down on a postcard - and to flourish in front of their benighted grandparents who once struggled to see how anything could be both mental *and* physical, and who lived in fear of "absent qualia" and "zombies".

So, seriously, what about those zombies? There's no question that the anxiety, that Dennett (2000) calls the "zombic hunch", continues to haunt the field. In fact it is clearly still there in the background of many if not all the present Commentaries. Even those who are trying to kick the habit cannot, it seems, quite leave zombies alone.

Van Gulick is open about his continuing concern. Having declared himself - after some appropriate and helpful reservations - largely in favour of the approach I've taken, nonetheless he all of a sudden raises the spectre of "absent qualia". "Couldn't one easily imagine systems [that had all the specified features] but lacked any phenomenal awareness,



any inner what's-it-like-to-be-ness. . . Simply imagine a robot that generates such responses to stimuli and monitors them. Such a system would seem to satisfy Humphrey's conditions, but that in itself would give us little reason to regard it as phenomenally conscious".

Van Gulick is of course entitled to question, at a scientific level, whether the specific features of my model are the right ones, so that implementing *those* features in a robot would make the robot conscious. (Although, be it said, as Dennett points out here, and as I discuss at some length in the book, that my model provides no simple blue-print for wiring up consciousness *de novo*: it is a model of how the functional pathways *developed in the course of evolution*, and no one is going to be able to implement these details without retracing much of the evolutionary history). All the same, I do not think I am misreading van Gulick when I say I detect, in his statement "perhaps were I to grasp [the conditions] more clearly I would see how they entail phenomenal consciousness", not so much a judicious withholding of judgment as a reluctance to believe that anything I or anyone else could offer would *ever* do the trick.

At a different level, I guess de Quincey is a closet zombist too. He goes further than van Gulick in declaring his support for what I'm doing (in fact he goes further than anyone, bar Dennett), but then he too suddenly jumps ship. He says that, at the last moment in my evolutionary story, I have helped myself to the idea of the "subjective present", when nothing in the functional account of subjectivity that I was developing entailed it. The implication is that we could have everything that my model does entail, without there being a phenomenal "present". But, since as we all know subjectively, the fact is there *is* a phenomenal present, this means that the potential for presentness was already latent in the design of the universe and has to be considered primitive - at least as a primitive *potential*. And for de Quincey (rather as for Chalmers) this clearly has to be a contingent fact about the universe, not a necessary one: logically it could have been otherwise (Van Gulick is right to correct me about what I said about Chalmers' position in the paper).

I do not know how to respond to this: except to say okay, but then show me how there could be the functional states in place but *not* the presentness. The idea of there being unrealised *potential* in the universe is of course basic to functionalist metaphysics. Thus we all agree there must have been, for example, the potential for there being "square roots" and "justice" and the "Journal of Consciousness Studies", and so on, long before the requisite functional states were ever realised. But, once they *were* realised, then surely these things had to exist necessarily. And the same for phenomenal consciousness and the subjective present. There's nothing contingent about it. I and you are conscious, not contingently, but *necessarily*.

Sometimes, I confess, that when faced with the only-up-to-a-point functionalism of van Gulick or de Quincey or most of my colleagues, I am almost relieved to come across the unrepentant in-your-face *anti-functionalism* of Stevan Harnad. At least with Harnad you know where you are. The sarcasm is transparent: “I can design and implement recursive self-sustaining loops fitting Humphrey’s description easily. Do they quicken with the light of consciousness?” No of course not, because: “If we characterise feelings computationally or functionally, we have simply begged the question and changed the subject”. Functionalism for Harnad is by definition “zombie functionalism”.

But, if I am glad to have Harnad state the enemy’s case so boisterously and scornfully as he does in his Commentary, it’s only because he thereby reveals the ultimate vacuity of his position. It goes nowhere. It makes no predictions. It generates no tests. Indeed, for Harnad it would actually be an argument against the legitimacy of any theory of consciousness that someone should even imagine that his theory could be *tested* by implementing the consciousness-producing architecture *in a machine*. Because if he were to interpret anything the machine actually *does* with its new architecture (anything at all) as evidence that the implementation has been successful, that would only show that his theory begged the question.

It’s as though Harnad has managed to turn Tertullian’s grand claim “I believe because it is impossible” into its corollary “I do not believe because it *is* possible”.

Of course this was also the ultimate argument used against Darwin. It may be true, the churchmen said, that it would have been *possible* for the living world to have been designed by natural selection. But don’t be fooled. God has arranged things to appear *as-if* designed by natural selection, just so as to test your faith in the fact that they have *not really* been so designed.

A hundred and fifty years later no one can be bothered with such sophistry. And there’s a moral there.

## REFERENCES

Bach-y-Rita, P. (1972), *Brain Mechanisms in Sensory Substitution*, (London: Academic Press).

Block, N. (1995), ‘On a confusion about a function of consciousness’, *Behavioral and Brain Sciences*, **18**, 227-247.

Carruthers, P. (2000), ‘Replies to critics: Explaining subjectivity’, *Psyche*, **6**, (3).

Dennett, D.C. (2000), 'The Zombic Hunch: Extinction of an Intuition?', *Philosophy*, (in press).

Frayn, M. (1965), *The Tin Men* (Boston: Little, Brown)

Gregory, R. L. (1996), 'What do qualia do?' *Perception*, **25**, 377-8.

Humphrey, N. K. (1968), *Two Studies in the Neuropsychology of Vision*, Ph.D. thesis, University of Cambridge.

Humphrey, N. K. (1974), 'Vision in a monkey without striate cortex: a case study', *Perception*, **3**, 241-255.

Humphrey, N. (1992), *A History of the Mind*, (London: Chatto & Windus).

Humphrey, N. (1995), 'Blocking out the distinction between sensation and perception: superblindsight and the case of Helen', *Behavioral and Brain Sciences*, **18**, 257-8.

Humphrey, N. (2000a), 'The privatization of sensation', in *The Evolution of Cognition*, ed. C. Heyes & L. Huber (Cambridge Ma: MIT Press), in press.

Humphrey, N. (2000b), 'Now you see it, now you don't', [commentary on Crick and Koch, 2000], *Neuro-psychoanalysis*, **2**, in press.

Huxley, A. (1954), *The Doors of Perception*, (New York: Harper and Row) p. 20.

Marcel, A. J. (1988), 'Phenomenal experience and functionalism', in *Consciousness in Contemporary Science*, ed. A.J. Marcel, and E. Bisiach (Oxford: Clarendon Press) pp. 121-158.

Michotte, A., Thinés, G., Crabbé, G. (1964/1991), 'Amodal completion of perceptual structures', in *Michotte's Experimental Phenomenology of Perception*, ed. G. Thinés, A. Costall, G. Butterworth (Hillsdale, NJ: Erlbaum) pp. 140-67.

Natsoulas, T. (1999), 'A rediscovery of presence', *Journal of Mind and Behavior*, **20**, 17-42.

O'Neill, D. K., Astington, J. W., Flavell, J. H. (1992), 'Young children's understanding of the role sensory experience plays in knowledge acquisition', *Child Development*, **63**, 474-90.

Oxbury, J. M., Oxbury, S. M., Humphrey, N. K. (1969), 'Varieties of colour anomia', *Brain*, **92**, 847-60

Paillard, J., Michel, F., Stelmach, G. (1983), 'Localization without content a tactile analogue of 'blind sight'', *Archives of Neurology*, **40**, 548-51.

Pirsig, R. (1974) *Zen and the Art of Motorcycle Maintenance* (London: Bodley Head).

Russell, B. (1919), *Introduction to Mathematical Philosophy*, (London: Allen and Unwin), p.71.

Solms, M. (1997), 'What is Consciousness?', *Journal of the American Psychoanalytical Association*, **45**: 681-703.

Stein, G. (1961), *The Autobiography of Alice B. Toklas* (New York: Random House)

Stoerig, P., Cowey, A. (1997) 'Blindsight in man and monkey', *Brain*, **120**, 535-59.