**The Intrinsic Quality of Experience**

Gilbert Harman

*Philosophical Perspectives*, Vol. 4, Action Theory and Philosophy of Mind (1990), 31-52.

Philosophical Perspectives, 4
Action Theory and Philosophy of Mind, 1990

# THE INTRINSIC QUALITY OF EXPERIENCE*

Gilbert Harman
Princeton University

ABSTRACT

There are three familiar and related arguments against psychophysical functionalism and the computer model of the mind. The first is that we are directly aware of intrinsic features of our experience and argues that there is no way to account for this awareness in a functional view. The second claims that a person blind from birth can know all about the functional role of visual experience without knowing what it is like to see something red. The third claims that functionalism cannot account for the possibility of an inverted spectrum. All three arguments can be defused by distinguishing properties of the object of experience from properties of the experience of an object.

## The Problem

Many philosophers, psychologists, and artificial intelligence researchers accept a broadly functionalist view of the relation between mind and body, for example, viewing the mind in the body as something like a computer in a robot, perhaps with massively parallel processing (as in Rumelhart and McClelland 1986). But this view of the mind has not gone unchallenged. Some philosophers and others object strenuously that functionalism must inevitably fail to account for the most important part of mental life, namely, the subjective feel of conscious experience.

The computer model of mind represents one version of func-

tionalism, although it is not the only version. In its most general form, functionalism defines mental states and processes by their causal or functional relations to each other and to perceptual inputs from the world outside and behavioral outputs expressed in action. According to functionalism, it is the functional relations that are important, not the intrinsic qualities of the stuff in which these relations are instanced. Just as the same computer programs can be run on different computers made out of different materials, so functionalism allows for the same mental states and events in beings with very different physical constitutions, since the very same functional relations might be instantiated in beings with very different physical makeups. According to functionalism, beliefs, desires, thoughts, and feelings are not limited to beings that are materially like ourselves. Such psychological states and events might also occur, for example, in silicon based beings, as long as the right functional relations obtained.

Functionalism can allow for the possibility that something about silicon makes it impossible for the relevant relations to obtain in silicon based beings, perhaps because the relevant events could not occur fast enough in silicon. It is even conceivable that the relevant functional relations might obtain only in the sort of material that makes up human brains (Thagard 1986; Dennett 1987, Chapter 9). Functionalism implies that in such a case the material is important only because it is needed for the relevant functional relations and not because of some other more mysterious or magical connection between that sort of matter and a certain sort of consciousness.

Various issues arise within the general functionalist approach. For one thing, there is a dispute about how to identify the inputs to a functional system. Should inputs be identified with events in the external environment (Harman 1988) or should they instead be identified with events that are more internal such as the stimulation of an organism's sensory organs (Block 1986)? There is also the possibility of disagreement as to how deterministic the relevant functional relations have to be. Do they have to be completely deterministic, or can they be merely probabilistic? Or might they even be simply nondeterministic, not even associated with definite probabilities (Harman 1973, pp. 51-53)?

I will not be concerned with these issues here. Instead, I will concentrate on the different and more basic issue that I have already mentioned, namely, whether this sort of functionalism, no matter

how elaborated, can account for the subjective feel of experience, for "what it is like" (Nagel 1974) to undergo this or that experience. Furthermore, I will not consider the general challenge, "How does functionalism account for X?" for this or that X. Nor will I consider negative arguments against particular functionalist analyses. I will instead consider three related arguments that purport to demonstrate that functionalism cannot account for this aspect of experience. I will argue that all three arguments are fallacious. I will say little that is original and will for the most part merely elaborate points made many years ago (Quine 1960, p. 235, Anscombe 1965, Armstrong 1961, 1962, and especially 1968, Pitcher 1971), points that I do not think have been properly appreciated. The three arguments are these:

> First, when you attend to a pain in your leg or to your experience of the redness of an apple, you are aware of an intrinsic quality of your experience, where an intrinsic quality is a quality something has in itself, apart from its relations to other things. This quality of experience cannot be captured in a functional definition, since such a definition is concerned entirely with relations, relations between mental states and perceptual input, relations among mental states, and relations between mental states and behavioral output. For example, "An essential feature of [Armstrong's functionalist] analysis is that it tells us nothing about the intrinsic nature of mental states... He never takes seriously the natural objection that we must know the intrinsic nature of our own mental states since we experience them directly" (Nagel 1970).
>
> Second, a person blind from birth could know all about the physical and functional facts of color perception without knowing what it is like to see something red. So, what it is like to see something red cannot be explicated in purely functional terms (Nagel 1974, Jackson 1982, 1986).
>
> Third, it is conceivable that two people should have similarly functioning visual systems despite the fact that things that look red to one person look green to the other, things that look orange to the first person look blue to the second, and so forth (Lycan 1973, Shoemaker 1982). This sort of spectrum inversion in the way things look is possible but cannot be given a purely functional description, since by

hypothesis there are no functional differences between the people in question. Since the way things look to a person is an aspect of that person's mental life, this means that an important aspect of a person's mental life cannot be explicated in purely functional terms.

## Intentionality

In order to assess these arguments, I begin by remarking on what is sometimes called the intentionality of experience. Our experience of the world has content—that is, it represents things as being in a certain way. In particular, perceptual experience represents a perceiver as in a particular environment, for example, as facing a tree with brown bark and green leaves fluttering in a slight breeze.

One thing that philosophers mean when they refer to this as the intentional content of experience is that the content of the experience may not reflect what is really there. Although it looks to me as if I am seeing a tree, that may be a clever illusion produced with tilted mirrors and painted backdrops. Or it may be a hallucination produced by a drug in my coffee.

There are many other examples of intentionality. Ponce de Leon searched Florida for the Fountain of Youth. What he was looking for was a fountain whose waters would give eternal youth to whoever would drink them. In fact, there is no such thing as a Fountain of Youth, but that does not mean Ponce de Leon wasn't looking for anything. He was looking for something. We can therefore say that his search had an intentional object. But the thing that he was looking for, the intentional object of his search, did not (and does not) exist.

A painting of a unicorn is a painting of something; it has a certain content. But the content does not correspond to anything actual; the thing that the painting represents does not exist. The painting has an intentional content in the relevant sense of "intentional."

Imagining or mentally picturing a unicorn is usefully compared with a painting of a unicorn. In both cases the content is not actual; the object pictured, the intentional object of the picturing, does not exist. It is only an intentional object.

This is not to suppose that mentally picturing a unicorn involves an awareness of a mental picture of a unicorn. I am comparing mentally picturing something with a picture of something, not with

a perception of a picture. An awareness of a picture has as its intentional object a picture. The picture has as its intentional object a unicorn. Imagining a unicorn is different from imagining a picture of a unicorn. The intentional object of the imagining is a unicorn, not a picture of a unicorn.

It is very important to distinguish between the properties of a represented object and the properties of a representation of that object. Clearly, these properties can be very different. The unicorn is pictured as having four legs and a single horn. The painting of the unicorn does not have four legs and a single horn. The painting is flat and covered with paint. The unicorn is not pictured as flat or covered with paint. Similarly, an imagined unicorn is imagined as having legs and a horn. The imagining of the unicorn has no legs or horn. The imagining of the unicorn is a mental activity. The unicorn is not imagined as either an activity or anything mental.

The notorious sense datum theory of perception arises through failing to keep these elementary points straight. According to that ancient theory, perception of external objects in the environment is always indirect and mediated by a more direct awareness of a mental sense datum. Defenders of the sense datum theory argue for it by appealing to the so-called argument from illusion. This argument begins with the uncontroversial premise that the way things are presented in perception is not always the way they are. Eloise sees some brown and green. But there is nothing brown and green before her; it is all an illusion or hallucination. From this the argument fallaciously infers that the brown and green Eloise sees is not external to her and so must be internal or mental. Since veridical, nonillusory, nonhallucinatory perception can be qualitatively indistinguishable from illusory or hallucinatory perception, the argument concludes that in all cases of perception Eloise is directly aware of something inner and mental and only indirectly aware of external objects like trees and leaves.

An analogous argument about paintings would start from the premise that a painting can be a painting of a unicorn even though there are no unicorns. From this it might be concluded that the painting is "in the first instance" a painting of something else that is actual, for example, the painter's idea of a unicorn.

In order to see that such arguments are fallacious, consider the corresponding argument applied to searches: "Ponce de Leon was searching for the Fountain of Youth. But there is no such thing. So

he must have been searching for something mental." This is just a mistake. From the fact that there is no Fountain of Youth, it does not follow that Ponce de Leon was searching for something mental. In particular, he was not looking for an idea of the Fountain of Youth. He already had the idea. What he wanted was a real Fountain of Youth, not just the idea of such a thing.

The painter has painted a picture of a unicorn. The picture painted is not a picture of an idea of a unicorn. The painter might be at a loss to paint a picture of an idea, especially if he is not familiar with conceptual art. It may be that the painter has an idea of a unicorn and tries to capture that idea in his painting. But that is to say his painting is a painting of the same thing that his idea is an idea of. The painting is not a painting of the idea, but a painting of what the idea is about.

In the same way, what Eloise sees before her is a tree, whether or not it is a hallucination. That is to say, the content of her visual experience is that she is presented with a tree, not with an idea of a tree. Perhaps, Eloise's visual experience involves some sort of mental picture of the environment. It does not follow that she is aware of a mental picture. If there is a mental picture, it may be that what she is aware of is whatever is represented by that mental picture; but then that mental picture represents something in the world, not something in the mind.

Now, we sometimes count someone as perceiving something only if that thing exists. So, if there is no tree before her and Eloise is suffering from a hallucination, we might describe this either by saying that Eloise sees something that is not really there or by saying that she does not really see anything at all but only seems to see something. There is not a use of "search for" corresponding to this second use of "see" that would allow us to say that, because there was and is no such thing as the Fountain of Youth, Ponce de Leon was not really searching for anything at all.

But this ambiguity in perceptual verbs does not affect the point I am trying to make. To see that it does not, let us use "see†" ("see-dagger") for the sense of "see" in which the object seen might not exist, as when Macbeth saw a dagger before him.[1] And let us use "see*" ("see-star") for the sense of "see" in which only things that exist can be seen. Macbeth saw† a dagger but he did not see* a dagger.

The argument from illusion starts from a case in which Eloise "sees"

something brown and green before her, although there is nothing brown and green before her in the external physical world. From this, the argument infers that the brown and green she sees must be internal and mental. Now, if "see" is "see†" here, this is the fallacy already noted, like that of concluding that Ponce de Leon was searching for something mental from the fact that there is no Fountain of Youth in the external world. On the other hand, if "see" is "see*" here, then the premise of the argument simply begs the question. No reason at all has so far been given for the claim that Eloise sees* something brown and green in this case. It is true that her perceptual experience represents her as visually presented with something brown and green; but that is to say merely that she sees† something brown and green, not that she sees* anything at all. (From now on I will suppress the † and * modification of perceptual verbs unless indication of which sense is meant is crucial to the discussion.)

Here, some philosophers (e.g. Jackson 1977) would object as follows:

> You agree that there is a sense in which Eloise sees something green and brown when there is nothing green and brown before her in the external world. You are able to deny that this brown and green thing is mental by taking it to be a nonexistent and merely intentional object. But it is surely more reasonable to suppose that one is in this case aware of something mental than to suppose that one is aware of something that does not exist. How can there be anything that does not exist? The very suggestion is a contradiction in terms, since "be" simply means "exist," so that you are really saying that there exists something that does not exist (Quine 1948). There are no such things as nonexistent objects!

In reply, let me concede immediately that I do not have a well worked out theory of intentional objects. Parsons (1980) offers one such theory, although I do not mean to express an opinion as to the success of Parson's approach. Indeed, I am quite willing to believe that there are not really any nonexistent objects and that apparent talk of such objects should be analyzed away somehow. I do not see that it is my job to resolve this issue. However this issue is resolved, the theory that results had better end up agreeing that Ponce de Leon was looking for something when he was looking for the Fountain

of Youth, even though there is no Fountain of Youth, and the theory had better *not* have the consequence that Ponce de Leon was looking for something mental. If a logical theory can account for searches for things that do not, as it happens, exist, it can presumably also allow for a sense of "see" in which Macbeth can see something that does not really exist.

Another point is that Eloise's visual experience does not just present a tree. It presents a tree as viewed from a certain place. Various features that the tree is presented as having are presented as relations between the viewer and the tree, for example, features the tree has from here. The tree is presented as "in front of" and "hiding" certain other trees. It is presented as fuller on "the right." It is presented as the same size "from here" as a closer smaller tree, which is not to say that it really looks the same in size, only that it is presented as subtending roughly the same angle from here as the smaller tree. To be presented as the same in size from here is not to be presented as the same in size, period.

I do not mean to suggest that the way the tree is visually presented as being from here is something that is easily expressed in words. In particular, I do not mean to suggest that the tree can thus be presented as subtending a certain visual angle only to someone who understands words like "subtend" and "angle" (as is assumed in Peacocke 1983, Chapter 1). I mean only that this feature of a tree from here is an objective feature of the tree in relation to here, a feature to which perceivers are sensitive and which their visual experience can somehow represent things as having from here.

Now, perhaps, Eloise's visual experience even presents a tree as seen by her, that is, as an object of her visual experience. If so, there is a sense after all in which Eloise's visual experience represents something mental: it represents objects in the world as objects of visual experience. But this does not mean that Eloise's visual experience in any way reveals to her the intrinsic properties of that experience by virtue of which it has the content it has.

I want to stress this point, because it is very important. Eloise is aware of the tree as a tree that she is now seeing. So, we can suppose she is aware of some features of her current visual experience. In particular, she is aware that her visual experience has the feature of being an experience of seeing a tree. That is to be aware of an intentional feature of her experience; she is aware that her experience has a certain content. On the other hand, I want to argue

that she is not aware of those intrinsic features of her experience by virtue of which it has that content. Indeed, I believe that she has no access at all to the intrinsic features of her mental representation that make it a mental representation of seeing a tree.

Things are different with paintings. In the case of a painting Eloise can be aware of those features of the painting that are responsible for its being a painting of a unicorn. That is, she can turn her attention to the pattern of the paint on the canvas by virtue of which the painting represents a unicorn. But in the case of her visual experience of a tree, I want to say that she is not aware of, as it were, the mental paint by virtue of which her experience is an experience of seeing a tree. She is aware only of the intentional or relational features of her experience, not of its intrinsic nonintentional features.

Some sense datum theorists will object that Eloise is indeed aware of the relevant mental paint when she is aware of an arrangement of color, because these sense datum theorists assert that the color she is aware of is inner and mental and not a property of external objects. But, this sense datum claim is counter to ordinary visual experience. When Eloise sees a tree before her, the colors she experiences are all experienced as features of the tree and its surroundings. None of them are experienced as intrinsic features of her experience. Nor does she experience any features of anything as intrinsic features of her experience. And that is true of you too. There is nothing special about Eloise's visual experience. When you see a tree, you do not experience any features as intrinsic features of your experience. Look at a tree and try to turn your attention to intrinsic features of your visual experience. I predict you will find that the only features there to turn your attention to will be features of the presented tree, including relational features of the tree "from here."

The sense datum theorists' view about our immediate experience of color is definitely not the naive view; it does not represent the viewpoint of ordinary perception. The sense datum theory is not the result of phenomenological study; it is rather the result of an argument, namely, the argument from illusion. But that argument is either invalid or question-begging, as we have seen.

It is very important to distinguish what are experienced as intrinsic features of the intentional object of experience from intrinsic features of the experience itself. It is not always easy to distinguish these things, but they can be distinguished. Consider the experience of

having a pain in your right leg. It is very tempting to confuse features of what you experience as happening in your leg with intrinsic features of your experience. But the happening in your leg that you are presented with is the intentional object of your experience; it is not the experience itself. The content of your experience is that there is a disturbance of a certain specific sort in your right leg. The intentional object of the experience is an event located in your right leg. The experience itself is not located in your right leg. If the experience is anywhere specific, it is somewhere in your brain.

Notice that the content of your experience may not be true to what is actually happening. A slipped disc in your back may press against your sciatic nerve making it appear that there is a disturbance in your right leg when there really is not. The intentional object of your painful experience may not exist. Of course, that is not to say there is no pain in your leg. You do feel something there. But there is a sense in which what you feel in your leg is an illusion or hallucination.

It is true that, if Melvin hallucinates a pink elephant, the elephant that Melvin sees does not exist. But the pain in your leg resulting from a slipped disc in your back certainly does exist.[2] The pain is not an intentional object in quite the way the elephant is. The pain in your leg caused by the slipped disc in your back is more like the afterimage of a bright light. If you look at a blank wall, you see the image on the wall. The image is on the wall, the pain is in your leg. There is no physical spot on the wall, there is no physical disturbance in your leg. The afterimage exists, the pain exists. When we talk about afterimages or referred pains, some of what we say is about our experience and some of what we say is about the intentional object of that experience. When we say the pain or afterimage exists, we mean that the experience exists. When we say that the afterimage is on the wall or that the pain is in your leg, we are talking about the location of the intentional object of that experience.

### Assessment of the First Objection

We are now in a position to reject the first of the three arguments against functionalism which I now repeat:

> When you attend to a pain in your leg or to your experience of the redness of an apple, you are aware of an intrinsic quality of your experience, where an intrinsic

> quality is a quality something has in itself, apart from its relations to other things. This quality of experience cannot be captured in a functional definition, since such a definition is concerned entirely with relations, relations between mental states and perceptual input, relations among mental states, and relations between mental states and behavioral output.

We can now see that this argument fails through confounding a quality of the intentional object of an experience with a quality of the experience itself. When you attend to a pain in your leg or to your experience of the redness of an apple, you are attending to a quality of an occurrence in your leg or a quality of the apple. Perhaps this quality is presented to you as an intrinsic quality of the occurrence in your leg or as an intrinsic quality of the surface of the apple. But it is not at all presented as an intrinsic quality of your experience. And, since you are not aware of the intrinsic character of your experience, the fact that functionalism abstracts from the intrinsic character of experience does not show it leaves out anything you are aware of.

To be sure, there are possible complications. Suppose David undergoes brain surgery which he watches in a mirror. Suppose that he sees certain intrinsic features of the firing of certain neurons in his brain and suppose that the firing of these neurons is the realization of part of the experience he is having at that moment. In that case, David is aware of intrinsic features of his experience. But that way of being aware of intrinsic features of experience is not incompatible with functionalism. Given a functionalist account of David's perception of trees, tables, and the brain processes of other people, the same account applies when the object perceived happens to be David's own brain processes. The awareness David has of his own brain processes is psychologically similar to the awareness any other sighted perceiver might have of those same brain processes, including perceivers constructed in a very different way from the way in which David is constructed.

According to functionalism, the psychologically relevant properties of an internal process are all functional properties. The intrinsic nature of the process is relevant only inasmuch as it is responsible for the process's having the functional properties it has. I have been considering the objection that certain intrinsic features of experience

must be psychologically relevant properties apart from their contribution to function, since these are properties we are or can be aware of. The objection is not just that we can become aware of intrinsic features of certain mental processes in the way just mentioned, that is, by perceiving in a mirror the underlying physical processes that realize those mental processes. That would not be an objection to functionalism. The objection is rather that all or most conscious experience has intrinsic aspects of which we are or can be aware in such a way that these aspects of the experience are psychologically significant over and above the contribution they make to function.

Of course, to say that these aspects are psychologically significant is not to claim that they are or ought to be significant for the science of psychology. Rather, they are supposed to be psychologically significant in the sense of mentally significant, whether or not this aspect of experience is susceptible of scientific understanding. The objection is that any account of our mental life that does not count these intrinsic properties as mental or psychological properties leaves out a crucial aspect of our experience.

My reply to this objection is that it cannot be defended without confusing intrinsic features of the intentional object of experience with intrinsic features of the experience. Apart from that confusion, there is no reason to think that we are ever aware of the relevant intrinsic features of our experiences.

There are other ways in which one might be aware of intrinsic features of our experience without that casting any doubt on functionalism. For example, one might be aware of intrinsic features of experience without being aware of them as intrinsic features of experience, just as Ortcutt can be aware of a man who, as it happens, is a spy without being aware of the man as a spy. When Eloise sees a tree, she is aware of her perceptual experience as an experience with a certain intentional content. Suppose that her experience is realized by a particular physical event and that certain intrinsic features of the event are in this case responsible for certain intentional features of Eloise's experience. Perhaps there is then a sense in which Eloise is aware of this physical process and aware of those intrinsic features, although she is not aware of them as the intrinsic features that they are.

Even if that is so, it is no objection to functionalism. The intrinsic features that Eloise is aware of in that case are no more

psychologically significant than is the property of being a spy to Ortcutt's perception of a man who happens to be a spy. The case gives no reason to think that there is a psychologically significant difference between Eloise's experience and the experience of any functional duplicate of Eloise that is made of different stuff from what Eloise is made of.

Similarly, if Eloise undertakes the sort of education recommended by Paul Churchland (1985) so that she automatically thinks of the intentional aspects of her experience in terms of their neurophysiological causes, then she may be aware of intrinsic features of her experience as the very features that they are. But again that would be no objection to functionalism, since it gives no reason to think that there is a psychological difference between Eloise after such training and a robot who is Eloise's functional duplicate and who has been given similar training (Shoemaker 1985). The duplicate now wrongly thinks of certain aspects of its experience as certain features of certain neurological processes—wrongly, because the relevant processes in the duplicate are not neurological processes at all.

Observe, by the way, that I am not offering any sort of positive argument that Eloise and her duplicate must have experiences that are psychologically similar in all respects. I am only observing that the cases just considered are compatible with the functionalist claim that their experiences are similar.

The objections to functionalism that I am considering in this paper claim that certain intrinsic properties of experience so inform the experience that any experience with different intrinsic properties would have a different psychological character. What I have argued so far is that this objection is not established by simple inspection of our experience.

## Perception and Understanding

Now, let me turn to the second objection, which I repeat:

> A person blind from birth could know all about the physical and functional facts of color perception without knowing what it is like to see something red. So, what it is like to see something red cannot be explicated in purely functional terms.

In order to address this objection, I have to say something about the functionalist theory of the content of mental representations and, more particularly, something about the functionalist theory of concepts. I have to do this because to know what it is like to see something red is to be capable of representing to yourself something's being red. You can represent that to yourself only if you have the relevant concept of what it is for something to be red. The blind person lacks the full concept of redness that a sighted person has; so the blind person cannot fully represent what it is for a sighted person to see something red. Therefore, the blind person cannot be said to know what it is like to see something red.

One kind of functionalist account of mental representation supposes that mental representations are constructed from concepts, where the content of a representation is determined by the concepts it contains and the way these concepts are put together to form that representation (Harman 1987). In this view, what it is to have a given concept is functionally determined. Someone has the appropriate concept of something's being red if and only if the person has available a concept that functions in the appropriate way. The relevant functioning may involve connections with the use of other concepts, connections to perceptual input, and/or connections to behavioral output. In this case, connections to perceptual input are crucial. If the concept is to function in such a way that the person has the full concept of something's being red, the person must be disposed to form representations involving that concept as the natural and immediate consequence of seeing something red. Since the blind person lacks any concept of this sort, the blind person lacks the full concept of something's being red. Therefore, the blind person does not know what it is like to see something red.

It is not easy to specify the relevant functional relation precisely. Someone who goes blind later in life will normally retain the relevant concept of something's being red. Such a person has a concept that he or she would be able to use in forming such immediate visual representations except for the condition that interferes in his or her case with normal visual perception. So, the right functional relation holds for such a person. I am supposing that the person blind from birth has no such concept; that is, the person has no concept of something's being red that could be immediately brought into service in visual representations of the environment if the person were suddenly to acquire sight.

We are now in a position to assess the claim that the person blind from birth could know all the physical and functional facts about color perception without knowing what it is like to see something red. I claim that there is one important functional fact about color perception that the blind person cannot know, namely, that there is a concept *R* such that when a normal perceiver sees something red in good lighting conditions, the perceiver has visual experience with a representational structure containing this concept *R*. The person blind from birth does not know that fact, because in order to know it the person needs to be able to represent that fact to him or herself, which requires having the relevant concepts. A key concept needed to represent that fact is the concept of something's being red, because the fact in question is a fact about what happens when a normal perceiver sees something red. Since the person blind from birth does not have the full concept of something's being red, the person cannot fully understand that fact and so cannot know that fact.

The blind person might know something resembling this, for example, that there is a concept *R* such that, when a normal perceiver sees something that reflects light of such and such a frequency, the perceiver has visual experience with a representational structure containing this concept *R*. But that is to know something different.

The person blind from birth fails to know what it is like to see something red because he or she does not fully understand what it is for something to be red, that is, because he or she does not have the full concept of something's being red. So, contrary to what is assumed in the second objection, the person blind from birth does not know all the functional facts, since he or she does not know how the concept *R* functions with respect to the perception of things that are red.

This response to the second objection appeals to a functionalism that refers to the functions of concepts, not just to the functions of overall mental states. There are other versions of functionalism that try to make do with references to the functions of overall mental states, without appeal to concepts. Some of these versions identify the contents of such states with sets of possible worlds (or centered possible worlds). These versions of functionalism cannot respond to the objection in the way that I have responded. It is unclear to me whether any satisfactory response is possible on behalf of such theories. For example, Lewis (1983) is forced to say that although

the person blind from birth lacks certain skills, e.g., the ability to recognize red objects just by looking at them in the way that sighted people can, this person lacks no information about visual perception. I am not happy with that response, since it is clearly false to say that the person blind from birth does not lack any information.

**Inverted Spectrum**

I now turn to the third objection to functionalism, which I repeat:

> It is conceivable that two people should have similarly functioning visual systems despite the fact that things that look red to one person look green to the other, things that look orange to the first person look blue to the second, and so forth. This sort of spectrum inversion in the way things look is possible but cannot be given a purely functional description, since by hypothesis there are no functional differences between the people in question. Since the way things look to a person is an aspect of that person's mental life, this means that there is an important aspect of a person's mental life that cannot be explicated in purely functional terms.

In order to discuss this objection, I need to say something more about how perceptual states function. In particular, I have to say something about how perceptual states function in relation to belief.

Perceptual experience represents a particular environment of the perceiver. Normally, a perceiver uses this representation as his or her representation of the environment. That is to say, the perceiver uses it in order to negotiate the furniture. In still other words, this representation is used as the perceiver's belief about the environment. This sort of use of perceptual representations is the normal case, although there are exceptions when a perceiver inhibits his or her natural tendency and refrains from using a perceptual representation (or certain aspects of that representation) as a guide to the environment, as a belief about the surroundings. The content of perceptual representation is functionally defined in part by the ways in which this representation normally arises in perception and in part by the ways in which the representation is normally used to guide actions (Armstrong 1961, 1968; Dennett 1969; Harman 1973).

The objection has us consider two people, call them Alice and Fred, with similarly functioning visual systems but with inverted spectra with respect to each other. Things that look red to Alice look green to Fred, things that look blue to Alice look orange to Fred, and so on. We are to imagine that this difference between Alice and Fred is not reflected in their behavior in any way. They both call ripe strawberries "red" and call grass "green" and they do this in the effortless ways in which normal perceivers do who have learned English in the usual ways.

Consider what this means for Alice in a normal case of perception. She looks at a ripe strawberry. Perceptual processing results in a perceptual representation of that strawberry, including a representation of its color. She uses this representation as her guide to the environment, that is, as her belief about the strawberry, in particular, her belief about its color. She expresses her belief about the color of the strawberry by using the words, "it is red." Similarly, for Fred. His perception of the strawberry results in a perceptual representation of the color of the strawberry that he uses as his belief about the color and expresses with the same words, "it is red."

Now, in the normal case of perception, there can be no distinction between how things look and how they are believed to be, since how things look is given by the content of one's perceptual representation and in the normal case one's perceptual representation is used as one's belief about the environment. The hypothesis of the inverted spectrum objection is that the strawberry looks different in color to Alice and to Fred. Since everything is supposed to be functioning in them in the normal way, it follows that they must have different beliefs about the color of the strawberry. If they had the same beliefs while having perceptual representations that differed in content, then at least one of them would have a perceptual representation that was not functioning as his or her belief about the color of the strawberry, which is to say that it would not be functioning in what we are assuming is the normal way.

A further consequence of the inverted spectrum hypothesis is that, since in the normal case Alice and Fred express their beliefs about the color of strawberries and grass by saying "it is red" and "it is green," they must mean something different by their color words. By "red" Fred means the way ripe strawberries look to him. Since that is the way grass looks to Alice, what Fred means by "red" is what she means by "green."

It is important to see that these really are consequences of the inverted spectrum hypothesis. If Alice and Fred meant the same thing by their color terms, then either (a) one of them would not be using these words to express his or her beliefs about color or (b) one of them would not be using his or her perceptual representations of color as his or her beliefs about color. In either case, there would be a failure of normal functioning, contrary to the hypothesis of the inverted spectrum objection.

According to functionalism, if Alice and Fred use words in the same way with respect to the same things, then they mean the same things by those words (assuming also that they are members of the same linguistic community and their words are taken from the common language). But this is just common sense. Suppose Alice and Humphrey are both members of the same linguistic community, using words in the same way, etc. Alice is an ordinary human being and Humphrey is a humanoid robot made of quite a different material from Alice. Common sense would attribute the same meanings to Humphrey's words as to Alice's, given that they use words in the same way. Some sort of philosophical argument is needed to argue otherwise. No such argument has been provided by defenders of the inverted spectrum objection.

Shoemaker (1982) offers a different version of the inverted spectrum objection. He has us consider a single person, call him Harry, at two different times, at an initial time of normal color perception and at a later time after Harry has suffered through a highly noticeable spectrum inversion (perhaps as the result of the sort of brain operation described in Lycan 1973, in which nerves are switched around so that red things now have the perceptual consequences that green things used to have, etc.) and has finally completely adapted his responses so as to restore normal functioning. Shoemaker agrees that Harry now has the same beliefs about color as before and means the same things by his color words, and he agrees that there is a sense in which strawberries now look to Harry the same as they looked before Harry's spectrum inversion. But Shoemaker takes it to be evident that there is another sense of "looks" in which it may very well be true that things do not look the same as they looked before, so that in this second sense of "looks" red things look the way green things used to look.

In other words, Shoemaker thinks it is evident that there may be a psychologically relevant difference between the sort of experience

Harry had on looking at a ripe strawberry at the initial stage and the experience he has on looking at a ripe strawberry at the final stage (after he has completely adapted to his operation). That is, he thinks it is evident that there may be a psychologically relevant difference between these experiences even though there is no functional difference and no difference in the content of the experiences.

Now, this may seem evident to anyone who has fallen victim to the sense datum fallacy, which holds that one's awareness of the color of a strawberry is mediated by one's awareness of an intrinsic feature of a perceptual representation. But why should anyone else agree? Two perceptual experiences with the same intentional content must be psychologically the same. In particular, there can be nothing one is aware of in having the one experience that one is not aware of in having the other, since the intentional content of an experience comprises everything one is aware of in having that experience.

I suggest that Shoemaker's inverted spectrum hypothesis will seem evident only to someone who *begins* with the prior assumption that people have an immediate and direct awareness of intrinsic features of their experience, including those intrinsic features that function to represent color. Such a person can then go on to suppose that the intrinsic feature of experience that represents red for Alice is the intrinsic feature of experience that represents green for Fred, and so forth. This prior assumption is exactly the view behind the first objection, which I have argued is contrary to ordinary experience and can be defended only by confusing qualities of the intentional objects of experience with qualities of the experience itself. Shoemaker's inverted spectrum hypothesis therefore offers no independent argument against functionalism.[3]

## Conclusion

To summarize briefly, I have described and replied to three related objections to functionalism. The first claims that we are directly aware of intrinsic features of our experience and argues that there is no way to account for this awareness in a functional view. To this, I reply that when we clearly distinguish properties of the object of experience from properties of the experience, we see that we are not aware of the relevant intrinsic features of the experience. The

second objection claims that a person blind from birth can know all about the functional role of visual experience without knowing what it is like to see something red. To this I reply that the blind person does not know all about the functional role of visual experience; in particular, the blind person does not know how such experience functions in relation to the perception of red objects. The third objection claims that functionalism cannot account for the possibility of an inverted spectrum. To this I reply that someone with the relevant sort of inverted spectrum would have to have beliefs about the colors of things that are different from the beliefs others have and would have to mean something different by his or her color terms, despite being a functionally normal color perceiver who sorts things by color in exactly the way others do and who uses color terminology in the same way that others do. Functionalism's rejection of this possibility is commonsensical and is certainly not so utterly implausible or counter-intuitive that these cases present an objection to functionalism. On the other hand, to imagine that there could be relevant cases of inverted spectrum without inversion of belief and meaning is to fall back onto the first objection and not to offer any additional consideration against functionalism.

**Notes**

*The preparation of this paper was supported in part by research grants to Princeton University from the James S. McDonnell Foundation and the National Science Foundation.

1. W. Shakespeare, Macbeth, Act II, Scene I: Is this a dagger which I see before me, The handle toward my hand? Come let me clutch thee. I have thee not, and yet I see thee still. Art thou not, fatal vision, sensible To feeling as to sight? or art thou but A dagger of the mind, a false creating, Proceeding from the heat oppressed brain?... ...I see thee still; And on thy blade and dudgeon gouts of blood, Which was not so before. There's no such thing; it is the bloody business which informs Thus to mine eyes.
2. I am indebted to Sydney Shoemaker for emphasizing this to me.
3. I should say that Shoemaker himself does not offer his case as an objection to what he calls functionalism. He claims that his version of functionalism is compatible with his case. But I am considering a version of functionalism that is defined in a way that makes it incompatible with such a case.

**References**

Anscombe, G. E. M. (1965) "The intentionality of sensation: a grammatical feature," *Analytical Philosophy*, second series, edited by R. J. Butler (Oxford, Blackwell); reprinted in Anscombe, G.E.M., *Metaphysics and the Philosophy of Mind: Collected Philosophical Papers, Volume II* (Minneapolis, Minnesota; University of Minnesota Press: 1981) pp. 3-20.

Armstrong, David M. (1961) *Perception and the Physical World* (London: Routledge and Kegan Paul).

Armstrong, David M. (1962) *Bodily Sensations* (London: Routledge and Kegan Paul).

Armstrong, David M. (1968) *The Materialist Theory of Mind* (London: Routledge and Kegan Paul).

Block, Ned (1986) "Advertisement for a semantics for psychology," *Midwest Studies in Philosophy* 10: 615-678.

Churchland, Paul (1985) "Reduction, qualia, and the direct introspection of mental states," *Journal of Philosophy* 82: 8-28.

Dennett, Daniel C. (1969) *Content and Consciousness* (London: Routledge and Kegan Paul).

Dennett, Daniel C. (1987) *The Intentional Stance* (Cambridge, Massachusetts: MIT Press).

Harman, Gilbert (1973) *Thought* (Princeton, New Jersey: Princeton University Press).

Harman, Gilbert (1987) "(Nonsolipsistic) conceptual role semantics," *New Directions in Semantics*, edited by Ernest LePore, London, Academic Press (1987) 55-81.

Harman, Gilbert (1988) "Wide functionalism," *Cognition and Representation*, edited by Stephen Schiffer and Susan Steele (Boulder, Colorado: Westview Press) 11-20.

Jackson, Frank (1977) *Perception: A Representative Theory* (Cambridge, England: Cambridge University Press).

Jackson, Frank (1982) "Epiphenomenal qualia," *Philosophical Quarterly* 32: 127-32.

Jackson, Frank (1986) "What Mary didn't know," *Journal of Philosophy* 83: 291-295.

Lewis, David K. (1983) "Postscript to 'Mad pain and martian pain'," *Philosophical Papers*, Volume 1, (New York: Oxford University Press) 130-132.

Lycan, William G. (1973) "Inverted spectrum" *Ratio* 15.

Nagel, Thomas (1970) "Armstrong on the mind," *Philosophical Review* 79, reprinted in Reading in the Philosophy of Psychology Volume 1, edited by Ned Block (Cambridge, Massachusetts: Harvard University Press).

Nagel, Thomas (1974) "What is it like to be a bat?" *Philosophical Review* 83: 435-450.

Parsons, Terence (1980) *Nonexistent Objects* (New Haven: Yale University Press).

Peacocke, Christopher (1983) *Sense and Content* (Oxford: Oxford University Press).

Pitcher, George (1971) *A Theory of Perception* (Princeton, New Jersey: Princeton University Press).

Quine, W. V. (1948) "On what there is," *Review of Metaphysics*, reprinted in *From a Logical Point of View* (Cambridge, Massachusetts; Harvard University Press: 1953).

Quine, W. V. (1960) *Word and Object* (Cambridge, Massachusetts: MIT Press).

Rumelhart, David E., and McClelland, James L. (1986) *Parallel Distributed Processing*, 2 volumes (Cambridge, Massachusetts: MIT Press).

Shoemaker, Sydney (1982) "The inverted spectrum," *Journal of Philosophy* 79: 357-81.

Shoemaker, Sydney (1985) "Churchland on reduction, qualia, and introspection," *PSA 1984*, Volume 2 (Philosophy of Science Association) pp. 799-809.

Thagard, Paul T. (1986) "Parallel computation and the mind-body problem," *Cognitive Science* 10: 301-318.