

## Author's Response

### Biology versus computation in the study of consciousness

Ned Block

Department of Philosophy, New York University, New York, NY 10003-6688.  
nb21@nyu.edu; [www.nyu.edu/gsas/dept/philo/faculty/block/](http://www.nyu.edu/gsas/dept/philo/faculty/block/)

**Abstract:** The distinction between phenomenal (P) and access (A) consciousness arises from the battle between biological and computational approaches to the mind. If  $P = A$ , the computationalists are right; but if not, the biological nature of P yields its scientific nature.

The target article focused on the distinction between P (for phenomenal) consciousness and A (for access) consciousness. P = experience. P-conscious qualities are the qualities of experience such as the phenomenal quality of pain or the sensation of red. A state is A conscious if it is poised for direct control of reasoning, speech, and action. The interest in the A/P distinction arises from the battle between two different conceptions of the mind, the *computational* and the *biological*. The computational approach supposes that all of the mind (including consciousness) can be captured with information processing notions such as computation and function in a system.

According to this view (often called functionalism by philosophers), the level of abstraction for understanding the mind is one that allows multiple realizations: just as one algorithm can be realized electrically or hydraulically, the mind can be realized biologically or electronically. The functionalist thinks that the right level of description for characterizing consciousness is the information processing level, not the level of realization of computation, namely, the biological level. The biological approach makes the opposite bet. If  $P = A$ , the functionalist side is right about consciousness. But if consciousness has a biological nature, then the realizations are what count, and we can expect that P and A will diverge.

I hypothesized that cases of P without A exist, but that A without P may not. In all my searching and reviewing suggestions of correspondents, I have seen only one case (in humans) that may well be a case of A without P. Hartmann et al. (1991) describe a case of "inverse Anton's syndrome," an adult whose primary visual cortex had been mostly destroyed, leaving a small island of primary visual cortex. (Thanks to Ralph Adolphs for drawing this to my attention.) This patient cannot discriminate whether the room is dark or illuminated, and he insists that he is blind. If stimuli are presented in the upper right visual field (which projects to the remnant of his primary visual cortex), however, he can recognize faces, facial emotions, and read single words. Yet the patient insists that he does not see anything. When asked how he knows what the word says or whose face it is, he says things like "It clicks" or "I feel it in my mind." There is no sign of hysteria or a psycho-social situation favoring blindness; that is, no reason to believe he is self-deceived. There is damage in the parietal lobes, including the left

inferior parietal region. Milner and Goodale (1995) have proposed that phenomenal consciousness requires ventral stream activity plus attention, and that the requisite attention can be blocked by parietal lesions. So perhaps this is a case of visual access without visual phenomenal consciousness. I hope that readers of this journal can comment on whether this is a genuine case of A without P.

### **R1. Tweaking the definition of “A”**

I certainly agree with **Chalmers**'s point that we should tweak the definition of “A” so as to avoid uninteresting cracks between P and A, that is uninteresting cases of P without A or A without P. Of course, it would be easy to redefine “A” in response to each crack between A and P, resulting in an ad hoc gerrymandered notion of A. Since P *has* an information-processing role (I assume), it would be trivial to claim that there are no cracks between P and *that* role. For example, in the target article I gave an example of P which does not result in A because of lack of attention. Assume that Crick and Koch (1995) are right that visual experience is a matter of activity in pyramidal cells of the lower cortical layers of the visual areas in the back of the head. Suppose further, (as Crick and Koch also suggest) that visual information is put in a position in which it can be used for reasoning and control of behavior by being transmitted to the frontal and prefrontal cortex (in the front of the head). So a conscious event in the visual cortex becomes A-conscious by virtue of transmitting information to the frontal and pre-frontal cortex, and those events in the front are later than the P-events in the back, since it takes time for the information to get to the front. If these ideas are right, a crack would appear to open up between P and A because of the myriad ways in which the information in the back might fail to affect the front in the appropriate way. Now many functionalists (especially the variety that hold functionalism as a conceptual truth) would not be bothered by this, for functionalism is prepared to count an event as P-conscious by virtue of its effects at *other times and places*. In fact, functionalism is prepared to include in the defining role of a P-conscious event processes that don't *actually* happen, but *would* happen under certain conditions. But such features, if used to frame a type of information processing, would make it far from a natural kind of information processing. If the claim that P = A is to be significant, A must be a genuine natural kind that is also a genuine informationprocessing analog of P. It was in this spirit that I defined A so as to rule out the kind of degraded access involved in blindsight as a case of genuine A. The blindsight patient cannot harness the information from the blind field without being told to guess and being given a set of alternatives. So it is best to think of access-consciousness as involving a form of access that is more full-blooded than what exists in blindsight. To rule out blindsight as a case of A without P, I defined a state as A-conscious if it is poised for *rational* control of reasoning, speech, and action. The word “rational” caused a great deal of misunderstanding and I conceded in the response to the first round of commentary that

it was a misleading choice. (I never meant to rule out control that involves poor reasoning – see especially my reply to Kobes in the original response.) “Control” does the job all by itself if understood properly: the information in the blindsight patient’s head about what he saw *influences*, but it does not control. In some publications, I have been defining a state as A-conscious if it is poised for *voluntary* or *direct* control. The blindsight patient’s guesses are voluntary, but the contents do not control the responses in a voluntary manner. They control via an indirect pathway involving guessing.

**Chalmers** proposes defining “A” as “direct availability for global control”; he expands on global control saying that he has in mind especially deliberate behaviors. His “deliberate” corresponds to my “voluntary” (and “rational”), and I think both play the same role in eliminating blindsight. No significant difference so far. Also, my *poised* and Chalmers’s *directly available* seem to do the same job. As I explained (Block 1995), the reason for “poised” was to rule out cases where access requires processing. For example, we all have a belief about what we had for breakfast this morning, but for many readers, that belief was quiescent until reading this sentence. If we make A a totally dispositional concept, a matter of mere accessibility, then quiescent or inactive beliefs will count as A without P. Chalmers’s “directly available” seems designed to do the same job, since as he explains, it is meant to eliminate contents that take some work to retrieve.

**Chalmers** suggests “global control” whereas I specify the kind of control in much more detail, including specifying that the kind of control of reasoning must involve inferential promiscuity, that is, free use of a representation as a premise in reasoning. I don’t see much difference here between Chalmers and me, but it is worth mentioning that the greater specificity does have an advantage. Consider the case I mentioned in the target article of a torture victim who represses the memories of torture. The memories exert a *global* effect on his behavior, causing him to react negatively to places and people that are similar to those involved in the torture; the memories cause slips, affect dreams, and create a global mood. Yet they are not A-conscious. The notion of inferential promiscuity is especially useful in seeing why not.

Now we come to a significant difference. Though our definitions of “A” seem more or less equivalent, there is a crucial difference in interpretation when it comes to thinking about my putative cases of P without A. I gave a number of examples that were designed to exploit the fact that access to a P-content can fail for a variety of reasons, including lack of attention and various forms of blockage. (I mentioned blockage due to repression, information processing limits, fragmentation of the self, and deactivation of centers of reasoning and planning by, for example, anesthesia.) If these cases are genuine cases of P, then they are cases of P without A, because some work would be required to access the blocked representations. Attention would have

to be focused or the blockage removed. **Chalmers** does not dispute that any of my cases are cases of P; rather, he tries to avoid such cases by saying “the information was directly available all along; it simply wasn’t accessed.” But he is trying to have his cake and eat it too, interpreting “directly available” as *poised* for access in order to rule *out* Aconsciousness of what I had for breakfast this morning (before it was mentioned) and as *merely potentially available* for access to rule *in* Aconsciousness in cases of inattention, repression, limits on information process, fragmentation, and anesthesia. The information about what I had for breakfast was potentially available for access only not accessed, yet not phenomenally conscious.

Perhaps **Chalmers** will say that accessing the information about what I had for breakfast this morning involves retrieval from memory – which is why it is not access-conscious – whereas the cases of P without A that I mentioned do not. But what about repression? Accessing the repressed images of torture involves retrieval from memory too, yet Chalmers wants to see them as access-conscious. In short, Chalmers regards the cases of inattention, repression, etc. as phenomenal and access-conscious, so he needs a merely potential notion of access. But he also needs a notion of access that is not merely potential to avoid cases such as a phenomenal appreciation of what I had for breakfast this morning (before it was mentioned above)—cases that are not phenomenal but are potentially available for access (and for phenomenality). No doubt there is some way of distinguishing between the ways that memory is involved in these two cases. But recall that a candidate definition of “A” must be non-ad-hoc as well as a genuine information processing image of P. To build into our definition of “A” a very fine grained condition distinguishing between two ways of accessing memory looks ad hoc, and it raises the question of why that difference involving memory ought to be included in an information processing image of P.

In sum, there are a variety of ways in which access to representations – both P and non-P representations – can be derailed. Anyone who wants to frame a definition of “A” that cuts between the P and non-P cases to avoid cracks between P and A owes us far more than **Chalmers** has provided. Moreover, P comes in a variety of degrees, of phenomenal flavors, and of representational contents. All would seem to affect the causal properties of P-states. But that raises the issue of whether the role of P has any unity apart from its dependence on the intensity, flavor, and representational properties of the P-states that have that role. Consider the kind *feet*, which, let us suppose, is a natural category. Now consider the causal role of feet, what affects them and how and what they affect. Feet are affected by concrete and high-heeled shoes and in turn affect the air-conditioners in gas pedal plants, the breeding of animals from which shoeleather is taken, and the stockprices of companies in the foot-jewelry industry. Is the *role* a natural item apart from the feet that mediate the causal relations? I doubt it, and I would guess that the same point

applies to the role of P-consciousness.

## **R2. Does consciousness have a function?**

The best explanation for the close correlation between P and A is that P is somehow involved in the machinery of A. By contrast, **Chalmers** favors epiphenomenalism. He objects to my claim that P greases the wheels of A on the ground that there is no conceptual entailment from neural stuff to P-consciousness, so there is no contradiction in the idea of a physical duplicate of me who is a zombie, that is, has no P-consciousness. His argument that P-consciousness must be redundant to the causal mechanisms of A-consciousness is that the zombie has the same physical causal machinery of A-consciousness as I do but has no P. Since the causal machinery works the same way with or without P, P does nothing.

But this argument takes mere logical possibilities much too seriously. Mere logical possibilities do not tell us what the real mechanisms are. Magic is logically possible in the sense of not contradictory. The scarecrow of Oz who thinks despite a head of straw is not contradictory, but one cannot move from that to any conclusion about the *actual* mechanisms of thinking. My car does not think and has no P-consciousness, but there is a consistently describable physical duplicate of it that is a sapient and sentient being whose thinking and P-consciousness plays a role in the operation of the car. In my car, the low-gas light goes on via a simple piece of machinery. That machinery is present in the magic “world”, but, in addition, there is another mechanism. In the magic “world”, the fact that the car wants to inform me of the empty tank plays a causal role that is parallel to the physical machinery but nonetheless causally efficacious. Both are causally efficacious; it is a case of overdetermination. The magic “world” is *merely* logically possible only in the sense that there is no contradiction in it. Sapience and sentience are present in one case, absent in the other. But no conclusion can be drawn about sapience and sentience having no effect.

Moving to a somewhat different topic, I agree with **Chalmers** that one can interpret much of the empirical work on consciousness that I criticized as assuming that P = A (that is, that P is a nondispositional state that provides the basis for the A-disposition). So some of this empirical work can be rescued in a post hoc way by making a distinction that the authors themselves did not see. I acknowledged this in the target article. But not all of this work is equally rescuable in that way. In particular, much of the reasoning I was criticizing has problems of the “trivial or false” variety. Witness Searle’s reasoning described in the target article and Crick and Koch’s reasoning that V1 is not part of the neural correlate of consciousness because V1 does not project to frontal cortex and projection to frontal cortex is required for direct control of behavior. This is trivial of A and false (or at least unjustified) for P. (See Block 1996b.) One final point: **Chalmers** notes that model 3 is implausible, apparently assuming that I thought otherwise. I indicated that model 3 might be difficult to refute empirically, not because I thought the model might actually be

right, but rather because of the usual problems with refuting epiphenomenalism. Refutations of the view always end up being more methodological than experimental. So called “simplicity” has to figure very strongly in refutation of such ideas.

### **R3. Consciousness and the self**

Many of the commentators in round one felt that neither A nor P corresponds very well to the intuitive notion of consciousness. The problem was that neither P nor A required that one’s *self* have access to one’s own conscious states. A-consciousness is a purely information-theoretic idea that does not explicitly involve the self. One can speak of one state controlling another without explicitly putting any self in the picture. Although I mentioned various connections with the self in talking about P, none loomed large. Both **Browne** and **Rosenthal** criticize me on this basis, as did many in round one. Church (1995), Harman (1995), Lycan (1995), Kitcher (1995), and Levine (1995) criticized my view explicitly on this ground, but many of the critics in round one were obliquely critical about this. For example, Baars (1995) expanded my P-consciousness as “personal consciousness” and Armstrong (1995) suggested that “A” would be better replaced by “I” for introspection. Officially, **Rosenthal**’s conclusion is that P entails A (unless one or the other is phony), so it is useless to look for cases of P without A. I say that this is his *official* conclusion because *actually* he thinks cases of P without A are completely obvious and uncontroversial. Rosenthal has adopted the misleading strategy of *redefining* both “P” and “A” so that the *P-redefined* entails the *A-redefined*, even though in his view, as in mine, *P does not entail A*. What is misleading about this procedure is that the redefinitions are not made explicit. I confess that my first thought on reading Rosenthal’s reply was that for the reason just mentioned, the disagreement between us was completely verbal. But on reflection, I see that the redefinitions he offers are natural expressions of the clash of our points of view about the importance of self-consciousness, and this clash is an important one to get clear about.

Let me explain. **Rosenthal** and I mean the same thing by “state with phenomenal content.” Phenomenal contents are specific types or categories of experience such as the experience of the sensation of red or the feeling of pain. In my terminology, a state with phenomenal content is just a P-conscious state; I do not distinguish between the two. But Rosenthal rejects the equivalence, *state with phenomenal content = P-conscious state*. His argument starts with the claim that we can have *unconscious* states that belong in one of these P-content categories, for example, an unconscious pain or an unconscious sensation of red in subliminal perception. Such unconscious pains and sensations, he notes, are not A-conscious. By my definitions (which I will be using here), they are cases of P without A, and so we see that Rosenthal accepts P without A in *my* senses of these terms as uncontroversial. Indeed, Rosenthal is much more liberal about P without A than I am. I think that there is only a very remote *possibility* that subliminal perception is P and

thus only a remote possibility that subliminal perception involves P-content without A. Suppose the letter “Q” is flashed too briefly for the subject to report on it, but long enough to influence later choices. Rosenthal seems to assume that such “perceptions” are states with phenomenal content. (I expect this is because he has a very thin notion of phenomenal content. But let us put this issue to one side, accepting with Rosenthal that there are uncontroversial cases of P, in my terms, without A.)

Here is where the basic difference in perspective comes in.

**Rosenthal** holds that these P-without-A states (in my senses of the terms) are not *conscious* states *at all*, for there is “nothing it’s like *for a subject* to be in that state.” In other words, P without A, if it exists, is not *real* consciousness because it need not involve access to the self, to the subject him or herself. So he rejects my notion of P (because it is not what he thinks of as *consciousness*). He holds that the cases of P without A are not real cases of P-*consciousness* without A. Since he thinks access to the self is required for genuine consciousness, he redefines “P” as what we might call “{P + self-access}.”

Okay, so that is part of the story. But we still haven’t seen how a redefined P will necessarily involve A. Does {P + self-access} entail A? No, because as **Rosenthal** notes, A is a purely information-processing notion that *also* involves no connection to the self. So Rosenthal changes my A too to what we might call “{A + self-access},” my A plus the added condition that the self has access to the state. He regards this as the pre-theoretic intuitive sense of A: “Much in Block’s discussion relies on this pretheoretic notion of A-consciousness, rather than the official connection with inference and the control of speech and action.” So Rosenthal’s claim that P entails A amounts to the claim that the redefined P, {P + self-access} entails the redefined A, {A + self-access}. There is more to the story here about why {P + self-access} entails {A + self-access}, but I will not go into it. My point is only that the claim that redefined-P entails redefined-A does not challenge my claim that P without A is at least conceptually possible. Those cases of P that *don’t involve self-access* may be the very cases that do not involve A. For example, in the target article I mentioned a case in which one is having an intense conversation oblivious to a loud noise, even though one has raised the volume of one’s voice to compensate for it. Once one notices the noise, one might realize that one was *hearing it all along*. This is plausibly a case of P without A, and one that does not involve self-access.

**Rosenthal** notes that A is dispositional. Being poised for direct mental and behavioral control is a disposition. But consciousness is not dispositional, he says. For one’s sensation of red to be conscious in his preferred sense, one must oneself be conscious of it, and that is not dispositional. I can agree that that nondispositionality of the P-conscious sensation of red shows that P-consciousness is not A-consciousness. But this does not show that we should reject A in favor of higher order thought. The sensation of red is

also P-conscious. And P consciousness, like higher order thought, is not dispositional.

The key problem with higher order thought as the main or only nondispositional notion of consciousness is that it is too intellectual. Consider the dog in pain mentioned by both Kitcher (1995) and me (Block 1995). Surely a dog with a pain that *hurts* (and therefore exerts direct control of behavior) is in a conscious state in a reasonable, intuitive sense of the term even if the dog has no higher order thought about the pain! And the dog may have a conscious pain in this sense even if it does not have a sufficient grip on the concept of pain or the concept of the self to *have* the thought that “I, myself have a pain.”

The verbal aspect of **Rosenthal**'s point can also be seen by noting that Rosenthal has no complaint against the naturalness or importance of what I call P-consciousness. For him, it is the category *state with phenomenal content*. And he makes no criticism of my notion of A except that, leaving out the self, it is not the “pretheoretic notion of A-consciousness.” The criticism is that neither P nor A deserve to be called categories of “*consciousness*.” So the verbal aspect is that the word “consciousness” should not be applied to them. But there is also an implicit substantive and nonverbal complaint, namely, that I have left out *the main thing* in a notion of consciousness. What is this main thing that I left out? For Rosenthal it is the higher order thought, one state being about another. I agree that higher order thought is important, but I have scanted it because both P and A are more primitive and fundamental. It is P that engenders the famous explanatory gap. We have a promising research program into the nature of thought.

There is no reason to suppose that higher order thought will not yield to it. But there is something else that I might be said to have left out. Armstrong, Baars, Church, Harman, Kitcher, Levine, and Lycan, all first round commentators, mention some sort of connection with the self. I will try to come to grips with this issue, starting with Browne's argument.

**Browne** regards the relations of access to the self as the heart of the intuitive conception of consciousness. He says that reducing this intuitive conception to A-consciousness will simply leave out the intuitive idea of access to the self. Recall that a representation is A-conscious to the extent that it is poised for direct mental and behavioral control. The informational relations involved in direct control of reasoning and action (e.g., informational promiscuity) make no mention of the self and do not in any explicit way clarify the intuitive notion of self-access. So, according to Browne, reducing the intuitive idea to A is not initially promising. The other alternative mentioned by Browne is the idea of reducing the intuitive idea of self-consciousness to P (or perhaps adopting a version of P that includes it). His objection to this idea is that it is unexplanatory. P does not help to explain anything about access to the self.

#### **R4. Deflationism about the self**

My disagreement with **Browne** (and many of the other commentators) hinges on my deflationism about the self.

(See White 1991, for a worked out picture along these lines.) This is not tantamount to being an eliminativist like Hume or Dennett. My view is that the upshot of work in cognitive psychology and cognitive neuroscience is that we (ourselves) are loose federations of centers of control and integration, and for this reason, the intuitive idea of the self as a monolithic integrated entity is an illusion. The conflict between the intuitive conception and the emerging scientific picture was first captured in a convincing manner in Nagel's 1971 paper, "Brain bisection and the unity of consciousness." Nagel argued that the fragmentation observed in split brain patients exists to some degree in normal people, and this challenges our intuitive concept of the self. This sort of idea has been widened and elaborated for many years now by many psychologists and neuropsychologists, for example by Gazzaniga and his colleagues (see also, Dennett 1991). Gazzaniga (1985) tries to explain many ubiquitous cognitive phenomena in terms of the relations among "sub-selves," especially the efforts of some sub-selves to rationalize the behavior of other sub-selves. The most impressive evidence involves cases where knowledge is accessible via one part of the body, but not another. Goodale and Milner (1992) note a double dissociation: some patients cannot describe the orientation of a slot but act appropriately towards it, others show the reverse. Marcel (1993) notes a situation in which blindsight patients can access information better if responding by button-presses than verbally, and better still by eye blinks. Such phenomena are observed not only in brain damaged patients, but also in normals.

So I take it that there is a good scientific basis for what might be called deflationism about the self; regarding the self as a loose federation. This fact is what underlies my disagreement with **Browne** and others. To begin, my notion of A-consciousness *does* involve the self, the only self that *really exists*. The self-consciousness that they hanker after is a mirage. For a representation to be informationally promiscuous, to directly control behavior and speech, *is* for it to be *self-conscious*, given what the self *really* is. The definition of access-consciousness is implicitly relativized to a system. For a representation to dominate activity within that system is for it to be *as self-conscious* as it *can* be. Browne's dissatisfaction with A because it leaves out the self depends on ignoring the relevant science. I said in the target article that one should take intuitions about consciousness very seriously. But these intuitions can only be taken seriously insofar as they do not conflict with scientific fact, and one of the few facts in this area is that the intuitive notion of the self is in large part illusory. So the dissatisfaction with A that many of the critics have expressed, that it does not involve any connection with the self, is a mistake. A does involve self-consciousness in the only sense in which self-consciousness is real.

#### **R5. Is A a kind of consciousness at all?**

**Bringsjord** points out that I waffle on whether A is a kind of consciousness that can exist without P. I expressed some

sympathy (but did not actually endorse) Searle's (1992) and Burge's (1996) claim that a zombie which has no P consciousness has no consciousness of any sort, even if it has the information processing aspects of P. I think they are on to something important about the ordinary notion of consciousness: P is the core and A is conceived of by many of us as a kind of consciousness only against a background of P (as I noted in the replies in the first round). But I have two reasons for seeing A as an independent kind of consciousness. First, I think we all use both A and P to some extent in thinking about consciousness. I refer the reader to the discussion of Searle in the target article. Searle officially denies that A is a kind of consciousness, but I have caught him using "consciousness" in the sense of A. There is also the curious fact I noted in the target article that many people appear to have a concept of consciousness in which A appears to be the core and P is a subsidiary sensory aspect of A that is not even necessary for consciousness. It would be surprising if there were no echo of this in those of us who officially see consciousness as P.

A second reason is that I am less concerned with our ordinary use of "conscious" than with the important scientific issue of the relation between P and its information processing image, namely, A. Are they phenomena of different sorts? Can one exist without the other? The ordinary concepts of consciousness are vague and even if Searle and Burge are right, not too much violence is done to the ordinary concept by treating A without P as a form of consciousness. [As Block (1996b) shows, both A and P are present in the pages of *Nature*.]

### **R6. How representational is P?**

Guzeldere & Aydede find my views on the representational properties of P incoherent. Let me summarize the relevant claims so you can judge for yourselves:

1. Some P-contents are not at all representational, or at least, there is nothing about P that *requires* that P-contents be representational. In the target article, I gave the example of orgasm, but I am not totally sure about it. What I am sure about is that what *matters* about the phenomenal content of orgasm is nothing representational.
2. So P-content is not representational *per se*. (This is just another way of saying that there is nothing about P that requires it to be representational.)
3. Some specific P contents *are* representational *per se*; that is, some specific P-contents have an essential representational aspect. The example I used was the image or visual experience of circles (as opposed to squares). I noted that it is a feature of these P contents that the squares are packable but the circles are not.
4. Some other specific P contents are representational, but not *per se*. According to me, the inverted spectrum thought experiment shows that the P-content that represents red might have represented green.

I think the appearance of incoherence that Guzeldere & Aydede are worried about comes from the ease of confusing the claim that P is not essentially representational with the

claim that some specific P-contents are essentially representational. Art is not essentially representational but some items of art are.

**Gamble** raises the interesting issue of how P could be representational at all. She says P is an intrinsic property whereas representation is relational. But why can't an intrinsic property represent via a relation? Consider the red color of a section of a map. Suppose the redness is an intrinsic property. Still, it can be used to represent altitude. Gamble says that cognitive science must treat P as a representation if it hopes to study it. I don't see why cognitive science can't study the *function* of something that is not representational. No doubt this depends on how one chooses to define "cognitive science." But using "cognitive science" so that Gamble is right, still some other field could study P, call it cognitive biology.

### **R7. Is there a fallacy?**

**Guzeldere & Aydede** say that Schacter's notion of consciousness is more like A than P. But their quotations do not seem to me to support this view. Guzelde & Aydede quote Schacter as speaking of "access to consciousness." Is this supposed to be *access to access-consciousness*? Charity requires rejecting this reading. My view is that consciousness is a mongrel concept containing elements of both P and A. Schacter (and Crick & Koch 1995b) are closer to P than A. But the important point is that by using a single notion of consciousness (that includes elements of both P and A), they end up with a dilemma: triviality or falsehood. This also applies to Crick and Koch (1995a). If they mean A, it is trivial that V1 is not conscious; but if they mean P it is perhaps false. Consider Searle (1992): the epileptics are missing "consciousness" and therefore flexibility. If it is P that is meant, the premise is very likely false. If A is meant, the reasoning is trivial. It is trivial that missing A leads to lack of flexibility because A *includes* flexibility in the relevant sense.

Searle does not make the P/A distinction, but if we make it, we can reinterpret him as saying that P is missing in the epileptics, and that explains the missing A. But even this much charity will not save his argument, since it is very implausible that they are missing P. **Bringsjord** tries to make it plausible that this happens all the time, for example, when we drive "automatically." But this is a very implausible view of automatic driving. Here is an experiment we can all perform. Next time you are going out on a long drive, get your companion to note when you seem to have spaced out and to ask you the following question: "What did you just see?" I will tell you my result: I remember the last car I passed, the last curve in the road and the like. I have been told that pilot work using the Nissan driving simulator (at the Nissan laboratory in Cambridge MA) yields the same result: a moving window of memory of about 30–45 seconds. (Unfortunately, I have been unable to confirm this report.) Bringsjord seems to assume that because there is no long term memory of P, there is no P.

What about Searle's contradiction? **Bringsjord** gives a

stunning application of the principle of charity in explaining away Searle's contradiction. I submit that my diagnosis (switching between using "consciousness" to mean A and P) was far more plausible.

### **R8. Representation and function**

**Gilman** and I are to some extent at cross purposes, as I can explain by distinguishing between representationism and functionalism. Functionalism is the view that the nature of experience can be completely captured by the role of experiences in the mental economy, how they affect other mental states and behavior, and how they are themselves affected by stimulation. Suppose that when I both touch and look at the corner of a cube, I have experiences in the two modalities with the same representational content but different phenomenal feels. One phenomenal feel in sight, another in touch, but no representational difference. This need not disturb a *functionalist*, since there are such large and obvious functional differences between sight and touch. A functionalist has the resources to explain the phenomenal difference. But a *representationist*, by contrast cannot accept experiences that have the same representational content but different phenomenal content, for representationism goes beyond functionalism in trying to cash out all phenomenal character in terms of the ways the world is represented to be. Similarly, a functionalist need not be troubled if the experience of orgasm has no representational content at all, for its functional role (e.g., its motivational role) can serve to distinguish that experience from other experiences.

As **Gilman** notes, I believe in a "nonrepresentational, nonfunctional notion of phenomenal consciousness." Although phenomenal consciousness represents and functions, it cannot be completely accounted for in these terms. However, I did *not* try to argue for the *nonfunctional* part in the target article. The strategy of the target article, was to try to put some of the controversies aside to discuss a distinction (between P and A) that was to some extent at least visible even if my position in those controversies is mistaken. However, I *did* argue that P-content goes beyond the representational. I did not give my strongest argument for that conclusion (namely, the Inverted Earth argument, presented in 1990; 1994; 1996) but I did make some brief remarks in that direction, discussing the impoverished representational content of orgasm (as compared with its truly impressive phenomenal character). And I also had a discussion of sensations with the same representational content in different modalities. My purpose was to head off an identification of P with A, one that surfaced in the commentaries of Armstrong, Dennett, Farah, and Tye, in the first round.

Here's why **Gilman** and I are largely at cross purposes. He argues against my point about the experience of orgasm partly by appealing to its functional properties. He says "Phenomenal contents may vary in a more fine-grained way than natural language labels for those contents, but is such variation obviously nonrepresentational *and nonfunctional?*"

He summarizes my remarks about the experience of orgasm as suggesting that “there is so much to the experience of orgasm that one couldn’t possibly exhaust ‘all that’ with a representational *or functional* account.” And he notes that there is no in-principle problem to “representational *or functional* accounts of the evaluative part of an experience.” (Emphasis added in all these quotations.)

Sure, the evaluative function of the experience of orgasm is entirely immune from my point that this experience is *representationally* impoverished; however, I wasn’t trying to argue against functionalism, but only against the stronger view: representationism.

We are not entirely at cross purposes, however. **Gilman** does also defend the representationist point of view. For example, he notes correctly that we cannot expect all of representational content to be expressible in natural language; for example, recognition dispositions often constitute a kind of content that is not expressible in English. But are we to take seriously the idea that the phenomenal character of orgasm is exhausted by a kind of *recognition*?

On the face of it, having the orgasm-experience and recognizing it are very different. Perhaps recognizing the experience *changes* the experience somewhat. But surely recognition does not wholly create the experience. (What about the first time?) And there is no plausibility in the idea that an orgasm experience requires any sort of categorization. Couldn’t an animal, or even a person, have something like that experience without the recognition?

#### **R9. P 5 A?**

I argued that just as the concept of water differs from the concept of H<sub>2</sub>O, so the concept of P and A differ. The real question, I suggested, was whether as a matter of empirical fact, just as water 5 H<sub>2</sub>O, so P 5 A. (Since A is dispositional whereas P is not, what this comes to is that all and only P-states have the A role.)

**Pöppel** presents evidence that 30–40 Hz oscillations (each one lasting roughly 30 msec) are the basis of consciousness.

For example, if a type of anesthesia is used that suppresses these oscillations, subjects feel that no time has elapsed when they wake up. (“When does the operation start?”) Types of anesthesia that do not suppress the oscillations promote implicit recall of tapes played under anesthesia.

(Patients exposed to a recording of a Robinson Crusoe story are much more likely to associate Crusoe with “Friday” after the operation; see Schwender et al. 1994). Pöppel mentions another interesting temporal matter: evidence for mechanisms of presemantic automatic sensory integration that take 2–3 seconds. Access to P must take place within such a 2–3 second window. So is the idea this? There are two basic mechanisms of consciousness; the 30–40 Hz oscillations underlie P, and the 2–3 second integration mechanism underlies A. I take it that with mechanisms that differ in their time scale in this way, we could have P without A. For a P event might occur and fade out before the integration required for A can take place.

**Noë** denies that the concepts of P and A differ. He argues

that perception intrinsically involves both P and A. Even if he is right about this, it falls short of the conclusion. Perception could essentially involve two nonidentical things. Moreover, I mentioned a number of nonperceptual cases. Recall the Freudian example of the repressed image of the red room in which the patient was tortured. I argued that the repressed image could be P without being A. (The case is hypothetical, but recall that we are talking about the *conceptual* possibility that P and A come apart.) But Noe sometimes appears to use “perception” to *mean* experience, namely, P. On this interpretation, there is no doubt that experience is intrinsically P. The only issue, then, is whether experience is intrinsically A, the issue of the next to last paragraph of Noe’s comment.

Noe gives two reasons why P contents must be A, but neither applies to nonperceptual cases like the Freudian case. The first is that experience by its nature has a rational import. Surely the repressed image potentially has a rational bearing, but one cannot use it unless it becomes A-conscious. The second is that he doubts that one would credit someone with P unless one were willing to credit the person with A too. But one might have all sorts of indirect evidence of the P content of the red image, including the person’s own testimony *after* the psychotherapy is successful and the image becomes A-conscious. The patient might tell us that once he recovered access to the image, he realized that he had always had the image, but the pain associated with it kept him from acknowledging it even to the point of realizing that he had it or realizing that it showed that he had been tortured. Even if one insists on the latter sort of evidence, there could be a period during which the image was P without being A. (Some models of memory, e.g., the Headed Records view of Morton, 1991, have room for such phenomena.)

**Mangan** agrees that there is a conceptual possibility of P diverging from A, but he is certain that in fact P  $\neq$  A. He seems to think that I argue as follows: a difference in concepts, therefore difference in fact. But that is not my argument. I say that we do not know whether P  $\neq$  A. There is certainly reason to take apparent cases of P without A (and one apparent case of A without P) seriously. Mangan says that research on P is doing well on the assumption that P  $\neq$  A. But is it really doing well when we have no idea how anything physical could have P, when we have proposals that the field seriously considers drawing on quantum mechanics, whose rationale seems to be that both quantum mechanics and consciousness are mysterious? Mangan mentions my analogy: perhaps P is like the liquid in a hydraulic computer, and A is like the computation. P is the hardware implementation of A. Mangan wonders whether P can “completely” implement A. But if the analogy is correct, then we have to wonder whether there are other implementations of A, just as a given computation may be realized electrically instead of mechanically. There can be hydraulic fluid without the hydraulic computer and an electronic version of the computer without any fluid. How

does Mangan rule out the analogous possibilities in the case of P and A?

**Bogen** wonders whether the right hemisphere might have A without P. He is sure it has A, and if his theory of P in terms of the ILN is right, it has P too. Perhaps some reader can shed more light on the issue. On dreaming, Bogen agrees with Revonsuo (first round) that dreams may be P without A. In dreaming, one's representations *are* poised to control behavior, but behavioral systems are paralyzed, so there is no behavior. Dream contents are A; so they do not provide a case of P without A.

### **R10. The explanatory gap**

**Van der Heijden et al.** think that the explanatory gap is made to seem wider than it is by assuming that, for example, roses are red and violets are blue. If you suppose that a rose is red, then, according to them, you have to suppose that red is "literally reproduced" in P-consciousness. And if red is "literally reproduced" in P-consciousness, it is no surprise that it seems almost impossible to explain P-consciousness in neural terms. They suggest that we give up the "color-color identity constraint" that insists that we have red both in the world and in the mind. Here is where they go wildly, unbelievably wrong. They say that we should give up the idea that a rose or anything else is ever red. The only redness, they say, is mental redness. But why not hold instead that roses are red, giving up the idea that red is "literally reproduced" in P-consciousness? Why not reject the "color-color identity constraint" by rejecting colors in the mind? Why not construe talk of red in the mind as a misleading way of expressing the fact that P-conscious states *represent* the world as red? And a representation of red need not itself be red (like the occurrences of the word "red" here). This idea is spelled out further in Block (1983) and Tye (1995, Ch. 4).

### **References**

- Armstrong, D. (1995) Perception-consciousness and action consciousness? *Behavioral and Brain Sciences* 18:247–48. [DG, DGa, DMR, rNB]
- Aydede, M. (1995) An analysis of pleasure vis-à-vis pain. Unpublished manuscript, University of Chicago. [GG]
- Baars, B. (1995) Evidence that phenomenal consciousness is the same as access consciousness. *Behavioral and Brain Sciences* 18:249. [rNB]
- (1988) *A cognitive theory of consciousness*. Cambridge University Press. [DJC]
- Bachmann, T. (1995) More empirical cases to break the accord of phenomenal and access-consciousness. *Behavioral and Brain Sciences*. 18:249–51. [DG]
- Berlucchi, G., Aglioti, S., Marzi, C. A. & Tassinari, G. (1995) Corpus callosum and simple visuomotor integration. *Neuropsychologia* 33:923–36. [JEB]
- Block, N. (1980) Troubles with functionalism. In: *Readings in philosophy of psychology*, vol. 1. Harvard University Press. [SB]
- (1983) Mental pictures and cognitive science. *The Philosophical Review* XCII 4:499–541. [rNB]
- (1990) Inverted earth. In: *Philosophical perspectives 4: Action theory and philosophy of mind*, ed. J. Tomberlin. Atascadero. [rNB]
- (1994) Consciousness. In: *A companion to philosophy of mind*, ed. S. Guttenplan. Blackwell. [rNB]
- (1995r) How many concepts of consciousness? *Behavioral and Brain Sciences* 18:272–84. [DG, DMR, rNB]
- (1995t) On a confusion about a function of consciousness. *Behavioral and Brain Sciences* 18:227–87. [DB, JEB, SB, DG, AVDH, EP]
- (1996a) Mental paint and mental latex. In: *Philosophical issues*, ed. E. Villanueva. Atascadero: Ridgeview. [rNB]
- (1996b) How to find the neural correlate of consciousness. *Trends in Neuroscience* 19:2. [NB]
- Bogen, J. E. (1993) The callosal syndromes. In: *Clinical neuropsychology*, 3rd ed., ed. K. M. Heilman & E. Valenstein. Oxford University Press. [JEB]

(1995a) On the neurophysiology of consciousness: 1. Overview. *Consciousness and Cognition* 4:52–62. [JEB]

(1995b) On the neurophysiology of consciousness: 2. Constraining the semantic problem. *Consciousness and Cognition* 4:137–58. [JEB]

Bogen, J. E. & Vogel, P. J. (1962) Cerebral commissurotomy in man: Preliminary case report. *Bulletin of the Los Angeles Neurological Society* 27:169–73. [JEB]

Bringsjord, S. (1992) *What robots can and can't be*. Kluwer. [SB]

(1995) Pourquoi Hendrik est – Is une menace pour la littérature g'en' er 'e'e par ordinateur? (trans. M. Lenoble). In: *Litt' erature et informatique la litt' erature g'en' er 'e'e par ordinateur*, ed. Alain Vuillerain. Artois Presses Universite. (English version forthcoming, MIT Press). [SB]

Burge, T. (1996) Two kinds of consciousness. In: *Consciousness: Philosophical and scientific debates*, ed. N. Block, O. Flanagan & G. G' uzeldere. MIT Press. [rNB]

Chalmers, D. J. (1996) *The conscious mind: In search of a fundamental theory*. Oxford University Press. [DJC, rNB]

Church, J. (1995) Fallacies or analyses? *Behavioral and Brain Sciences* 18:251–52. [GG, rNB]

Cooper, L. A. & Shepard, R. N. (1973) Chronometric studies of the rotation of mental images. In: *Visual information processing*, ed. W. G. Chase. Academic Press. [SB]

Corballis, M. C. (1995) Visual integration in the split brain. *Neuropsychologia* 33:937–59. [JEB]

Crick, F. & Koch, C. (1990) Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences* 2:263–75. [DJC]

(1995a) Are we aware of neural activity in primary visual cortex? *Nature* 375:121–23. [rNB]

(1995b) Why neuroscience may be able to explain consciousness (sidebar). *Scientific American* 12(95):92. [rNB]

Delacour, J. (1995) An introduction to the biology of consciousness. *Neuropsychologia* 33:1061–74. [JEB]

Dennett, D. (1991) *Consciousness explained*. Little Brown. [GG]

(1995) The path not taken. *Behavioral and Brain Sciences* 18:252–53. [AN]

Dretske, F. (1969) *Seeing and knowing*. University of Chicago Press. [GG]

(1981) *Knowledge and the flow of information*. MIT Press. [GG]

(1995) *Naturalizing the mind*. MIT Press. [GG]

Farah, M. J. (1994) Visual perception and visual awareness after brain damage: A tutorial overview. In: *Consciousness and unconscious information processing: Attention and performance*, ed. C. Umiltà & M. Moscovitch. MIT Press. [DJC]

(1995) Is consciousness of perception really separable from perception? *Behavioral and Brain Sciences* 18:254–55. [DG]

Gazzaniga, M. (1985) *The social brain*. Basic Books. [rNB]

Gazzaniga, M., Fendrich, R. & Wessinger, C. (1994) Blindsight reconsidered. *Current Directions in Psychological Science* 3(3):93–6. [DG]

Gazzaniga, M. & LeDoux, J. E. (1978) *The integrated mind*. Plenum. [rNB]

Gerstner, G. E. & Fazio, V. A. (1995) Evidence of a universal perceptual unit in mammals. *Ethology* 101:89–100 [EP]

Goodale, M. & Milner, A. D. (1992) Separate visual pathways for perception and action. *Trends in Neuroscience* 15:20–25. [rNB]

Graham, G. (1995) Guilty consciousness. *Behavioral and Brain Sciences* 18(2):255–56. [SB]

G' uzeldere, G. (1997) *The many faces of consciousness*. Ph.D. dissertation, Stanford University. [GG]

Hardin, C. L. (1988) *Colour for philosophers*. Hackett Publishing Company. [AVDH]

Harman, G. (1990) The intrinsic quality of experience. In: *Philosophical perspectives, vol. 4*, ed. J. Tomberlin. Ridgeview. [GG]

(1995) Phenomenal fallacies and confluents. *Behavioral and Brain Sciences* 18:256–57. [rNB, SB, DG, GG]

Hartmann, J. A. et al. (1991) Denial of visual perception. *Brain and Cognition* 16:29–40. [rNB]

Humphrey, N. (1992) *A history of the mind*. Simon & Schuster. [AVDH]

Huxley, T. H. (1866) Lessons in elementary psychology. Quoted in: *A history of the mind*, by N. Humphrey, 1992. Simon & Schuster. [AVDH]

Katz, L. D. (1995) On distinguishing phenomenal consciousness from the representational functions of mind. *Behavioral and Brain Sciences* 18:258–59. [GG]

Kinsbourne, M. (1995) The intralaminar thalamic nuclei: Subjectivity pumps or attention-action coordinators? *Consciousness and Cognition* 4:167–71. [JEB]

Kitcher, P. (1995) Triangulating phenomenal consciousness. *Behavioral and Brain Sciences* 18:259–60. [AVDH, DMR, rNB]

Kobes, B. W. (1995) Access and what it is like. *Behavioral and Brain Sciences* 18:260. [GG]

Kowal, S., O'Connell, D. C. & Sabin, E. J. (1975) Development of temporal patterning and vocal hesitations in spontaneous narratives. *Journal of Psycholinguistic Research* 4:195–207. [EP]

Lackner, J. & Garrett, M. (1973) Resolving ambiguity: Effects of biasing context

in the unattended ear. *Cognition* 1:359–37. [DMR]

Lambek, J. (1961) How to program an infinite abacus. *Canadian Mathematical Bulletin* 4:295–302. [SB]

Levine, J. (1995) Phenomenal access: A moving target. *Behavioral and Brain Sciences* 18:261. [DG, GG, rNB]

Libet, B. (1993) The neural time factor in conscious and unconscious events. In: *Experimental and theoretical studies of consciousness* (Ciba Foundation Symposium 174). Wiley. [DJC]

Lloyd, D. (1995) Access denied. *Behavioral and Brain Sciences* 18(2):261–62. [SB]

Lycan, W. G. (1995) We've only just begun. *Behavioral and Brain Sciences* 18:262–63. [DG, DG, DGa, GG, AVDH, DMR, rNB]

Madler, C. & Pöppel, E. (1987) Auditory evoked potentials indicate the loss of neuronal oscillations during general anaesthesia. *Naturwissenschaften* 75:42–43. [EP]

Mangan, B. (1993a) Dennett, consciousness, and the sorrows of functionalism. *Consciousness and Cognition* 2(1):1–17. [BM]

(1993b) Taking phenomenology seriously: The “fringe” and its implications for cognitive research. *Consciousness and Cognition* 2(2):89–108. [BM]

Marcel, A. J. (1993) Slippage in the unity of consciousness. In: *Experimental and theoretical studies of consciousness*. Wiley. [rNB]

Mates, J., Müller, U., Radil, T. & Pöppel, E. (1994) Temporal integration in sensorimotor synchronization. *Journal of Cognitive Neuroscience* 6:332–40. [EP]

McDowell, J. (1994) The content of perceptual experience. *Philosophical Quarterly* 44:190–205. [AN]

Milner, A. D. & Goodale, M. A. (1995) *The visual brain in action*. Oxford University Press. [rNB]

Morton, J. (1991) Cognitive pathologies of memory: A headed records analysis. In: *Memories, thoughts, and emotions: Essays in honor of George Mandler*, ed. W. Kessen, A. Ortony & F. Craik. Erlbaum. [rNB]

Nagel, T. (1971) Brain bisection and the unity of consciousness. *Synthese* 22:396–413. [AVDH, rNB]

(1974) What is it like to be a bat? *Philosophical Review* 83:435–50. [AVDH]

Natsoulas, T. (1995) How access-consciousness might be a kind of consciousness. *Behavioral and Brain Sciences* 18(2):264–65. [SB]

Pani, J. R. (1982) A functionalist approach to mental imagery. 23rd Annual Psychonomic Society Meeting. [SB]

Pare, D. & Llinás, R. (1995) Consciousness and preconscious processes as seen from the standpoint of sleep-waking cycle neurophysiology. *Neuropsychologia* 33:1155–68. [JEB]

Peterson, L. B. & Peterson, M. J. (1959) Short-term retention of individual items. *Journal of Experimental Psychology* 58:193–98. [EP]

Pöppel, E. (1970) Excitability cycles in central intermittency. *Psychologische Forschung* 34:1–9. [EP]

(1988) *Mindworks: Time and conscious experiences*. Harcourt Brace Jovanovich. (Orig. 1985: *Grenzen des Bewusstseins*, dva, Stuttgart). [EP]

(1989) Taxonomy of the subjective: An evolutionary perspective. In: *Neuropsychology of visual perception*, ed. J. Brown. Erlbaum. [EP]

(1994) Temporal mechanisms in perception. *International Review of Neurobiology* 37:185–202. [EP]

Pöppel, P., Held, R. & Frost, D. (1973) Residual visual function after brain wounds involving the central visual pathways in man. *Nature* 243:295–96. [EP]

Revonsuo, A. (1995) Conscious and nonconscious control of action. *Behavioral and Brain Sciences* 18(2):265–66. [JEB, SB]

Rey, G. (1995) Block's philosophical anosognosia. *Behavioral and Brain Sciences* 18:266–67. [GG]

Rosenthal, D. M. (1986) Two concepts of consciousness. *Philosophical Studies* 49:329–59. [DMR]

(1990) Why are verbally expressed thoughts conscious? Report no.32, Center for Interdisciplinary Research. University of Bielefeld. [DMR]

(1993) Thinking that one thinks. In: *Consciousness*, ed. M. Davies & G. Humphreys. Blackwell. [DMR]

Schleidt, M., Eibl-Eibesfeldt, I. & Pöppel, E. (1987) A universal constant in temporal segmentation of human short-term behavior. *Naturwissenschaften* 74:289–90. [EP]

Schwender, D., Madler, C., Klasing, S., Peter, K. & Pöppel, E. (1994) Anesthetic control of 40 Hz brain activity and implicit memory. *Consciousness and Cognition* 3:129–47. [EP]

Searle, J. (1983) *Intentionality*. Cambridge University Press. [SB]

(1992) *The rediscovery of the mind*. MIT Press. [NB]

Shallice, T. (1972) Dual functions of consciousness. *Psychological Review* 79:383–93. [DJC]

(1988a) *From neuropsychology to mental structure*. Cambridge University Press. [DJC]

(1988b) Information-processing models of consciousness: Possibilities and problems. In: *Consciousness in contemporary science*, ed. A. Marcel & E. Bisiach. Oxford University Press. [DJC]

Shepard, R. N. (1995) What is an agent that it experiences P-consciousness? And

what is P-consciousness that it moves an agent? *Behavioral and Brain Sciences* 18:267–68. [GG]

Sperling, G. (1960) The information available in brief visual presentations. *Psychological Monographs* 74. [DJC, DMR]

Sperry, R. W. (1974) Lateral specialization in the surgically separated hemispheres. In: *Neuroscience 3rd study program*, ed. F. O. Schmitt & F. G. Worden. MIT Press. [JEB]

Stich, S. P. (1978) Beliefs and subdoxastic states. *Philosophy of Science* 45:499–518. [DB]

Tye, M. (1995a) Blindsight, orgasm and representational overlap. *Behavioral and Brain Sciences* 18:268–69. [DG]

(1995b) *Ten problems of consciousness*. MIT Press. [GG, rNB]

Vollrath, M., Kazenwadel, J. & Krüger, H.-P. (1992) A universal constant in temporal segmentation of human speech. *Naturwissenschaften* 79:479–80. [EP]

White, S. (1991) *The unity of the self*. MIT Press. [rNB]

Zaidel, E. (1978) Concepts of cerebral dominance in the split brain. *Cerebral correlates of conscious experience*, ed. Buser & Rougeul-Buser. Elsevier/North-Holland Biomedical Press. [JEB]

Zaidel, E., Zaidel, D. W. & Bogen, J. E. (1996) Disconnection syndrome. In: *The Blackwell dictionary of neuropsychology*, ed. J. G. Beaumont, R. Keneally & M. Rogers. Blackwell. [JEB]

(1995) *Behavioral and Brain Sciences*. Editorial commentary BBS 18:272. [SB]