

## CONSCIOUSNESS OUTSIDE THE HEAD

François Tonneau  
*University of Guadalajara*

**ABSTRACT:** Brain-centered theories of consciousness seem to face insuperable difficulties. While some philosophers now doubt that the hard problem of consciousness will ever be solved, others call for radically new approaches to conscious experience. In this article I resurrect a largely forgotten approach to consciousness known as neorealism. According to neorealism, consciousness is merely a part, or cross-section, of the environment. Neorealism implies that all conscious experiences, veridical or otherwise, exist outside of the brain and are wholly independent of being perceived or not; nonveridical perceptions of the environment over an arbitrarily short period of time are supposed to be objective constituents of the environment over a more extended time scale. I argue here that neorealism fares at least as well as brain-centered theories of consciousness on a number of fundamental issues. On one fundamental issue—the nature of the relation between veridical and nonveridical perceptions—neorealism outperforms its competitors.

*Key words:* neorealism, direct realism, consciousness, cross-section, environment

Asked to remember what my grandmother looks like (she lives thousands of miles from me), I can evoke an image of her in my mind—an image that shares at least some of the perceptible features of my grandmother, such as the shape of her hands or the color of her eyes. Not all of her features are present (the mental image is much vaguer), but at least some are. Yet no process in my head could have these features. If consciousness is a brain process (Place, 1956), how could my conscious experience of my grandmother have *these* features—such as the color of her eyes—that no brain process could have?

The answer is, I believe, obvious. If consciousness is a brain process, then conscious experience *cannot* have the features that in fact it has (Shaffer, 1963), for these features are features of the environment (such as the colors and shapes of surrounding objects) and not of brain processes. That conscious experience contains environmental properties has a further, inescapable consequence, for if consciousness presents us with actual features *of surrounding objects*, then consciousness must be located where these very features and objects are themselves located: *in the environment*. If my experience of my grandmother

---

**AUTHOR'S NOTE:** Parts of this argument are scattered in time at the Annual Conference of the Experimental Analysis of Behavior Group (York, England, April 1990), the Annual Meeting of the Association for Behavior Analysis (Atlanta, GA, USA, May 1991), and the Seventh Biannual Symposium on the Science of Behavior (Guadalajara, México, February 2002). I thank José Burgos, Alan Costall, Fred Dretske, and Nicholas Thompson for their comments. Please address all correspondence to François Tonneau, CEIC, Universidad de Guadalajara, 12 de Diciembre 204, Col. Chapalita, CP 45030, Guadalajara, Jalisco, México. Email: ftonneau@udgserv.cencar.udg.mx.

presents perceptible features that my grandmother actually has, then my experience of my grandmother must be located where my grandmother *herself* is located: outside of my head.

Of course, if consciousness did not actually present us with properties of external objects but only with representations thereof, or internal activities related to (but distinct from) these properties, one could perhaps maintain that consciousness is inside the head. However, I will argue in this article that conscious experience presents us not with covert surrogates of properties of the environment but with these properties themselves. I will also argue that any other view of consciousness leads to intolerable, self-refuting consequences. The conclusion is as obvious as it is inescapable: Consciousness is outside the head (and, more generally, outside the body). Consciousness is in the environment.

The notion that consciousness resides in the environment itself was central to an approach to psychology conceived at the beginning of the last century and known as the New Realism or neorealism (Holt et al., 1910, 1912).<sup>1</sup> The neorealists believed that a person's conscious experience at any instant was the part of the environment acting on this person at this instant—and nothing else. We shall see that in order to provide a plausible account of consciousness, this part, or *cross-section*, of the environment must be conceptualized in a rather abstract and complicated way; in any event, according to neorealism (e.g., Holt, 1914) consciousness is identical with not the entire environment but only with a part of it.

Many veridical perceptions include parts of one's body (e.g., hands and limbs) as well as environmental objects. Neorealism assumes that parts of the body enter consciousness in the same way that any part of the environment can enter consciousness. Hence, the body has no special status in the neorealist theory of conscious experience. In contrast, most theories other than neorealism assume that consciousness depends on the existence of, and somehow reside in, complex biological systems such as bodies or brains. According to neorealism, however, consciousness does *not* depend on and does *not* reside in biological systems such as bodies or brains. Consciousness can be, and often is, located entirely outside anyone's body and is never located in anyone's brain. What is true is that sometimes brains are located in consciousness. When looking at the brain of a rabbit, for example, properties of this brain (the rabbit's) are present in my consciousness. But the properties of the rabbit's brain that appear in my consciousness are not in *my* brain. They are outside my brain, in the environment.

As far as veridical perception is concerned, neorealism is just a form of direct realism consistent with common sense. When veridically perceiving a rabbit, common sense has it that I am aware of actual features of the rabbit itself as opposed to internal representations or activities in my brain or my body. Neorealism goes beyond direct realism by extending the thesis of independence and externality of consciousness to all experiences, however unreal or subjective they might appear to be. Neorealism maintains that dreams, memories, mental

---

<sup>1</sup> For a good survey of Holt's neorealism, its antecedents, and its relations to later ecological psychology, see Heft (2001, ch. 2).

## CONSCIOUSNESS OUTSIDE

images, illusions, and hallucinations, no less than veridical perceptions, exist outside the head, independent of being perceived or not (Nunn, 1909-1910).

Direct realism with respect to veridical perception may be controversial, but at least it seems to make sense. Neorealism, in contrast, just seems crazy. How could dreams, for example, reside in the environment? How could hallucinations exist objectively outside of the head? And how could neorealism address all of the evidence that seems to show that consciousness is a brain process? Intuitions to the contrary notwithstanding, neorealism can answer these questions, and I will try to explain how. I will also examine how neorealism fares on various fundamental issues in comparison with the *internal view* of consciousness. By *internal view*, I mean the belief that a person's conscious experiences do not exist in the environment, independent of being perceived or not, but depend on this person and reside in this person. The internal view includes virtually any materialist theory of consciousness, from the radical behaviorist view of private events (Skinner, 1974) to central-state identity theory to recent functionalist accounts (e.g., Dennett, 1991).<sup>2</sup>

I will compare neorealism and the internal view on five fundamental issues. First, the problem of *bodily dependency*: Conscious experiences seem to depend on body states (brain states for example). How can this be if neorealism is true? Second, the problem of *environmental silence*: Conscious contents (in mental imagery, for example) can change, even in an unchanging environment (say, a dark room). This kind of dissociation suggests that consciousness cannot reside in the environment and therefore that neorealism is false. Third, the problem of *secondary qualities*: At a most basic level, physics seems to have no room for colors, odors, and sounds. How then could colors, odors, and sounds (for example) reside in the very environment that physics describes? Fourth, the problem of *illusion* or error: In a variety of cases that range from mild (mental imagery) to severe (hallucinations), the contents of experience seem to be completely distinct from the environment. How could direct realism in general, and neorealism in particular, explain such discrepancies? Finally, and most importantly, the problem of *veridical and nonveridical perceptions*: How are veridical and nonveridical conscious contents related, and what does this relation imply for a theory of consciousness?

The version of neorealism that I will defend involves two main notions. The first one, the notion of cross-section, was a central part of Holt's (1912) own neorealism. The second one, the notion of a higher-order property of temporally extended environmental variables, derives from modern behavior analysis (Baum, 1989; Himeline, 1981) but is otherwise fully consistent with the original neorealist program. Applying the concept of temporal molarity to the environment is precisely what the first neorealists lacked to maintain their strict realism and

---

<sup>2</sup> Some readers might be surprised to see Skinner's (1974) radical behaviorism as an example of the internal view. My suggestion is nevertheless correct. Radical behaviorism does identify dreams, mental images, and other "private events" with covert occurrences inside the body. For further discussion of the similarity between Skinnerian behaviorism and cognitive theories of consciousness, see Schnaitter (1984).

give a clear account of hallucinations and dreams as overt environmental objects. Now that this concept is available (see below), we can build neorealism anew and see where it leads; what I will call neorealism in this article is simply the resulting effort.

Partly for historical reasons (e.g., Smith, 2000) and partly because some of the associated phenomena (e.g., mental images and dreams) seem especially puzzling, most of my arguments and examples will involve visual perception. Obviously, much additional work would be needed to cover other sensory modalities and phenomena such as pain and bodily sensations, but I hope that the ways in which neorealism could be extended to such cases will be reasonably clear.<sup>3</sup>

### The Concept of Cross-Section

The neorealist notion of cross-section (Holt, 1912, 1914) can be applied to any physical system. Consider a reference system and a relation R in which this system and parts of its surroundings could participate. A *cross-section* is a function of the system state that takes its values in the environment of the reference system. The value of this function at any moment is called the *content* of the cross-section and consists in the part of the environment that fulfills R.

A concrete example (Figure 1) shows how simple the concept of cross-section really is. In Figure 1, the reference system is a transparent box with a searchlight mounted on a wheel inside the box. Turning the wheel causes the searchlight to illuminate various objects (A, B, C) in the environment outside the box. Here the reference system is the box with the wheel and its searchlight; the relation R that defines the cross-section consists in being illuminated; and the content of the cross-section is the illuminated object. At the moment shown in Figure 1, the object A is the content of the cross-section.

Which object the searchlight illuminates obviously depends on the orientation of the wheel inside the box. As long as the wheel remains in the position shown in Figure 1, for example, object A will remain illuminated. Turning the wheel slightly to the right, however, would cause the searchlight to illuminate B instead of A; and turning the wheel even more to the right would cause the searchlight to illuminate C instead of B. Over time, any particular series of orientations of the wheel will be perfectly correlated with a particular series of illuminated objects in the environment:

O<sub>1</sub>    O<sub>1</sub>    O<sub>1</sub>    O<sub>2</sub>    O<sub>3</sub>    O<sub>2</sub>    O<sub>3</sub>    O<sub>1</sub>    (1)

A    A    A    B    C    B    C    A    (2)

Here the top series (1) shows the successive states of the reference system (the positions of the wheel inside the box), whereas the bottom series (2) shows the successive contents of the cross-section (a particular series of illuminated objects);

---

<sup>3</sup> For a neorealist conception of pain see McKenzie (1968).

## CONSCIOUSNESS OUTSIDE

time flows from left to right. The correspondence between the two series illustrates a crucial property of cross-sections: As long as the reference system remains in the same state, the content of the cross-section does not change (see the first three instants of Series 1 and 2). Conversely, *changes in the state of the reference system entail corresponding changes in the content of the cross-section* (see the last five instants of Series 1 and 2).

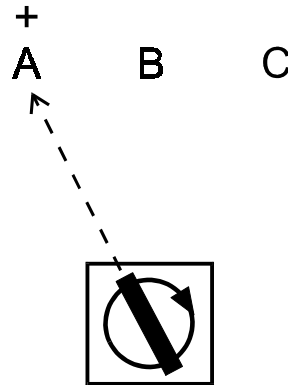


Figure 1. A simple example of cross-section. The reference system is a transparent box (black square) with a searchlight (filled rectangle) mounted on a wheel inside the box (open circle). Turning the wheel causes the searchlight to illuminate different objects (A, B, C) in the environment. At this particular moment, the searchlight illuminates the object A; hence A is the content of the cross-section defined by the relation of illumination (symbolized by an arrow between the box and the object). A plus sign indicates the current content of the cross-section.

Why this correspondence should hold is easily explained. Because the content of a cross-section is defined by its relation to a reference system, a given object (A) will be able to sustain this relation only when the system is in a particular state. Changing the state of the system will disable the relation between A and the reference system and enable *another* object to stand in the same relation, the net result being that the cross-section will move from one object to another. Anybody who has ever explored a landscape with a lantern intuitively understands how a cross-section works.

Notice that in Series 1 and 2 the changes in the reference system and those of the cross-section are of entirely different kinds. The terms ( $O_1, O_2, O_3$ ) of Series 1 are the successive states of a single object—the reference system. The terms (A, B, C) of Series 2, however, are *different* objects coming to entertain a single relation to the reference system. These objects (A, B, C, etc.) enter the illuminated area one after the other then return to obscurity while retaining their separate identities. As the cross-section moves from A to B, for example, obviously A does not transform into B; neither does A crumble or vanish. Rather, as the wheel inside the box

moves from one *state* to another, the relation of illumination is automatically transferred from one *object* (A) to another (B).

In what follows I will speak of *state changes* and *sectional shifts* to contrast these two kinds of change. While state changes involve a single *thing* moving from one *state* to another, sectional shifts involve a single relation moving from one *thing* to another. The concept of cross-section points to the simple fact that state changes in a reference system cause sectional shifts in its surroundings.

Another point worth remembering is the difference between the dependency of “the object” and the dependency of “which object” (Montague, 1912, p. 295). In Figure 1, which object the searchlight illuminates obviously depends on the orientation of the wheel (if the wheel changes its orientation *another* object would be illuminated); however, *the* object that the searchlight illuminates does not depend on the orientation of the wheel. This object (say, A) resides in the *environment*, not inside the wheel, and its existence does not depend in any way on the wheel, the box, or its content. Analogously, which object I see depends on the orientation of my neck, but *the* object that I see (a table) does not depend on the orientation of my neck, and this object is certainly not in my neck.

This distinction holds of any cross-section. Which object is the content of a cross-section depends on the state of the reference system, but *the* object that is the content of the cross-section does not depend on the state of the reference system. From the dependency of “which object” (a selective dependency) one cannot infer the dependency of the object (a creative dependency). Wrongly inferring the latter from the former has led to serious mistakes in theories of consciousness.

### Phenomenal Properties

Locating dreams and mental images in the environment requires more than a bare appeal to the properties of cross-sections. What is needed, and what was largely missing from the early neorealist program (Holt et al., 1910, 1912), is a conception of the environment in temporally extended terms, supplemented by various hypotheses about higher-order features of sequences of environmental events.

Perhaps I should begin by stating two general assumptions about properties that I will make use of. First, I assume that properties exist and consist of universals; properties can be instantiated by more than one particular.<sup>4</sup> Second, I assume that properties are not free-floating but are located in space and time, where their bearers are themselves located. Armstrong (1978a) gives many arguments to support the first assumption. In favor of the second assumption (defended by Donagan, 1963, p. 222), I note that at least some *prima facie* properties and their locations are open to inspection. We can certainly *point* to colors and shades to indicate where they are located. In Figure 2, for example, it will be evident to a majority of nonphilosophers that grayness is located where the

---

<sup>4</sup> An alternative would be to adopt an ontology of tropes (i.e., nonrepeatable properties) with primitive resemblances among them. See Armstrong (1989, ch. 6).

## CONSCIOUSNESS OUTSIDE

triangle and the square are located, and nowhere else. Because they occupy scattered portions of space, two properties might seem to bear incompatible spatial relations to each other, but the apparent incompatibility is easily resolved by relativizing spatial relations to places. Thus, grayness is to the left of blackness in the upper part of Figure 2 and to the right of blackness in the lower part; over the figure as a whole, grayness is neither to the left nor to the right of blackness.

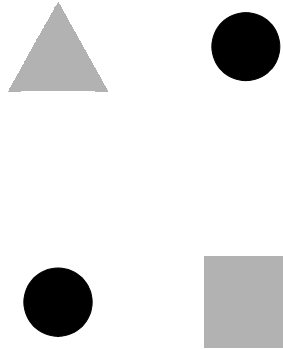


Figure 2. In this picture the location of environmental properties such as grayness and blackness is open to plain inspection.

In this article I will identify phenomenal properties with higher-order features of a person's cumulative *path*, defined as the entire sequence of environmental events or objects with which the person has ever interacted (Tonneau, 2001, pp. 21-23). A path is clearly an environmental entity, no more "subjective" or "internal" than one's house or one's country. Over time, a person's path grows by incorporating more and more extended portions of the environment (Tonneau, 2001, Figure 5a), and at any moment the path includes numerous past objects and events as its temporal parts.<sup>5</sup> In Figure 3a, for instance, the environmental path existing from  $t_0$  to  $t$  includes the object E (represented by a shaded area) as one its temporal parts (a temporal part coincident with time  $t_1$ ).

Let us assume that E, a temporal part of the path, has a physical property  $x$ . By "physical property" I mean one of the properties postulated or measured by (current or "completed") physics, and I exclude putative secondary properties such as smells, sounds, and colors. The property  $x$  could, for example, consist in being an electromagnetic wave of such-and-such frequency. If E has the property  $x$ , then the path that includes E (Figure 3a) has another property: the property of having a temporal part with the property  $x$ . These two properties,  $x$  and *having-a-temporal-part-with-the-property-x*, being properties of different particulars, have different locations in space and time, as argued above. The first property,  $x$ , being a property of the object E, is located where E is located: at time  $t_1$ . The second property, however, *having-a-temporal-part-with-the-property-x*, being a property of the path,

---

<sup>5</sup> In philosophy what I am calling a "path" is often known as a space-time *worm* (e.g., Sider, 1996).

is located where the *path* itself is located: over the whole interval from  $t_0$  to  $t$  (see Figure 3a).

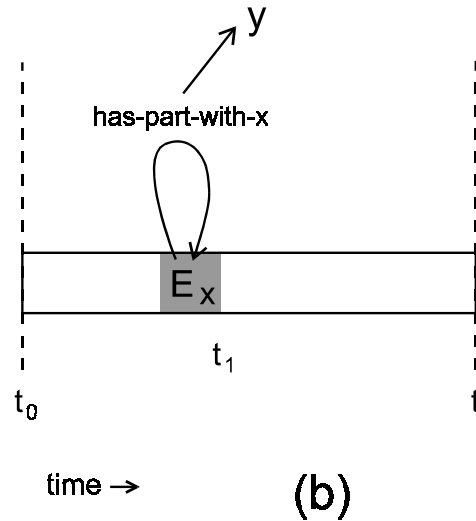
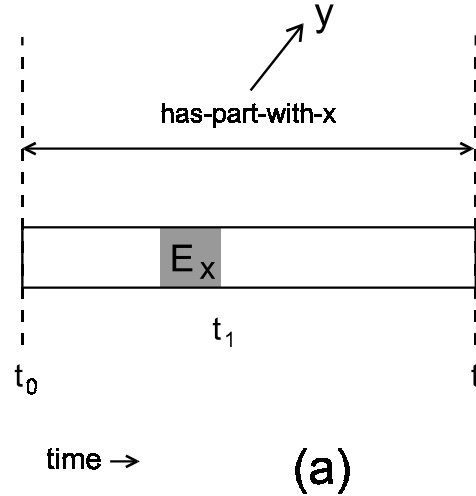


Figure 3. (a) An environmental path present over the time interval  $[t_0, t]$ . The object  $E$  (represented by a shaded area with the physical property  $x$ ), present at time  $t_1$ , is one of the many temporal parts of the path. Because it includes  $E$ , the path has the property of having a temporal part with the property  $x$ . The location of this property of the path is indicated by a horizontal arrow from  $t_0$  to  $t$ . This property has, in turn, a higher-order property  $y$ . See main text for more details; also see Tonneau, 2001, Figure 5a. (b) The analysis applied to the entire path in panel (a) also applies to the object  $E$  itself. Time flows from left to right.



I need one additional assumption to fit phenomenal properties in this picture. I will assume that the path's property of having-a-temporal-part-with-the-property- $x$  has itself a higher-order property  $y$ . This third property,  $y$ , being a property of a property located where the path is located, must also be located where the path is located (Figure 3a). I assume that this third property,  $y$ , is a phenomenal property. If the physical property  $x$  consists in having such-and-such surface spectral reflectance, for example, then  $y$ , a property of the path's property of having a temporal part with spectral reflectance  $x$ , is presumably a *color*.<sup>6</sup>

Thus, we have three distinct properties. The first property,  $x$ , is a physical property of an object or event  $E$ . The second property, having-a-temporal-part-with-the-property- $x$ , is a property of a sequence of physical events or histories of physical objects (a path). The third property,  $y$ , is a property of the second property, hence a property of a property of a path of physical events or histories of physical objects. The first property is located where the object or event  $E$  is located, whereas the second and third properties are spread over the entire path. The second and third properties share the same location (the path), but the second property is a property of the path, whereas the third property is a property of a property of the path.

Other putative cases of properties with properties are easy to devise. The property of being red and square, for example, is a complex property, as is the property of being blue and triangular. Thus, these two properties share the higher-order property of being a complex property (Armstrong, 1978b); however, a property's property of being complex is what Armstrong (1978b) calls a "formal" property (p. 138). Formal properties presumably include mathematical, logical, and mereological properties. The notion that *phenomenal* properties (e.g., colors and sounds) are higher-order properties of properties of the environment is more controversial and can be justified only through its role in a coherent account of consciousness.

It is important to understand how the present approach differs from common views on primary and secondary qualities. The account I propose does acknowledge a distinction between primary qualities (typically mass, charge, and spin) and at least some of the secondary qualities present in conscious experience such as colors and sounds; however, in the neorealist account the secondary qualities are as objective as the primary ones (Nunn, 1909-1910). Secondary qualities, being properties of the property of having a temporal part with such-and-such primary property, exist in the environment, fully independent of being

---

<sup>6</sup> I mention surface spectral reflectance only as a possible candidate for the property that "underlies" colors in the sense that  $x$  "underlies"  $y$  in Figure 3. What stimulus properties (or complex combinations thereof) actually do the job remains to be determined—assuming that the job can indeed be done, which many proponents of the internal view have denied. Most versions of the internal view assume that colors are in some sense creations of the brain (e.g., McGilvray, 1994); for a recent view to the contrary, see Byrne and Hilbert (2003). Notice that whereas Byrne and Hilbert (2003) identify colors with physical properties ( $x$ ), neorealism as I defend it identifies colors with properties (such as  $y$  in Figure 3) of the property of having a temporal part with such-and-such physical property  $x$ .

perceived or not. They are neither mind- nor brain-dependent. Were we and all other color perceivers to vanish, for example, colors would not vanish (contrast with Locke, 1690/1924, book II, ch. viii). Neither are secondary qualities (dispositions) to produce internal sensations. According to neorealism, the “secondary” qualities of the environment are just as occurrent and objective as the primary qualities.

Furthermore, primary properties ( $x$  in Figure 3a) should correspond to particular secondary properties in temporally extended portions of the environment ( $y$  in Figure 3a). The property that has  $y$  as a higher-order property might, of course, consist not in having a temporal part with a single property  $x$  (Figure 3a) but in having a temporal part with several properties ( $x_1, x_2$ , etc.) standing in a variety of physical relations to one another. In either case, the assumed lawful correspondence between  $x$  and  $y$  entails a supervenience of phenomenal properties on physical properties: no phenomenal difference without a physical difference. This supervenience thesis, however, is in itself compatible with a broad range of views on the mind–body problem (Kim, 1998, pp. 9-15). What distinguishes the neorealist view is the assumed mind-independent location of phenomenal properties in the public environment.

Properties such as shape raise interesting problems. Are shapes and spatial relations primary or secondary qualities? Shapes are clearly present in conscious experience, but they are also the topic of much theorizing in physics and chemistry. Imagine that in Figure 3a the path includes a temporal part with the physical properties  $x_1$  and  $x_2$  standing in a spatial relation  $r_{12}$ . The path has the property of having a temporal part with the property  $x_1$ , and this property has a higher-order property  $y_1$ . Similarly, the path has the property of having a temporal part with the property  $x_2$ , and this property has a higher-order property  $y_2$ . The path also has a third, structural property  $P$ : the property of having a temporal part with the properties  $x_1$  and  $x_2$  standing in a spatial relation  $r_{12}$ . I assume that this property  $P$  has a higher-order, structural property: the property  $S$  of having  $y_1$  and  $y_2$  standing in a particular spatial relation  $R_{12}$ . The property  $S$ , I suggest, is a phenomenal shape.

One might maintain that  $r_{12}$  and  $R_{12}$  are two different kinds of relations. Any difference between the two types of relations would legitimize the traditional distinction between “physical space” and “phenomenal space” or spaces—a distinction that Feigl (1958) thought “indispensable” (p. 407). If the distinction between “physical” space and “phenomenal” space is legitimate,  $r_{12}$  should qualify as a *physical* spatial relation and  $R_{12}$  as a *phenomenal* spatial relation. Perhaps the distinction does not hold; perhaps  $r_{12}$  and  $R_{12}$  are spatial relations of the same type. I will not try here to adjudicate this difficult issue. The important point is that whether  $R_{12}$  and  $r_{12}$  turn out to be of identical types or not, according to neorealism phenomenal shapes (involving  $R_{12}$ ) are as objective and mind-independent as physical shapes (involving  $r_{12}$ ), and both exist in the environment.

Although Figure 3a suggests that phenomenal properties (such as  $y$ ) exist only over the extended time frame of a person’s path, the analysis developed so far can be reapplied downward to the original object (E) with its physical property  $x$

(Figure 3b). This object can itself be seen as a short history which, like the entire path, has not only the property  $x$  but also the property of having a (proper or improper) temporal part with the property  $x$ . This last property of  $E$  also has the higher-order property  $y$ , hence not only the whole path (Figure 3a) has a property with the property  $y$ , but the object  $E$  itself (or its own short history) has a property with the property  $y$  (Figure 3b).

In sum, if an object  $E$  has a historical property with a phenomenal property  $y$ , any portion of the environment extended enough to include  $E$  will have a historical property with the same phenomenal property  $y$ . I assume that these properties can become part of a person's consciousness by acting on (having *effects* on, causing *changes* in) this person. How the relevant effects or changes should be defined is a complex issue that is discussed later in more detail. These changes certainly involve neural and behavioral dimensions. A property may enter consciousness when this property causes some specific reactions inside the brain or, alternatively, causes changes in one's dispositions to behave (however understood). The two criteria are clearly linked to each other; the brain reactions that should count as a criterion of consciousness (Holt's "specific response") are presumably those that result, at least potentially, in overt behavior (Holt, 1914, p. 312).

To frame the discussion, I will temporarily adopt a physiological criterion of property action and identify consciousness with the part of the environment that evokes specific changes in someone's brain (Holt, 1914). A person's conscious experience at any time will be a set (or bundle) of properties present in the environment and acting on this person from the outside. This bundle of properties will be a cross-section of the public environment lawfully related to (but located outside) the person's brain.

### **The Argument from Bodily Dependency**

Recall the list of questions I put forward as a challenge to any adequate theory of consciousness. The first issue was that of bodily dependency: How could consciousness be in the environment if it depends on the body or the brain? Invoking bodily dependencies as a proof of the internal location of consciousness remains one of the commonest arguments against direct realism and in favor of the internal view (e.g., Metzger, 1972). The argument might involve mere correlations between bodily events and conscious experiences. Most often, however, the argument relies on actual dependencies between bodily events or states (from fever to brain injuries) and conscious events or states (from blurred vision to rank hallucinations). At first sight such dependencies seem devastating for neorealism: Why should something that is outside the body change whenever the body changes?

Yet the argument from bodily dependency proves nothing. The facts of dependency show merely that having one's body or brain in a particular physical state  $O_1$  causes one to be conscious of a content  $A$ , whereas having one's body or brain in another state  $O_2$  causes one to be conscious of *another* content,  $B$ . This apparent dependency is fully compatible with an environmental location for  $A$  and

B. According to neorealism, an object or part of the environment<sup>7</sup> is present in consciousness only if the object causes specific changes in the body or the brain; however, an object might be able to do so only if (when) the person's body or brain is in a particular state.

In Figure 4, for example, the object A can cause a change in the body or the brain only when the latter is in an initial state  $O_1$  (upper panel), whereas the object B can cause a similar change only when the body or the brain is in an initial state  $O_2$  (lower panel). Putting a person's body or brain in one of these two states (either  $O_1$  or  $O_2$ ) will therefore enable either the object A or the object B to cause a brain change. If neorealism is correct, in the upper panel of Figure 4 the person's conscious content will *be* the object A, whereas in the lower panel the person's conscious content will *be* the object B. These two objects being *different* objects, the conscious experiences will indeed differ; A and B are just the different contents of a cross-section.

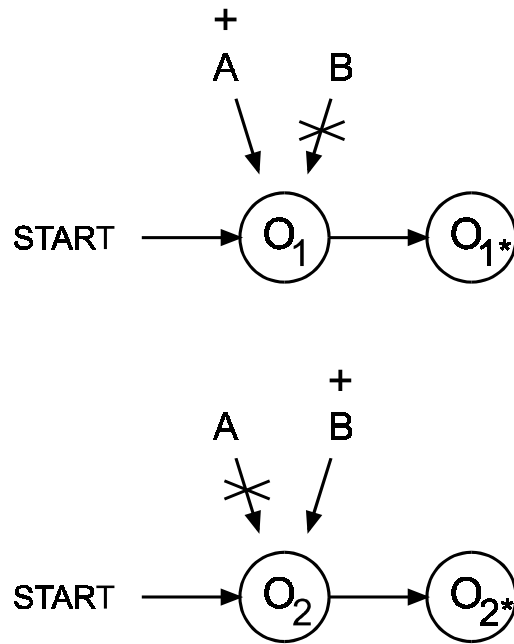


Figure 4. The baseline dependency of environmental action. The environmental object A can cause a change in the body (or the brain) only when the latter is in the starting state  $O_1$  (upper panel), whereas the environmental object B can cause a similar change only when the body or the brain is in the starting state  $O_2$  (lower panel). In each panel, the external object able to cause a state transition in the body or the brain is designated by a + sign.

<sup>7</sup> I speak of environmental *objects* for convenience; however, remember that the entities present in consciousness are conceived of more properly as sets of environmental properties rather than as objects in a substantive sense.

## CONSCIOUSNESS OUTSIDE

The argument from bodily dependency thus stems from a deep confusion. From the (indisputable) fact that at a given time which entity is a person's conscious experience depends on the state of this person's body or brain, it is inferred that *the* entity which is the experience depends on the state of the person's body or brain—and then that this experience is located in the body or the brain. As we have seen, however, this kind of argument is simply absurd (Montague, 1912, p. 295). Which object I designate, for example, depends on the orientation of my finger; but *the* object that I designate does not depend on the orientation of my finger, and the object is certainly not inside my finger. The argument from bodily dependency proves nothing; if the internal view accounts for the facts of bodily dependency, so does neorealism.

### The Argument from Environmental Silence

The second problem on my list is that of environmental silence. Conscious experience can change even in a constant environment, whereas the only thing that changes is one's body or brain. Hence (so the argument goes) these changes of conscious experience cannot take place in the environment. These changes must take place in the body or the brain, as the internal view maintains.

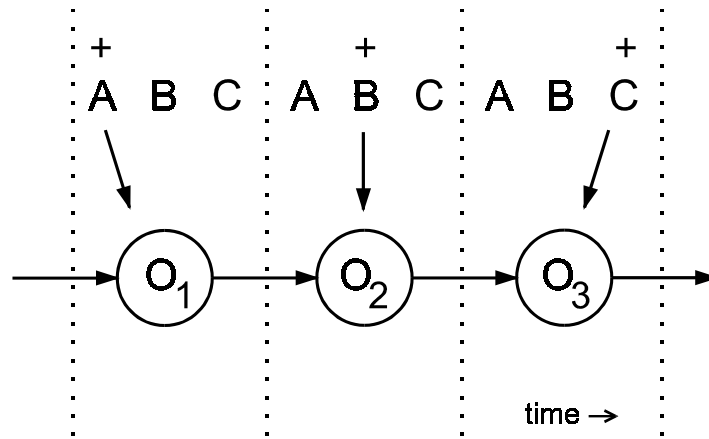


Figure 5. Evolution of a cross-section. The very fact that a given object (say, A) causes the body or the brain to move into a new state disables the current causal relation and enables another one. As the body (or brain) moves from one state to another ( $O_1$ ,  $O_2$ ,  $O_3$ , etc.), the relational property that consists in acting on it is automatically transferred from one object to another (A, B, C, etc.). The active object at any time is designated by a + sign (compare with Figures 1 and 4). Time flows from left to right. Dotted lines separate successive time slices.

Again, the argument is flawed. In this case the flaw arises from the failure to distinguish sectional shifts from state transitions (see The Concept of Cross Section section of this article). If changes in conscious experience were state

transitions, the argument would indeed be correct (pending the falsity of substance dualism); however, neorealism denies that a change of conscious experience is a state transition. According to neorealism, a person's conscious contents are not the successive states of a single thing (which, granted, would have to be the body or the brain) but *different* things (or, more exactly, different sets of properties; see footnote 7) coming one after the other to act on this person, and these things, located in the environment, are not necessarily changing (undergoing a state transition). Thus, in Figure 5, conscious experience does change over time, being A, B, and C in succession, whereas an unchanging conscious experience would have to be A, A, and A, for example. Yet none of the environmental entities A, B, or C changes (undergoes a state transition), and the same holds of the environment as a whole (Figure 5). Environmental silence proves nothing.

### The Argument from Physics

My third question concerned secondary qualities and their place in the environment. How could phenomenal properties be located in the objective environment (as neorealism maintains) if this environment is the subject matter of physics and if physics finds no room for phenomenal properties? Doesn't this physical absence of phenomenal properties from the environment prove neorealism to be wrong and the internal view to be right?

Cognitive psychologists have often used this sort of argument to buttress the notion that perception takes place in the brain (e.g., Robertson, 1987). From the premise that "physics does not disclose knowledge of qualitative properties such as those that occur in 'seeing'" (Weimer, 1974, p. 426), for example, Weimer (1974) concludes that conscious experience is "wholly within our own nervous system" (p. 426).

Yet conceived as a point in favor of the internal view, the argument from physics is obviously flawed because "our own nervous system" is as much a part of the physical world as the environment is. If physics prohibits the environment to entertain phenomenal properties, then physics also prohibits the brain to entertain phenomenal properties. If colors cannot be properties of chairs because chairs follow the laws of physics, colors cannot be properties of brains because brains also follow the laws of physics. Incidentally, a relational approach to perception leaves the problem untouched: If colors cannot emerge from chairs interacting with chairs, then colors cannot emerge from *brains* interacting with chairs, because all follow the laws of physics.

Or are we to believe, perhaps, that chairs are physical but that brains are not? That chairs follow the laws of physics but that brains do not? After arguing on physicalist grounds that there are no stimulus relations in the environment (Robertson, 1987, p. 300), Robertson (1987) suddenly invokes various "levels of analysis" of reality to legitimize the postulation of a "cognitive" brain (p. 301). It is this cognitive brain that is supposed to construct stimulus relations and "impose" them on the environment. If this convenient concept of "level of analysis" is appropriate to the brain, however, why not apply it to the environment in the first

place and assume that the latter, after all, does contain stimulus relations? If the argument from physics applies to the environment, then the argument from physics applies to the brain. If the brain is exempt from the argument, so is the environment.

Again we reach a stalemate. Now the entire argument might be viewed as an argument for dualism (or better, pluralism), not the internal view, for if physics does not include the properties we experience daily, then some pluralism of properties is called for. The neorealist approach sketched in Figure 3 fills the bill by appealing to third-order properties of the environment. Because physicists study material systems that usually differ from us,<sup>8</sup> it is perhaps not surprising that physics fails to mention the environmental properties to which we in fact react.

### **The Argument from Illusion**

Any adequate theory of consciousness should accommodate illusion or error, the fourth issue on my list. Alleged discrepancies between conscious experience and the “real world” (or some version of it) have led psychologists and philosophers to reject direct realism and to endorse the internal view (Parker, 1967). The argument from illusion thus moves from one premise to one conclusion. The premise is that the environment does not (or cannot) contain nonveridical perceptions such as mental images or hallucinations. The conclusion is that neorealism is false and that the internal view is correct.

Irrespective of any refutation of neorealism, however, the argument from illusion does not support the internal view at all; the conscious occurrences that proponents of the internal view cannot find in a person’s environment *cannot be found in this person either*. When I imagine my grandmother, nothing in my brain has the shape of her hands, yet the shape of her hands exists and *is* present in my experience (whether or not I choose to call this experience a “mental image”); to deny this fact is to deny the very existence of my experience, hence the very existence of the problem that the internal view was supposed to solve. How, then, could my experience be a brain process? The answer is obvious: My experience *cannot* be a brain process, and no amount of fumbling or feigning anesthesia will help.

The suggestions of Place (1956) and Smart (1959) on brain states, in particular, will not help. Place (1956) and Smart (1959) suggested that when we report the mental image of an environmental entity X, we just report that we are in a brain state similar to the brain state we are in when we actually perceive X. The problem for the internal view, however, is not to find a brain state with the features of *another* brain state, even “the brain state we are in when we actually perceive X,” for we are *not* aware of the actual features of these brain states, as Smart himself noted (1959, p. 150). The problem for the internal view is instead to find a

---

<sup>8</sup> Notice, though, that physicists study (temporal) parts of ourselves. Tracking the effects of temporary environmental occurrences into the brain and in real-time implies moving away from the historical properties of the environment (historical properties that remain outside the head) and *their* properties.

brain state with the perceptible features of my experience of X (the shape of my grandmother's hands, for example), but no brain state has such features.

Note that calling the relevant brain states "representations" (e.g., Shepard, 1984) will change nothing to the issue because these putative representations, being in the brain, *will not* have the features that conscious experience has. As Holt (1914) put it:

According to [the representative theory of knowledge], consciousness or knowledge contains almost none. . . of those properties that are actually in knowledge. . . "Oh, but they are represented!" Indeed, and according to your theory by shapeless representations of shape, motionless representations of motion, colourless representations of colour, and odourless representations of odour! Whereas the fact is that my knowledge is neither shapeless, motionless, colourless nor odourless. (pp. 141-142)

Holt's (1914) remark that no theory could "more flagrantly explain away what it professes to explain" (p. 141) is as valid today as it was ninety years ago.

Not only is the internal view false, but if my Figure 3 is correct, the premise of the argument from illusion is false too. According to Figure 3, any perceptible property  $y$  of (a property) of an object is also a perceptible property of (a property) of the path. In veridical perception, the set of properties  $y$  associated with a single object  $E$  acts on us and becomes identical with our conscious experience (a cross-section of the environment). In memories and mental images, the same properties  $y$  that are also properties of (properties of) the path over an extended time scale, again enter consciousness.

Because these properties (see Figure 3) are present in the temporally extended environment, our being aware of them requires no epistemic mediators such as internal or neural representations. A memory of the environment is as objective and external as one's veridical perception, and both are outside the head; the only difference is that the former occupies a broader portion of space-time. The same evidence that is said to show that the brain stores "representations" of past events or objects (e.g.,  $E$  in Figure 3) can just as well be taken as evidence that the brain is a response mechanism for temporally extended properties of the environment (see Gibson, 1979, p. 253; Wilcox & Katz, 1981).

In nonveridical perceptions such as dreams, the environmental properties acting on us are again objective properties  $y$  of (properties of) the environment over an extended time scale; however, these properties (of properties of) the path arise from the latter's including more than one object as past constituents. If the path includes three objects  $E_1$ ,  $E_2$ , and  $E_3$ , for example, then (in agreement with Figure 3) respective higher-order properties such as  $y_1$ ,  $y_2$ , and  $y_3$  will be properties of properties of the path, all present over the entire path and available in the objective environment to act on the dreamer. Because these properties are normally associated with different objects, the resulting bundles are often bizarre, looking like parts of an object (say,  $E_1$ ) combined with parts of another (say,  $E_2$ ). The dynamics of the dream are those of a cross-section (Figure 5): As a particular bundle of environmental properties ( $y_1$ ,  $y_2$ , and  $y_3$ , for example) acts on the brain,



the state of the latter necessarily changes, which disables the causal action of the previous bundle and enables another bundle to act.

Thus, many things conspire to give apparent support to the internal view. Dream contents seem to depend on (and are certainly correlated with) the dreamer's brain state and change in a perfectly constant environment, and they include entities (such as pink elephants) that could not possibly (or so it seems) exist in the environment. Clearly there are no pink elephants in the environment, but there are property bundles such as <pink, elephant-shape>, and it is these temporally extended bundles that appear in dreams. These property bundles ( $y_1, y_2, y_3$ , etc.) are also the bundles of properties of properties that any real pink elephant (or any elephant painted pink) would have. The mechanics of cross-sections takes care of the rest. The hypothesis that dreams are outside the head makes perfect sense.

In fact, this hypothesis makes *more* sense than the internal view, since neorealism, of all theories of consciousness, predicts that the features present in nonveridical perception should be features of the *environment* (Figure 3a). I take it as obvious that neorealism is right. It is the features of the environment in which we live (and not of our brain) that are present in mental imagery and thought. *It is the features of our surroundings, from trees to water to smoke to houses, that day after day and night after night appear in our dreams, our memories, and our mental images.* Neorealism explains why. The internal view cannot.

### **Veridical and Nonveridical Perceptions**

Neorealism explains the nature of the relation between veridical and nonveridical perceptions—the last and most important question on my list—by appealing to temporally extended properties of the environment. Because the environment over an extended time scale has properties of properties that are identical with properties of properties of past objects (Figure 3), any feature  $y$  of the environment of which we are aware in veridical perception can later reappear in a mental image, a hallucination, or a dream.

Admittedly, in normal circumstances we easily distinguish our veridical perceptions from nonveridical ones, but this lack of confusion arises from the fact that usually some properties of the former are missing from the latter. Dreams, for example, tend to lack the coherence of veridical visual perception (as discussed above). As more and more features are shared, however, veridical and nonveridical perceptions become indistinguishable; in hallucination no distinction can be made. The existence of hallucinations (and illusions) that are indistinguishable from veridical perceptions strongly suggests that nonveridical and veridical perceptions are made of the same basic constituents.

The problem knowing what these basic constituents are. Neorealism claims that they are objective properties of (properties of) the environment. This realist stance agrees with common sense with respect to veridical perceptions (e.g.,

Stroud-Drinkwater, 1994) but differs from common sense<sup>9</sup> with respect to nonveridical ones. In contrast, both the internal view and common sense assume that nonveridical perceptions reside somewhere inside the body, probably inside the brain, but because the perceptible features that appear in our nonveridical perceptions can (and do) appear in veridical perception as well, the latter, no less than the former, must reside inside the brain. The internal view ends up implying that we are never aware of *anything* except our brain!

Some proponents of the internal view have endorsed this neural solipsism with enthusiasm;<sup>10</sup> recall Weimer's (1974) claim that the entirety of conscious experience exists inside the brain (p. 426). Is neural solipsism a coherent theory? If the apparent features of the trees I see through the window, for example, are creations of my brain (and not actual features of these trees, assuming that there are indeed such trees), then shouldn't the apparent features of the *brains* I see in neuroscience laboratories similarly be creations of my brain (and not actual features of *these* brains)? But then what *are* the actual features of these brains?

Of the proponents of the internal view, Feigl (1958) has been most explicit on this troubling issue. Feigl (1958) points out that whenever he locates experience into a "brain," by "brain" he does *not* mean the "appearance of the gray mass of the brain as one perceives it when looking into an opened skull" (p. 454). This appearance is not the physical brain in which Feigl wants to locate the totality of his conscious experience. No, this "appearance" is merely a phenomenal brain located, like other phenomenal objects such as phenomenal tables and chairs, in his physical brain.

Without this distinction between physical and phenomenal brains, the internal view would collapse. For no brain, physical or otherwise, could be a proper part of itself; indeed, according to Feigl (1958) all sorts of phenomenal objects (besides phenomenal brains) are inside our physical brain. Furthermore, there are parts of our conscious experience, such as phenomenal tables and chairs, that are not part of any phenomenal brain, as Feigl himself emphasizes (1958, p. 455). So Feigl's physical brain is not a perceptible, phenomenal brain (p. 454). Neither is this physical brain located physically where and only where perceptible brains (as opposed to perceptible tables and chairs) are located; all sorts of phenomenal objects are located physically in this physical brain. This physical brain is not even located *phenomenally* where phenomenal brains are located because the assumed physical brain has no phenomenal properties, hence no phenomenal location.

So what is this physical brain that includes all our perceptible world and exhibits none of the features of any perceptible brain? I have no idea, and I think nobody has. The notion of a physical brain in Feigl's theory seems unintelligible to me. Instead of a "physical brain," Feigl (1958) might just as well have postulated a

---

<sup>9</sup> Common sense or some parochial version of it. In some cultures dreams are conceived as literal encounters with the environment, for example because the soul leaves the body during sleep and visits its surroundings (Tedlock, 1987, pp. 5-8).

<sup>10</sup> Here it is the internal view's turn to defy common sense. What do you prefer to believe: that dreams and hallucinations exist in the environment, or that *everything* you are ever aware of resides inside your brain?

## CONSCIOUSNESS OUTSIDE

“physical cauliflower” devoid of “the appearance of the white mass of a cauliflower as one perceives it when looking at a cauliflower.” In what sense could the postulated “cauliflower” qualify as a cauliflower at all, and how could it comprise, include, exhibit, or present not only phenomenal chairs but phenomenal cauliflowers as well? The notion of a “physical cauliflower” that shares none of the features of perceptible cauliflowers and comprises all that we are ever aware of is unintelligible or vacuous. So is Feigl’s notion of a “physical brain.”

Admittedly, Feigl’s (1958) story does not *seem* vacuous or unintelligible, but only because its readers fail to abide by Feigl’s dichotomy of physical objects and appearances and lapse into identifying phenomenal and physical brains. Feigl (1958) himself could not do otherwise. After calling the distinction between “physical” and “phenomenal” spaces “indispensable” (p. 407), he freely admits that “in our rooms or in a landscape we *perceive* directly at least some of this [physical space]” (p. 408). If we perceive directly at least some of this physical space, presumably we also perceive directly at least some of the brains it contains. Why, indeed, would Feigl (1958) want to locate consciousness in a physical *brain* if we did not perceive physical brains at all? Certainly it is Feigl’s own experience with perceptible, *phenomenal* brains that led him to conclude (perhaps through the argument of bodily dependency and other fallacies) that consciousness resides in physical brains (and not, say, physical cauliflowers).

If phenomenal brains *are* identified with (at least part of) physical brains, however, then it again becomes obvious that the internal view is false because phenomenal/physical chairs (for example) are never located, physically or otherwise, in our phenomenal/physical brain. Yet features of these chairs can be present in consciousness and *are* present whenever we see a chair or imagine a chair. Hence, consciousness is outside the head, and with or without physical brains the internal view is false. To sum up, if the proponents of the internal view claim no correlation between physical brains and phenomenal brains, then their theory is unintelligible. If some minimal correlation is admitted, then their theory is false.

The same seeds of self-destruction are contained in any approach that rejects direct perceptual realism. If ordinary, perceptible chairs, being products of one’s brain or mind or soul, have none of the features that they seem to have, then ordinary, perceptible brains fall in the same trap. Therefore, the very formulation of the internal view undermines the actual evidence about brains that the internal view employs to argue that consciousness is brain-dependent in one way or another (Katz & Frost, 1979). Dualism in its various forms falls prey to similar problems. The only theory of consciousness to avoid them is neorealism.

### **Eight Objections**

The reader probably still has questions whose answers did not appear in the text, or did not appear clearly, so I want to survey possible objections to neorealism (some of them I have actually heard) and indicate what the answers are

or might be. The last question to be discussed is especially important and should provoke further developments in neorealism.

*Objection 1.* “Your assumption that properties are located in space and time seems reasonable, especially when these properties are open to inspection. Clearly properties are neither ghosts nor other-worldly entities; however, the notion that properties could have phenomenal *properties* such as shapes and colors does not make any sense. How could a *property* be green or triangular? *Properties* are neither shaped nor colored. Objects are.”

Reply. Neorealism does not assume that all properties are properties of properties. It does assume that the properties of which we are conscious (colors and shapes, for example) are properties of properties. Absurd, you say, but how could you know? Peering at a nearby table will not help. All that we find in our experience of the environment are perceptible properties, so there is no way of knowing by inspection whether these properties are properties of objects (or portions of space–time) or properties of properties of these objects. Of course, if we were aware not only of the properties of which we are aware but *also* of their bearers, we could easily settle the issue. Unfortunately we are *not* aware of these bearers, so the issue remains wide open; for all that we know, the bearers of phenomenal properties could be properties. No absurdity here.

*Objection 2.* “If we are not aware of the bearers of phenomenal properties, then we are never aware of physical objects as such. Your neorealism reintroduces a gap between physical objects and our consciousness of them. The theory is not as realist as it seems.”

Reply. Neorealism could not account for nonveridical perception without some degree of independence of perceptible properties from physical objects. This (partial) independence is accomplished by letting phenomenal properties be properties of any portion of the environment that has these objects as temporal parts. Thus, the realism of neorealism cannot be *completely* naive. However, the fact remains that in neorealism the properties that are part of consciousness exist in the environment *independent* of being perceived or not. Hence, what is perceived from the environment has the characteristic of complete independence from consciousness that we normally (and correctly) attribute to environmental objects.

*Objection 3.* “Neorealism locates consciousness in the public environment. But this is as wrong as locating consciousness inside the skull or the brain. The fact is that we do not locate our experience anywhere. Experience has no location.”

Reply. From the fact that we do not “locate” our dreams (say) it does not follow that they have no location. Right now, for example, I have no idea where you are, but surely you are somewhere. Neorealism argues that experience (even during dreams and hallucinations) is located in the environment. Furthermore, neorealism has a pretty obvious explanation for why we do not “locate” our experience anywhere. To be aware of the spatial location of an entity A one needs to be aware of spatial relations between A and other entities such as B, which is possible only if A and B are both present in the cross-section of consciousness.

There is one entity, however, the location of which we *cannot* perceive—the cross-section itself. For to be aware of its location we would need to be aware of

the relation of the cross-section to entities outside it, which is of course impossible because we cannot be aware of anything that is not part of the cross-section.

*Objection 4.* “Neorealism cannot be true of dreams. I take it as obvious that dreams exist inside the brain. Any claim to the contrary defies common sense.”

Reply. There is no evidence whatever that dreams take place inside the brain (see The Argument from Bodily Dependency section of this article, for example), and the evidence that they do not is overwhelming because what appears in dreams are constituents of the *environment*. Not even subjective evidence supports the notion that dreams are inside the body; we do not feel our dreams inside ourselves. Rather, we feel ourselves inside our dreams, which is exactly what neorealism predicts (also see footnotes 9 and 10).

*Objection 5.* “Perhaps you are right about dreams. But what of mental images? I do not feel the objects I imagine to be outside me. At least I do not feel them outside me as I feel this table over there to be outside me. Neorealism is false.”

Reply. Admittedly, in veridical perception the objects we are aware of seem to be located in the environment, whereas in imagery they do not, but there is a simple reason for this. When perceiving an object (as opposed to imagining an object) “we are conscious not merely of it but of the space and time which intervenes [sic] between it and ourselves” (Montague, 1912, p. 291). In mental imagery the intervening surfaces are missing. It is not surprising, therefore, that imagined objects do not seem to be in the environment: there is no way that they could be. Conversely, *if* a mental image included some intervening space along with the imagined objects, then the latter *would* seem to be in the environment; there is no way that they could not be. Hallucinations, I assume, are images of this sort: bundles of temporally extended features of the environment (as in Figure 3) that comprise coherent distances to one’s body. Significantly, artists and writers often deny that their experiences are personal creations and instead attribute them to an independent, external reality (see the quotations in Kantor & Smith, 1975, p. 359). No doubt their testimonies will be dismissed as irrelevant or wrong by defenders of the internal view, but the testimonies remain, and they support neorealism.

*Objection 6.* “My consciousness of some features of an object (say, a tree) cannot be identical with these features. For these features, being features of the tree can persist in the environment even if my consciousness of them vanishes. Neorealism is false.”

Reply. One might just as well say that in Figure 1, the illuminated object cannot be identical with A because the searchlight could cease to illuminate A. But why couldn’t the illuminated object *be* A in Figure 1? When dealing with the successive contents of a cross-section, identity is relative to times. Two years ago my address *was* Cubilete 245. Now my address *is* Ermita 1493. The content of a

cross-section (being the value of a function) can be an object A at a particular instant and another object B one second later.<sup>11</sup>

*Objection 7.* “Your theory confuses the objects of our experiences with our having these experiences. Neorealism may be true of the former, but it is false of the latter. Having an experience is a perceptual process or *activity*, and this activity cannot be identical with an environmental object.”

Reply. My theory does not “confuse” the objects of experience with our “having” such experiences. Rather, it identifies the former with constituents of the environment and *denies* the existence of the latter as a separate entity. According to neorealism, perceiving an object consists not of this object and a putative activity of perceiving, but of this object and nothing else. Seeing a chair is a chair or a portion of it—period.

If you doubt that neorealism is correct, look at any nearby object (say, a pencil) and think about what is actually present in your conscious experience. If you manage to find in it a “process” or an “activity” distinct from the very pencil you are looking at, then you must be a remarkable individual. I, for one, have never been aware of any “activity” of having an experience that was not this experience itself. As Moore (1922) acknowledged, experience is “diaphanous” (p. 25); it reveals its object and nothing else (Sartwell, 1995), as neorealism predicts. Certainly, experience does not present us with any of the brain processes which occur inside us when acted upon by an external object, so why believe, aside from our experiences themselves, in “activities” of experiencing that nobody has ever been aware of? As to the claim that on logical grounds our “experiencing” an object cannot be identical with this object itself, see Objection 6.

*Objection 8.* “A cross-section must be defined by reference to a particular relation R (the “criterion” of the cross-section). Different relations, different cross-sections. So what is this particular relation R that defines the cross-section of consciousness? You write that we are conscious of the environmental objects or properties that “act” on us, but how should the relevant actions be defined? Admitting any causal action of the environment will not do. If I lay unconscious on the floor and a heater raises my temperature little by little, the heater acts on me (causes changes in me), but it is not part of my consciousness. Are we to believe, perhaps, that in any causal relation between two entities A and B (say, two water molecules) that B is conscious of A?”

The objection is well founded and poses a crucial problem (Bhattacharya, 1942). On the one hand, a narrow criterion of consciousness (such as causing the neural change XYZ) will pass the test of the heater but raise the question of why

---

<sup>11</sup> The difference between first- and third-person accounts of consciousness is also relevant here. When saying that Arthur is conscious of a particular object, I do more than just mentioning this object. In addition, I assume a particular relation R between Arthur and this object (which relation holds and serves as the criterion of consciousness is discussed in Objection 8). From a third-person perspective, therefore, Arthur’s consciousness comprises the object and this relation as well. From *Arthur’s* perspective, however, the consciousness of the object is identical with the *object* to which Arthur is related (as discussed in the reply to Objection 7). It is this *first-person perspective*, fundamental to conscious experience, that neorealism is designed to address.

## CONSCIOUSNESS OUTSIDE

XYZ, of all conceivable effects of the environment, should be associated with phenomenal experience. Why XYZ and not XYW or XYU? On the other hand, a broad causal criterion that admits any effect of the environment as characteristic of consciousness will eliminate the “why” question, but it will not pass the test of the heater (and others like it). Bhattacharya (1942) asks rhetorically, “Is a barometer conscious of good or bad weather?” (p. 187).

Of the two horns of the dilemma I choose the second. The aim of neorealism is to explain consciousness without making any appeal to special processes; admitting a highly complex criterion XYZ for the cross-section would be a step backward. Thus, I assume that in any causal relation between two entities A and B, B is conscious of A (Montague, 1912). Water molecules *are* conscious of each other. A barometer *is* conscious of the weather (or more exactly, atmospheric pressure).

If you believe that consciousness is a complex psychological process, the idea of attributing consciousness to water molecules seems absurd—but remember, according to neorealism consciousness is *not* a complex psychological process. Consciousness is not even “psychological” or “mental.” A thing’s consciousness is the part of the environment that acts causally on this thing; water molecules and barometers are causally related to the environment no less than we are. Attributing consciousness to barometers (for example) in no way denies the existence of enormous psychological differences between barometers and us. There *are* psychological differences between barometers and us, but the difference is not that we are conscious and that they are not. The difference lies in *what* we are conscious of and what we do about it. Presumably barometers are not conscious of much and do not do much either, a fact we can establish by studying how barometers behave and what their behavior depends on.

A strict causal criterion of consciousness, in turn, has crucial implications for what we *are*. In the few lines that follow I will barely sketch these implications; the direction in which they point is relatively clear, but filling the details will prove a challenging task for neorealism. Assume again that I lay completely unconscious on the floor and that a heater raises my temperature. If I am *not* conscious of the heater, then (granting the validity of a strict causal criterion of consciousness) it must be because *the heater does not act on me*, appearances to the contrary notwithstanding. The heater acts on my body or on my brain; but the heater does not act on *me*. My *body* is conscious of the heater, but *I* am not.

If I am not my body, then what am I? We can get an idea of if by examining the criteria of consciousness used in the literature. These criteria are clearly *behavioral* (e.g., Allport, 1994). They may involve, for example, a readiness for integrated action derived from, and in turn changing, previously established causal relations between environment and behavior (a change commonly discussed in terms of “memory” or “learning”). These historical factors are absent from the example of the heater: The effect of the heater on my body (increasing its temperature) does not change and does not depend upon previously established causal relations between the environment and my behavior.

If these suggestions are correct (and clearly they will need further refinement), then we know why, lying *unconscious* on the floor, I am not aware of the heater that raises my temperature. I lay *unconscious* on the floor to the extent that current environmental effects on my body do not depend upon, and do not restructure or change, previously established causal relations between environmental variables and my behavior. Indisputably the heater acts on my body, but the heater does not act on *me* because I *am* a cumulative, modifiable, historically integrated pattern of causal relations between the environment and my body's behavior (Holt, 1915; Rachlin, 1994).

### Conclusion

The recent history of the internal view is one of ever more impressive technical advances in neuroscience combined with the ever-growing suspicion that these advances leave the hard problem of consciousness entirely untouched (see Chalmers, 1995). That we have brains and that these can be studied scientifically is not in dispute; certainly Descartes granted that much. How the qualities that exist in conscious experience could be products or aspects of the *brain* nevertheless remains as incomprehensible as ever. When theoretical puzzles reach such levels of unintelligibility, it is time to reconsider our basic assumptions and look for what went wrong. Dretske (1999) mentions two assumptions that are especially hard to reconcile:

- (1) Conscious perceptual experiences exist inside a person (probably somewhere in the brain)
- (2) Nothing existing inside a person has (or needs to have) the properties one is aware [of] in having these experiences. (p. 103)<sup>12</sup>

I take (2) to be a fact and (1) to be false. Conscious perceptual experiences do *not* exist inside a person. There is no evidence to support the internal view over neorealism because the facts of bodily dependency and environmental silence (among others) can be explained equally well by neorealism through the mechanism of the cross-section.

Moreover, the evidence *for* neorealism and *against* the internal view is enormous. As soon as the internal view is abandoned, all sorts of previously incomprehensible facts start to make sense. Why does conscious experience look so different from the brain? The answer is obvious: Because conscious experience is not in the brain. Why does conscious experience present us with features of the environment? Because conscious experience is in the environment. How can the environmental features that appear in veridical perceptions also reappear in our dreams and hallucinations? Because our dreams and hallucinations reside where our veridical perceptions themselves reside—in the environment. Any perceptible

---

<sup>12</sup> For readability two footnote numbers have been deleted from the quotation.



## CONSCIOUSNESS OUTSIDE

property of an environmental object (Figure 3b) is also a property of the environment over an extended time scale (Figure 3a).

All of these clues converge toward the same answer: Consciousness is outside the head. If this conclusion is correct, then our theoretical efforts should be geared toward a better understanding of our surroundings (see Honderich, 1998; Thomas, 2001; Unger, 1999) and of their relations to behavior. The version of neorealism I have defended relies on historical and temporally extended features of the environment; consciousness (ours at least) is a historical phenomenon. The only difference between a veridical perception and a mental image or a dream lies in the temporal extension of the active environment.

Clearly, neorealism is unconventional in *some* ways. Some price has to be paid for a unified account of veridical and nonveridical perceptions, but the price that the internal view eventually needs to pay is far worse than that of neorealism. By lodging all conscious experiences in the human body or brain, the internal view ends up destroying the very evidence that led to its formulation as a theory of consciousness.

Some proponents of the internal view have recently admitted that they have no idea whatsoever how their theory could be true, while maintaining that it is nonetheless true in ways that we will never be able to grasp (McGinn, 1989). I submit that it is time to move forward.

## References

- Allport, A. (1994). What concept of consciousness? In A. J. Marcel & E. Bisiach (Eds.), *Consciousness in contemporary science* (pp. 159-182). Oxford, England: Clarendon Press.
- Armstrong, D. M. (1978a). *Universals and scientific realism: Vol. I. Nominalism and realism*. New York: Cambridge University Press.
- Armstrong, D. M. (1978b). *Universals and scientific realism: Vol. II. A theory of universals*. New York: Cambridge University Press.
- Armstrong, D. M. (1989). *Universals: An opinionated introduction*. Boulder, CO: Westview.
- Baum, W. M. (1989). Quantitative prediction and molar description of the environment. *Behavior Analyst*, 12, 167-176.
- Bhattacharya, C. (1942). An object view of mind: Holt and Russell. *The Philosophical Quarterly* [India. An organ of the Indian Institute of Philosophy and the Indian Philosophical Congress], 18, 183-205.
- Byrne, A., & Hilbert, D. R. (2003). Color realism and color science. *Behavioral and Brain Sciences*, 26, 3-21.
- Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2, 200-219.
- Dennett, D. C. (1991). *Consciousness explained*. Boston, MA: Little, Brown & Company.
- Donagan, A. (1963). Universals and metaphysical realism. *The Monist*, 47, 211-246.
- Dretske, F. (1999). The mind's awareness of itself. *Philosophical Studies*, 95, 103-124.
- Feigl, H. (1958). The "mental" and the "physical." In H. Feigl, M. Scriven & G. Maxwell (Eds.), *Minnesota studies in the philosophy of science: Vol. 2. Concepts, theories,*

- and the mind-body problem* (pp. 370-497). Minneapolis, MN: University of Minnesota Press.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.
- Heft, H. (2001). *Ecological psychology in context: James Gibson, Roger Barker, and the legacy of William James's radical empiricism*. Mahwah, NJ: Erlbaum.
- Hineline, P. N. (1981). The several roles of stimuli in negative reinforcement. In P. Harzem & M. D. Zeiler (Eds.), *Advances in analysis of behaviour: Vol. 2. Predictability, correlation, and contiguity* (pp. 203-246). Chichester, England: Wiley.
- Holt, E. B. (1912). The place of illusory experience in a realistic world. In E. B. Holt, W. T. Marvin, W. P. Montague, R. B. Perry, W. B. Pitkin, & E. G. Spaulding (Eds.), *The new realism: Coöperative studies in philosophy* (pp. 303-373). New York: Macmillan.
- Holt, E. B. (1914). *The concept of consciousness*. New York: Macmillan.
- Holt, E. B. (1915). *The Freudian wish and its place in ethics*. New York: Henry Holt.
- Holt, E. B., Marvin, W. T., Montague, W. P., Perry, R. B., Pitkin, W. B., & Spaulding, E. G. (1910). The program and first platform of six realists. *Journal of Philosophy, Psychology and Scientific Methods*, 7, 393-401.
- Holt, E. B., Marvin, W. T., Montague, W. P., Perry, R. B., Pitkin, W. B., & Spaulding, E. G. (1912). *The new realism: Coöperative studies in philosophy*. New York: Macmillan.
- Honderich, T. (1998). Consciousness as existence. In A. O'Hear (Ed.), *Current issues in philosophy of mind* (pp. 137-155). Cambridge, England: Cambridge University Press.
- Kantor, J. R., & Smith, N. W. (1975). *The science of psychology: An interbehavioral survey*. Chicago, IL: Principia Press.
- Katz, S., & Frost, G. (1979). The origins of knowledge in two theories of brain: The cognitive paradox revealed. *Behaviorism*, 7(2), 35-44.
- Kim, J. (1998). *Mind in a physical world: An essay on the mind-body problem and mental causation*. Cambridge, MA: MIT Press.
- Locke, J. (1924). *An essay concerning human understanding* (A. S. Pringle-Pattison, Ed.). Oxford, England: Clarendon Press. (Original work published 1690).
- McGilvray, J. A. (1994). Constant colors in the head. *Synthese*, 100, 197-239.
- McGinn, C. (1989). Can we solve the mind-body problem? *Mind*, 98, 349-366.
- McKenzie, J. C. (1968). The externalization of pains. *Analysis*, 28, 189-193.
- Metzger, W. (1972). Critical remarks to J. J. Gibson's conception of "direct" visual perception, i.e., of revived prephysiological realism. In J. R. Royce & W. W. Rozeboom (Eds.), *The psychology of knowing* (pp. 233-236). New York: Gordon and Breach.
- Montague, W. P. (1912). A realistic theory of truth and error. In E. B. Holt, W. T. Marvin, W. P. Montague, R. B. Perry, W. B. Pitkin, & E. G. Spaulding (Eds.), *The new realism: Coöperative studies in philosophy* (pp. 251-300). New York: Macmillan.
- Moore, G. E. (1922). The refutation of idealism. In *Philosophical studies* (pp. 1-30). New York: Harcourt, Brace & Co.
- Nunn, T. P. (1909-1910). Are secondary qualities independent of perception? *Proceedings of the Aristotelian Society*, 10, 191-218.
- Parker, F. H. (1967). Realism—without error? *The Monist*, 51, 224-237.
- Place, U. T. (1956). Is consciousness a brain process? *British Journal of Psychology*, 47, 44-50.

## CONSCIOUSNESS OUTSIDE

- Rachlin, H. (1994). *Behavior and mind: The roots of modern psychology*. New York: Oxford University Press.
- Robertson, L. C. (1987). A cognitive approach to behavior. In S. Modgil & C. Modgil (Eds.), *B. F. Skinner: Consensus and controversy* (pp. 295-305). Barcombe, Lewes, England: Falmer Press.
- Sartwell, C. (1995). Radical externalism concerning experience. *Philosophical Studies*, 78, 55-70.
- Schnaitter, R. (1984). Skinner on the "mental" and the "physical." *Behaviorism*, 12(1), 1-14.
- Shaffer, J. (1963). Mental events and the brain. *Journal of Philosophy*, 60, 160-166.
- Shepard, R. N. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review*, 91, 417-447.
- Sider, T. (1996). All the world's a stage. *Australasian Journal of Philosophy*, 74, 433-453.
- Skinner, B. F. (1974). *About behaviorism*. New York: Knopf.
- Smart, J. J. C. (1959). Sensations and brain processes. *Philosophical Review*, 68, 141-156.
- Smith, A. D. (2000). Space and sight. *Mind*, 109, 481-518.
- Stroud-Drinkwater, C. (1994). The naive theory of colour. *Philosophy and Phenomenological Research*, 54, 345-354.
- Tedlock, B. (1987). Dreaming and dream research. In B. Tedlock (Ed.), *Dreaming: Anthropological and psychological interpretations* (pp. 1-30). New York: Cambridge University Press.
- Thomas, N. J. T. (2001). Color realism: Toward a solution to the "hard problem." *Consciousness and Cognition*, 10, 140-145.
- Tonneau, F. (2001). Equivalence relations: A critical analysis. *European Journal of Behavior Analysis*, 2, 1-128. (Includes commentary.)
- Unger, P. (1999). The mystery of the physical and the matter of qualities: A paper for Professor Shaffer. *Midwest Studies in Philosophy*, 23, 75-99.
- Weimer, W. B. (1974). Overview of a cognitive conspiracy: Reflections on the volume. In W. B. Weimer & D. S. Palermo (Eds.), *Cognition and the symbolic processes* (pp. 415-442). Hillsdale, NJ: Erlbaum.
- Wilcox, S., & Katz, S. (1981). A direct realistic alternative to the traditional conception of memory. *Behaviorism*, 9, 227-239.

