BEHAVIORAL AND BRAIN SCIENCES (2007) **30**, 481–548
Printed in the United States of America
doi: 10.1017/S0140525X07002786

Consciousness, accessibility, and the mesh between psychology and neuroscience

Ned Block
Department of Philosophy, New York University, New York, NY 10003
ned.block@nyu.edu

Abstract: How can we disentangle the neural basis of phenomenal consciousness from the neural machinery of the cognitive access that underlies reports of phenomenal consciousness? We see the problem in stark form if we ask how we can tell whether representations inside a Fodorian module are phenomenally conscious. The methodology would seem straightforward: Find the neural natural kinds that are the basis of phenomenal consciousness in clear cases – when subjects are completely confident and we have no reason to doubt their authority – and look to see whether those neural natural kinds exist within Fodorian modules. But a puzzle arises: Do we include the machinery underlying reportability within the neural natural kinds of the clear cases? If the answer is "Yes," then there can be no phenomenally conscious representations in Fodorian modules. But how can we know if the answer is "Yes"? The suggested methodology requires an answer to the question it was supposed to answer! This target article argues for an abstract solution to the problem and exhibits a source of empirical data that is relevant, data that show that in a certain sense phenomenal consciousness overflows cognitive accessibility. I argue that we can find a neural realizer of this overflow if we assume that the neural basis of phenomenal consciousness does not include the neural basis of cognitive accessibility and that this assumption is justified (other things being equal) by the explanations it allows.

Keywords: access consciousness; accessibility; change blindness; consciousness; mind/body problem; NCC; phenomenal consciousness; refrigerator light illusion; reportability; unconscious; vegetative state; working memory

1. Introduction

In *The Modularity of Mind*, Jerry Fodor argued that significant early portions of our perceptual systems are modular in a number of respects, including that we do not have cognitive access to their internal states and representations of a sort that would allow reportability (Fodor 1983; see also Pylylshyn 2003; Sperber 2001). For example, one representation that vision scientists tend to agree is computed by our visual systems is one which reflects sharp changes in luminosity; another is a representation of surfaces (Nakayama et al. 1995). Are the unreportable representations inside these modules phenomenally conscious? Presumably there is a fact of the matter. But since these representations are cognitively inaccessible and therefore utterly unreportable, how could we know whether they are conscious or not? It may seem that the appropriate methodology is clear in principle even if very difficult in practice: Determine the natural kind (Putnam 1975; Quine 1969) that constitutes the neural basis of phenomenal consciousness in completely clear cases – cases in which subjects are completely confident about their phenomenally conscious states and there is no reason to doubt their authority– and then determine whether those neural natural kinds exist inside Fodorian modules. If they do, there are conscious within-module representations; if they don't, there are not. But should we include the machinery underlying reportability *within* the natural kinds in the clear cases? Apparently, in order to decide whether cognitively inaccessible and therefore unreportable representations inside modules are phenomenally conscious, we have to have decided already whether phenomenal consciousness includes the cognitive accessibility underlying reportability. So it looks like the inquiry leads in a circle. I will be calling this problem "the methodological puzzle of consciousness research."

The first half of this article is about the methodology of breaking out of this circle. The second half brings empirical

NED BLOCK is Silver Professor in the Departments of Philosophy and Psychology, and in the Center for Neural Science, at New York University. He was formerly Chair of the MIT Philosophy Program. He is a Fellow of the American Academy of Arts and Sciences and a recipient of fellowships from the Guggenheim Foundation, the National Endowment for the Humanities, the American Council of Learned Societies and the National Science Foundation. He is a past President of the Society for Philosophy and Psychology, a past Chair of the MIT Press Cognitive Science Board, and past President of the Association for the Scientific Study of Consciousness. *The Philosophers' Annual* selected his papers as one of the "ten best" in 1983, 1990, 1995 and 2002. The first of two volumes of his collected papers came out in 2007.

© 2008 Cambridge University Press 0140-525X/08 $40.00

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

evidence to bear on actually breaking out of it, using the principle that other things being equal, a mesh between psychology and neuroscience is a reason to believe the theory that leads to the mesh.

2. Two illustrations

Before giving a more precise statement of the methodological puzzle, I'll give two illustrations that are intended to give the reader a feel for it.

Nancy Kanwisher and her colleagues (Kanwisher 2001; Tong et al. 1998) have found impressively robust correlations between the experience of faces and activation at the bottom of the temporal lobe, usually in the subject's right hemisphere in what they call the "fusiform face area." One method that has been used to investigate the neural basis of face perception exploits a phenomenon known as "binocular rivalry" (see Koch 2004, Ch. 16). Presented with a face-stimulus to one eye and a house stimulus to the other, the subject experiences a face for a few seconds, then a house, then a face, and so on. Examination of the visual processing areas of the brain while the face/house perceptual alternation is ongoing, found stronger shifts with the percept in the fusiform face area than in other areas. The fusiform face area lights up when subjects are experiencing seeing a face and not when subjects are experiencing seeing a house, despite the fact that the stimuli are unchanging. The fusiform face area also lights up when subjects imagine faces (O'Craven & Kanwisher 2000).

In highly constrained experimental situations, observers viewing functional magnetic resonance imaging (fMRI) recordings are 85% accurate in telling whether subjects in a scanner are seeing faces or houses (Haynes & Rees 2006). However, Rafi Malach and his colleagues (Hasson et al. 2004) have been able to get similar results from free viewing of movies by correlating activations in a number of subjects (see also Bartels & Zeki 2004).

There has been some dispute as to what exactly the fusiform face area is specialized for, but these issues can be put aside here. (See Grill-Spector et al. 2006, 2007; Kanwisher 2006, 2007; Tsao et al. 2006).

No one would suppose that activation of the fusiform face area all by itself is sufficient for face-experience. I have never heard anyone advocate the view that if a fusiform face area were kept alive in a bottle, that activation of it would determine face-experience – or any experience at all (Kanwisher 2001). The *total* neural basis of a state with phenomenal character C is itself sufficient for the instantiation of C. The *core* neural basis of a state with phenomenal character C is the *part* of the total neural basis that distinguishes states with C from states with other phenomenal characters or phenomenal contents,[1] for example the experience as of a face from the experience as of a house. (The core neural basis is similar to what Semir Zeki [Zeki 2001; Zeki & Bartels 1999] has called an *essential node.*) So activation of the fusiform face area is a candidate for the core neural basis – not the total neural basis – for experience as of a face (see Block 2005; Chalmers 2000; Shoemaker 1981).

For purposes of this target article, I adopt the physicalistic view (Edelman 2004) that consciousness is identical to its total neural basis, rather than John Searle's view that consciousness is determined by but not identical to its neural basis (McLaughlin 1992; Searle 1992). The issue of this article is not physicalism versus dualism, but rather, whether consciousness includes the physical functions involved in the cognitive accessibility that underlies reportability.

What is the total minus core neural basis? That is, what is the neural background required to make a core neural basis sufficient for a phenomenally conscious experience? There is some evidence that there is a single neural background of all experience involving connections between the cortex and the upper brain stem including the thalamus (Churchland 2005; Laureys 2005; Llinás 2001; Llinás et al. 1998; Merker 2007; Tononi & Edelman 1998). This background can perhaps be identified with what Searle (2005) calls the "unified conscious field." Perhaps the most convincing evidence is that disabling connections to the thalamus seems the common core of what different general anesthetics do (Alkire & Miller 2005). Although Merker (2007) does not make the distinction between core and total, he presents evidence that children born pretty much without a cortex can have the conscious field with little or nothing in the way of any conscious contents: that is, they have the total without much in the way of core neural bases.

Nancy Kanwisher (2001) and Dan Pollen (2003; in press) argue that activation of areas of the brain involved in spatio-temporal binding is required for perceptual phenomenology. Of course some states that have phenomenology, for example, emotions and thoughts, are not experienced as spatially located. But Kanwisher and Pollen may be right about temporal aspects of visual experience. Further, Antonio Damasio (1999) and Pollen argue that all experience requires a sense of self, partly based in the posterior parietal lobe. If true, this would be part of the background.

At the risk of confusing the reader with yet another distinction, it is important to keep in mind the difference between a causal condition and a constitutive condition. For example, cerebral blood flow is causally necessary for consciousness, but activation of the upper brainstem is much more plausibly a constitutive condition, part of what it is to be conscious. (What does "constitutive" mean? Among other things, *constituent*: Hydrogen is partially constitutive of water since water is composed of hydrogen and oxygen.) The main issue of this article is whether the cognitive access underlying reportability is a constitutive condition of phenomenal consciousness.

Here is the illustration I have been leading up to. There is a type of brain injury which causes a syndrome known as *visuo-spatial extinction.* If the patient sees a single object on either side, the patient can identify it, but if there are objects on both sides, the patient can identify only the one on the right and claims not to see the one on the left (Aimola Davies 2004). With competition from the right, the subject cannot attend to the left. However, as Geraint Rees has shown in two fMRI studies of a patient identified as "G.K.," when G.K. claims not to see a face on the left, his fusiform face area (on the right, fed strongly by the left side of space) lights up almost as much as when he reports seeing the face (Driver & Vuilleumier 2001; Rees et al. 2000, 2002b). Should we conclude that G.K. has face experience that – because of lack of attention – he does not know about? Or that the

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

fusiform face area is not the whole of the core neural basis for the experience, as of a face? Or that activation of the fusiform face area is the core neural basis for the experience as of a face but that some other aspect of the total neural basis is missing? How are we to answer these questions, given that all these possibilities predict the same thing: no face report?

I will use the phrase "core neural basis of the experience" instead of Frances Crick's and Christof Koch's "NCC," for neural correlate of consciousness. Mere correlation is too weak. At a minimum, one wants the neural underpinnings of a match of content between the mental and neural state (Chalmers 1998; Noë & Thompson 2004).

3. The puzzle

The following is a principle that will be appealing to many (though it is not to me): Whatever it is about a state that makes it unreportable, would also preclude its being phenomenally conscious. We can call this the Phenomenally Conscious → Reportable Principle, or for short, the Phenomenal → Reportable Principle. But how could we test the Phenomenal → Reportable Principle? If what we mean by a "direct" test is that we elicit reports from subjects about unreportable states, then a direct test will always be negative. And it might seem that there could not be an indirect test either, for an indirect test would have to be based on *some* direct method, that is, a method of investigating whether a state is phenomenally conscious independently of whether it is reportable – a method that apparently does not exist.

Here is a brain-oriented version of the point: Suppose empirical investigation finds a neural state that obtains in all cases in which a phenomenally conscious state is reportable. Such a neural state would be a candidate for a core neural basis. Suppose in addition, that we find that the putative core neural basis is present sometimes when the state is unreportable because mechanisms of cognitive access are damaged or blocked. Would that show the existence of unreportable phenomenal consciousness? No, because there is an alternative possibility: that we were *too quick* to identify the core neural basis. Perhaps the supposed core neural basis that we identified is necessary for phenomenal consciousness but not quite sufficient. It may be that whatever it is that makes the state unreportable also makes it unconscious. Perhaps the cognitive accessibility mechanisms underlying reportability are a constitutive part of the core neural basis, so that without them, there cannot be a phenomenally conscious state. It does not seem that we could find any evidence that would decide one way or the other, because any evidence would inevitably derive from the reportability of a phenomenally conscious state, and so it could not tell us about the phenomenal consciousness of a state which cannot be reported. So there seems a fundamental epistemic (i.e., having to do with our knowledge of the facts rather than the facts themselves) limitation in our ability to get a complete empirical theory of phenomenal consciousness. This is the methodological puzzle that is the topic of this article.

Note that the problem cannot be solved by giving a definition of "conscious." Whatever definition one offers of this and other terms, the puzzle can be put in still other terms – there would still be the question, does what it is like to have an experience include whatever cognitive processes underlie our ability to report the experience?

The problem does not arise in the study of, for example, water. On the basis of the study of the nature of accessible water, we can know the properties of water in environments outside our light cone – that is, in environments that are too far away in space and time for signals traveling at the speed of light to reach us. We have no problem in extrapolating from the observed to the unobserved, and even unobservable in the case of water, because we are antecedently certain that our cognitive access to water molecules is not part of the constitutive scientific nature of water itself. In homing in on a core neural basis of reportable episodes of phenomenal consciousness, we have a choice about whether or not to include the aspects of those neurological states that underlie reportability *within* the core neural basis. If we do, then unreportable phenomenally conscious states are ruled out; if we do not, unreportable phenomenally conscious states are allowed. Few scientifically minded people in the twenty-first century would suppose that water molecules are partly constituted by our cognitive access to them (Boghossian 2006), but few would be sure whether phenomenal consciousness is or is not partly constituted by cognitive access to it. It is this asymmetry that is at the root of the methodological puzzle of phenomenal consciousness.

This issue – whether the machinery of cognitive accessibility is a constitutive part of the nature of phenomenal consciousness – is the focus of this target article. I will not mention evidence concerning inaccessible states within Fodorian modules, or whether G.K. has face experience, but I do claim to show that the issue of whether the cognitive accessibility underlying reportability is part of the constitutive nature of phenomenal consciousness can be resolved empirically and that we already have evidence for a negative answer.

I now turn to a consideration of reportability, but first I want to mention one issue that will not be part of my discussion. Readers are no doubt familiar with the "explanatory gap" (Levine 1983; Nagel 1974), and the corresponding "hard problem" of phenomenal consciousness (Chalmers 1996): the problem of explaining why the neural basis of a given phenomenal quality is the neural basis of that phenomenal quality rather than some other phenomenal quality or none at all. No one has any idea what an answer would be, even a highly speculative answer. Is the explanatory gap an inevitable feature of our relation to our own phenomenology? Opinions differ (Churchland 1994; McGinn 1991). I argue that we can make at least some progress on solving the methodological puzzle even without progress in closing the explanatory gap.

I have been talking about consciousness versus reportability, but reportability is not the best concept to use in thinking about the puzzle.

4. Cognitive accessibility versus reportability

Empirical evidence about the Phenomenal → Reportable Principle *seems* unobtainable, but that is an illusion: *that* principle is clearly false even though another closely

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

related principle is problematic. If a locked-in subject loses control of the last twitch, all mental states can become unreportable. There has been progress in using electrodes implanted in the brain, and less intrusively, electroencephalographic (EEG) technology to enable patients to communicate with the outside world. But if the patient is not trained with these technologies before the total loss of control of the body, these technologies may not work. (See the articles on this topic in the July 2006 issue of *Nature*.)

There is a distinct problem with the Phenomenal → Reportable Principle, namely that a person who is not paralyzed may lose all ability to produce or understand language, and so not have the language capacity required for reporting. In some forms of this syndrome (profound global aphasia), subjects clearly have phenomenal states – they can see, they have pain, and they can make clear what they want and don't want in the manner of a pre-linguistic child – but they are totally without the ability to report in any non-extended sense of the term. (Come to think of it, the same point applies to pre-linguistic children and animals.) And if an aphasic *also* had locked-in syndrome, the unfortunate conjunctively disabled person would be doubly unable to report conscious states. But there is no reason to think that conscious states would magically disappear. Indeed, given that aphasia is fairly common and locked-in syndrome, though infrequent, is not rare, no doubt there have been such conjunctive cases.

Of course there can be nonverbal reports: giving a thumbs-up and shaking one's head come to mind. But not every behavioral manifestation of cognitive access to a phenomenal state is a report, except in an uninterestingly stretched version of the term. Reportability is a legacy of behaviorism that is less interesting than it has seemed. The more interesting issue in the vicinity is not the relation between the phenomenal and the reportable, but rather the relation between the phenomenal and the cognitively accessible.

Adrian Owen and colleagues (Owen et al. 2006) report that a patient who, at the time of testing, satisfied the criteria for a vegetative state, responded to requests to imagine a certain activity in a way indistinguishable from normal patients on an fMRI scan. Her premotor cortex was activated upon being asked to imagine playing tennis, and her parahippocampal place area was activated on being asked to imagine walking through rooms in her house. Paul Matthews objected that the brain activity could have been an associative response to the word "tennis," but Owen counters that her response lasted 30 seconds – until he asked her to stop (Hopkin 2006). In an accompanying article in *Science*, Lionel Naccache insists on behavioral criteria for consciousness. He says, "Consciousness is univocally probed in humans through the subject's report of his or her own mental states" and notes that Owen and colleagues "did not directly collect such a subjective report" (Naccache 2006b). But the evidence is that the patient is capable of an intentional act, namely, the act of imagining something described. That should be considered no less an indication – though of course a fallible indication – of consciousness than an external behavioral act. As an editorial in *Nature* suggests, instead of "vegetative state" we should say "outwardly unresponsive." (*Nature*, Editorial 2006).

In the rest of this article, I will be talking about cognitive accessibility instead of reportability. Reportability is a behavioristic ladder that we can throw away.

In previous papers (Block 1995b; 2001; 2005), I have argued that there can be phenomenally conscious states that are not cognitively accessible. (I put it in terms of phenomenal consciousness without access consciousness.) But I am mainly arguing for something weaker here. Cognitive accessibility could be a causally necessary condition of phenomenal consciousness without being a constitutive part of it. Bananas constitutively include CH_2O molecules but not air and light. Still, without air and light, there could be no bananas – they are causally necessary. The focus here is on whether accessibility is constitutively necessary to phenomenal consciousness, not whether it is causally necessary.

5. Why the methodological puzzle matters

I will mention two ways in which it matters whether we can find out whether phenomenal consciousness includes cognitive accessibility. First, if we cannot get evidence about this, we face a fundamental limit in empirical investigation of the neural basis of phenomenal consciousness – we cannot tell whether the putative core neural basis we have found is the neural basis of phenomenal consciousness itself or the neural basis of *phenomenal consciousness wrapped together with the cognitive machinery of access to phenomenal consciousness*.

Second, there is a practical and moral issue having to do with assessing the value of the lives of persons who are in persistent vegetative states. Many people feel that the lives of patients in the persistent vegetative state are not worth living. But do these patients have experiences that they do not have cognitive access to? It is not irrational to regard a rich experiential life – independently of cognitive access to it – as relevant to whether one would want the feeding tube of a loved one removed.

6. Phenomenal consciousness and Awareness

We may suppose that it is platitudinous that when one has a phenomenally conscious experience, one is in some way aware of having it. Let us call the fact stated by this claim – without committing ourselves on what exactly that fact is – the fact that phenomenal consciousness requires Awareness. (This is awareness in a special sense, so in this section I am capitalizing the term.) Sometimes people say Awareness is a matter of having a state whose content is in some sense "presented" to the self or having a state that is "for me" or that comes with a sense of ownership or that has "me-ishness" (as I have called it; Block 1995a).

Very briefly, three classes of accounts of the relation between phenomenal consciousness and Awareness have been offered. Ernest Sosa (2002) argues that all there is to the idea that in having an experience one is necessarily aware of it is the triviality that in having an experience, one experiences one's experience just as one smiles one's smile or dances one's dance. Sosa distinguishes this minimal sense in which one is automatically aware of one's experiences from noticing one's experiences, which is not

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

required for phenomenally conscious experience. At the opposite extreme, David Rosenthal (2005) has pursued a cognitive account in which a phenomenally conscious state requires a higher order thought to the effect that one is in the state. That is, a token experience (one that can be located in time) is a phenomenally conscious experience only in virtue of *another* token state that is about the first state. (See also Armstrong 1977; Carruthers 2000; and Lycan 1996 for other varieties of higher order accounts.) A third view, the "Same Order" view says that the consciousness-of relation can hold between a token experience and *itself*. A conscious experience is reflexive in that it consists in part in an awareness of itself. (This view is discussed in Brentano 1874/1924; Burge 2006; Byrne 2004; Caston 2002; Kriegel 2005; Kriegel & Williford 2006; Levine 2001, 2006; Metzinger 2003; Ross 1961; Smith 1986).

The same order view fits both science and common sense better than the higher order view. As Tyler Burge (2006) notes, to say that one is necessarily aware of one's phenomenally conscious states should not be taken to imply that every phenomenally conscious state is one that the subject notices or attends to or perceives or thinks about. Noticing, attending, perceiving, and thinking about are all cognitive relations that need not be involved when a phenomenal character is present to a subject. The mouse may be conscious of the cheese that the mouse sees, but that is not to say that the mouse is conscious of the visual sensations in the visual field that represent the cheese, or that the mouse notices or attends to or thinks about any part of the visual field. The ratio of synapses in sensory areas to synapses in frontal areas peaks in early infancy, and likewise for relative glucose metabolism (Gazzaniga et al. 2002, p. 642–43). Since frontal areas are likely to govern higher-order thought, low frontal activity in newborns may well indicate lack of higher-order thoughts about genuine sensory experiences.

The relevance of these points to the project of the target article is this: the fact of Awareness can be accommodated by either the same order view or the view in which Awareness is automatic, or so I will assume. Hence, there is no need to postulate that phenomenal consciousness requires cognitive accessibility of the phenomenally conscious state. Something worth calling "accessibility" may be intrinsic to any phenomenally conscious state, but it is not the cognitive accessibility that underlies reporting.

The highly ambiguous term "conscious" causes more trouble than it is worth in my view. Some use the term "conscious" so as to trivially include cognitive accessibility. To avoid any such suggestion I am from here on abandoning the term "phenomenal consciousness" (which I think I introduced [Block 1990; 1992]) in favor of "phenomenology."

In the next section, I discuss the assumption underlying the methodological puzzle, and in the section after, how to proceed if we drop that assumption.

7. Correlationism

Correlationism says that the ultimate database for phenomenology research consists in reports which allow us to find correlations between phenomenal states and features, on the one hand, and scientifically specifiable states and features – namely, neural states and features – on the other. These reports can be mistaken, but they can be shown to be mistaken only on the basis of other reports with which they do not cohere. There is no going beyond reports.

One version of correlationism is stated in David Papineau's (2002) *Thinking about Consciousness*, in which he says:

> If the phenomenal property is to be *identical* with some material property, then this material property must be both necessary and sufficient for the phenomenal property. In order for this requirement to be satisfied, the material property needs to be present in all cases where the human subjects report the phenomenal property – otherwise it cannot be necessary. And it needs to be absent in all cases where the human subjects report the absence of the phenomenal property – otherwise it cannot be sufficient. The aim of standard consciousness research is to use these two constraints to pin down unique material referents for phenomenal concepts. (Papineau 2002, p. 187)

Consider, for example, what an adherent of this methodology would say about patient G.K. mentioned earlier. One kind of correlationist says we have misidentified the neural basis of face experience and so some aspect of the neural basis of face experience is missing. That is, either the activation of the fusiform face area is not the core neural basis for face experience, or, if it is, then in extinction patients some aspect of the total neural basis outside the core is missing. Another kind of correlationist does not take a stand on whether G.K. is having face experience, saying that we cannot get scientific evidence about it.

So there are two versions of correlationism. *Metaphysical* correlationism – the first version just mentioned – says that there is (or can be) an answer to the sort of question I have raised about G.K. and that answer is no. The metaphysical correlationist thinks that the cognitive access relations that underlie the subject's ability to report are a part of what constitutes phenomenology, so there could not be phenomenology without cognitive accessibility (Papineau 1998).

Epistemic correlationism says that G.K. might be having face experience without cognitive accessibility, but that the issue is not scientifically tractable. According to epistemic correlationism, cognitive accessibility is intrinsic to our knowledge of phenomenology but not necessarily to the phenomenal facts themselves. Epistemic correlationism is more squarely the target of this article, but I will say a word about what is wrong with metaphysical correlationism.

Why does the metaphysical correlationist think G.K. cannot be having face experience? Perhaps it is supposed to be a conceptual point: that the very concepts of phenomenology and cognitive accessibility make it incoherent to suppose that the first could occur without the second. Or it could be an empirical point: the evidence (allegedly) shows that the machinery of cognitive accessibility is part of the machinery of phenomenology. I have discussed the conceptual view elsewhere (Block 1978; 1980).

The neuroscientists Stanislas Dehaene and Jean-Pierre Changeux (2004) appear to advocate epistemic correlationism. (References in the passage quoted are theirs but in this and other quotations to follow citations are in the style of this journal.) Dehaene and Changeux write:

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

> We shall deliberately limit ourselves, in this review, to only one aspect of consciousness, the notion of *conscious access* . . . Like others (Weiskrantz 1997), we emphasize *reportability* as a key property of conscious representations. This discussion will aim at characterizing the crucial differences between those aspects of neural activity that can be reported by a subject, and those that cannot. According to some philosophers, this constitutes an "easy problem" and is irrelevant to the more central issues of phenomenology and self-awareness (e.g., Block 1995b). Our view, however, is that conscious access is one of the few empirically tractable problems presently accessible to an authentic scientific investigation. (Dehaene & Changeux 2004, p. 1145–1146)

Kouider et al. (2007) say: "Given the lack of scientific criterion, at this stage at least, for defining conscious processing without reportability, the dissociation between access and phenomenal consciousness remains largely speculative and even possibly immune to scientific investigation" (p. 2028). (*Access-consciousness* was my term for approximately what I am calling "cognitive accessibility" here.)

In a series of famous papers, Crick and Koch (1995) make use of what appears to be metaphysical correlationism. They argue that the first cortical area that processes visual information, V1, is not part of the neural correlate of phenomenology because V1 does not *directly* project to the frontal cortex. They argue that visual representations must be sent to the frontal cortex in order to be reported and in order for reasoning or decision-making to make use of those visual representations. Their argument in effect makes use of the hidden premise that part of the constitutive function of visual phenomenology is to harness visual information in the service of the *direct* control of reasoning and decision-making that controls behavior.

Jesse Prinz (2000) argues for the "AIR" theory, for *attended intermediate representations.* The idea is that "consciousness arises when intermediate-level perception representations are made available to working memory via attention." Because of the requirement of connection to working memory, this is a form of metaphysical correlationism.

David Chalmers (1998) endorses epistemic correlationism. He says,

> Given the very methodology that comes into play here, we have no way of definitely establishing a given NCC as an independent test for consciousness. The primary criterion for consciousness will always remain the functional property we started with: global availability, or verbal report, or whatever. That's how we discovered the correlations in the first place. 40-hertz oscillations (or whatever) are relevant *only* because of the role they play in satisfying this criterion. True, in cases where we know that this association between the NCC and the functional property is present, the NCC might itself function as a sort of "signature" of consciousness; but once we dissociate the NCC from the functional property, all bets are off. (Chalmers 1998)

Victor Lamme (2006) gives the example of the split-brain patient who says he does not see something presented on the left, but nonetheless can draw it with his left hand. There is a conflict between normal criteria for conscious states. Lamme says that "preconceived notions about the role of language in consciousness" will determine our reaction and there is no objective truth about which view is right." He argues for "letting arguments from neuroscience override our intuitive and introspective notion of consciousness," using neuroscientific considerations to motivate us to define "consciousness" as recurrent processing, in which higher areas feed back to lower areas, which in turn feed forward to the higher areas again, thereby amplifying the signal. He doesn't claim the definition is correct, just that it is the only way to put the study of consciousness on a scientific footing. Although Lamme does not advocate correlationism in either its metaphysical or epistemic forms, his view depends on the idea that the only alternative to epistemic correlationism is neurally based postulation.

Often philosophers – Hilary Putnam (1981) and Dan Dennett (1988; 1991) come to mind – argue that two views of the facts about consciousness are "empirically indistinguishable" – and then they in effect conclude that it is better to say that there are no such facts than to adopt epistemic correlationism. One example is Putnam's thought experiment: We find a core neural basis for some visual experience, but then note that if it occurs in the right hemisphere of a split-brain patient, the patient will say he doesn't see anything. If we restore the corpus callosum, the patient may then say he remembers seeing something. But we are still left with two "empirically indistinguishable" hypotheses, that the hypothesis of the core neural basis is correct so the memory is veridical and, alternatively, that the memory is false.

I will give an empirical argument that we can achieve a better fit between psychology and neuroscience if we assume that the perspectives just described are wrong.

8. An alternative to epistemic correlationism

The alternative I have in mind is just the familiar default "method" of inference to the best explanation, that is, the approach of looking for the framework that makes the most sense of all the data, not just reports (Harman 1965; Peirce 1903, Vol. V, p. 171).

The reader may feel that I have already canvassed inference to the best explanation and that it did not help. Recall that I mentioned that the best explanation of all the data about observed water can give us knowledge of unobserved – even unobservable – water. I said that this approach does not apply straightforwardly to phenomenology. The reasoning that leads to the methodological puzzle says that inevitably there will be a choice about whether to include the neural basis of cognitive access within the neural basis of phenomenology. And that choice – according to this reasoning – cannot be made without some way of measuring or detecting phenomenology independently of cognitive access to it. But we don't have any such independent measure. As I noted, there is a disanalogy with the case of water, since we are antecedently certain that our access to information about water molecules is not part of the natural kind that underlies water molecules themselves. But we are not certain (antecedently or otherwise) about whether our cognitive access to our own phenomenology is partly constitutive of the phenomenology. Without antecedent knowledge of this – according to the reasoning that leads to the methodological puzzle – we cannot know whether whatever makes a phenomenal state cognitively inaccessible also renders it non-phenomenal.

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

Here is the fallacy in that argument: The best theory of *all* the data may be one that lumps phenomenology with water molecules as things whose constitutive nature does not include cognitive access to it. To hold otherwise is to suppose – mistakenly – that there are antecedent views – or uncertainties in this case – that are not up for grabs.

Perhaps an analogy will help. It might seem, offhand, that it is impossible to know the extent of errors of measurement, for any measurement of errors of measurement would have to be derived from measurement itself. But we can build models of the sources of measurement error and test them, and if necessary we can build models of the error in the first level models, and so on, stopping when we get a good predictive fit. For example, the diameter of the moon can be measured repeatedly by a number of different techniques, the results of which will inevitably vary about a mean. But perhaps the diameter of the moon is itself varying? The issue can be pursued by simultaneously building models of source of variation in the diameter itself and models of error in the various methods of measurement. Those models contain assumptions which can themselves be further tested.

The puzzle of how it is possible to use measurement itself to understand errors of measurement is not a deep one. As soon as one sees the answer, the problem of principle falls away, although it may be difficult to build the models in practice. I do not believe that the same is true for the methodological puzzle. One reason is the famous "explanatory gap" that I mentioned earlier. There may be reasonable doubt whether the method of inference to the best explanation can apply in the face of the explanatory gap. A second point is that with the demise of verificationism (Uebel 2006), few would think that the nature of a physical magnitude such as length or mass is constitutively tied to our measurement procedures. The mass of the moon is what it is independently of our methods of ascertaining what it is. But verificationism in the case of consciousness is much more tempting – see Dan Dennett's "first person operationism" (Dennett 1991) for a case in point. Lingering remnants of verificationism about phenomenology do not fall away just because someone speaks its name.

The remainder of this article will describe evidence that phenomenology overflows cognitive accessibility, and a neural mechanism for this overflow. The argument is that this mesh between psychology and neuroscience is a reason to believe the theory that allows the mesh. The upshot is that there are distinct mechanisms of phenomenology and cognitive accessibility that can be empirically investigated.

9. Phenomenology overflows accessibility

George Sperling (1960) showed subjects arrays of alphanumeric characters; for example, three rows of four characters, for 50 msec, followed by a blank field. Subjects said that they could see all or almost all of the characters and this has also been reported in replications of the experiment (Baars 1988, p. 15). The phenomenology of a version of the experiment was described by William James in his *Principles of Psychology*: "If we open our eyes instantaneously upon a scene, and then shroud them in complete darkness, it will be as if we saw the scene in ghostly light through the dark screen. We can read off details in it which were unnoticed whilst the eyes were open" (James 1890); and it may be what Aristotle was talking about when he said, "even when the external object of perception has departed, the impressions it has made persist, and are themselves objects of perception" (Aristotle in Ross 1955, 460b).

When Sperling asked subjects to say what letters they had seen, subjects were able to report only about 4 of the letters, less than half of the number of letters they said they could see. (This result was first reported by Cattell [1885] – I am indebted to Patrick Wilken [2001]). Did the subjects really see all or almost all the shapes as they said? Sperling's clever idea was to test whether people really did see all or almost all of the characters and whether the phenomenology persists after the stimulus was turned off by playing a tone soon after the array was replaced by a blank. Subjects were to report the top row if the tone was high, the bottom row if the tone was low, and the middle row in case of an intermediate tone. The result was that subjects could report all or almost all the characters in any given row. Versions of this type of experiment have been done with as many as 32 alphanumeric characters with similar results (Sligte et al. 2008). An attractive picture of what is going on here – and one that I think makes the most sense of the data – is that although one can distinctly see all or almost all of the 9–12 objects in an array, the processes that allow one to conceptualize and identify the specific shapes are limited by the capacity of "working memory," allowing reports of only about 4 of them. That is, the subject has experiences as of specific alphanumeric shapes, but cannot bring very many of them under specific shape or alphanumeric concepts (i.e., representations) of the sort required to report or make comparisons. The subject can bring them under a general concept like "alphanumeric character" – which is why the subjects can report that they have seen an array of alphanumeric characters – but not under the more specific concepts required to identify which alphanumeric character. Interestingly, Sperling found the same results whether he made the exposure of the grid as short as 15 msec or as long as 500 msec.

Sperling's experiment is often described as showing that a "visual icon" persists after the stimulus is turned off. However as Max Coltheart (1980) notes, this term is used ambiguously. In my terms, the ambiguity is between (1) phenomenal persistence and (2) persistence of accessible information concerning the stimulus. Since these are the very notions whose empirical separation is the topic of this article, the term *icon* is especially unfortunate and I will not be using it further.[2]

The idea that one does in fact phenomenally register many more items than are (in a sense) accessible and that the phenomenology persists beyond the stimulus is further tested in a combination of a change "blindness" paradigm with a Sperling-like paradigm (Landman et al. 2003).

First, I will sketch the change "blindness" paradigm. In these experiments, a photograph is presented briefly to subjects, followed by a blank, followed sometimes by an identical photograph but other times by a similar but not identical photograph, followed by another blank. Then

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press



Figure 1. Compare this with Figure 4 without looking at the two figures side by side. There is a difference between the two pictures that can be hard to be aware of, a fact that motivates the appellation (a misnomer in my view) "Change Blindness."

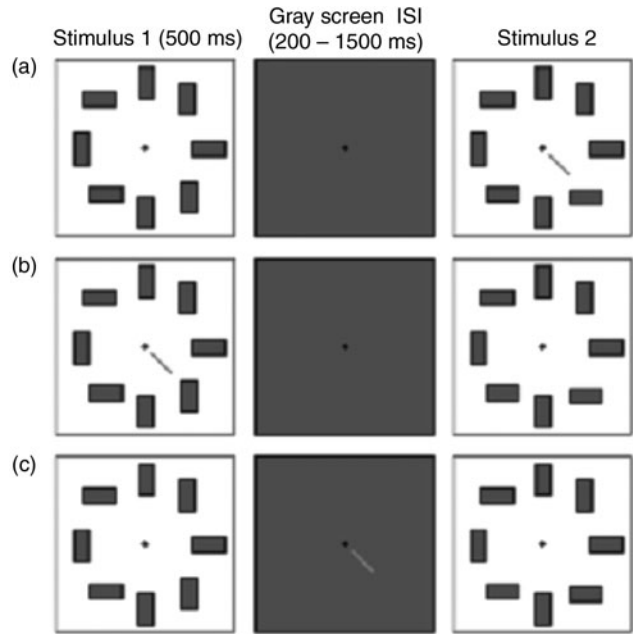


Figure 2. Landman et.al.'s (2003) paradigm combining change "blindness" with Sperling's (1960) experiments on iconic memory. The rectangles are displayed here as line drawings but the actual stimuli were defined by textures. (From Lamme 2003.)

the cycle starts over. One can get the idea by comparing Figure 1 and Figure 4 (p. 494) *without* placing them side by side. When the two photographs differ, they usually differ in one object that changes color, shape, or position, or appears or disappears. The surprising result is that subjects are often unaware of the difference between the two pictures, even when the changed region takes up a good deal of the photographic real estate. Even with 50 repetitions of the same change over and over again, people are often unaware of the change. It is widely agreed that the phenomenon is an attentional one. The items that change without detection have been shown to be items that the subjects do not attend to. But the controversial question – to be discussed later – is whether the phenomenon is one of inattentional blindness or inattentional inaccessibility.[3]

Now for the experiment by Landman et al. (2003). The subject is shown 8 rectangles for half a second as in (a) of Figure 2. There is a dot in the middle which the subject is supposed to keep looking at. (This is a common instruction in visual perception experiments and it has been found, using eye-tracking, that subjects have little trouble maintaining fixation.) The array is replaced by a blank screen for a variable period. Then another array appears in which a line points to one of the objects which may or may not have changed orientation. The subject's task is to say whether the indicated rectangle has changed orientation. In the example shown in Figure 2, there is an orientation change. Using statistical procedures that correct for guessing, Landman et al. computed a standard capacity measure (Cowan's K; see Cowan 2001) showing how many rectangles the subject is able to track. In (a), subjects show a capacity of 4 items. Thus, the subjects are able to deploy working memory so as to access only half of the rectangles despite the fact that in this as in Sperling's similar task, subjects' reported phenomenology is of seeing all or almost all of the rectangles. This is a classic "change blindness" result. In (b), the indicator of the rectangle that may or may not change comes on in the first panel. Not surprisingly, subjects can get almost all of the orientations right: their capacity is almost 8. The crucial manipulation is the last one: the indicator comes on during the blank after the original rectangles have gone off. If the subjects are continuing to maintain a visual representation of the whole array – as subjects say they are doing – the difference between (c) and (b) will be small, and that is in fact what is observed. The capacity measure in (c) is between 6 and 7 for up to 1.5 seconds after the first stimulus has been turned off, suggesting that subjects are able to maintain a visual representation of the rectangles. This supports what the subjects say, and what William James said, about the phenomenology involved in this kind of case. What is both phenomenal and accessible is that there is a circle of rectangles. What is phenomenal but in a sense not accessible, is all the specific shapes of the rectangles. I am taking what subjects say at face value (though of course I am prepared to reject what subjects say if there is evidence to that effect). Whether that is right will be taken up in the section 11.

Subjects are apparently able to hold the visual experience for up to 1.5 seconds – at least "partial report superiority" (as it is called) lasts this long – considerably longer than in the Sperling type experiments in which the production of 3 to 4 letters for each row appears to last at most a few hundred msecs. The difference (Landman et al. 2003) is that the Sperling type of experiment requires a good enough representation for the subjects to actually volunteer what the letters were, whereas the Landman et al. methodology only requires a "same/different" judgment. Yang (1999) found times comparable to Landman et al.'s, using similar stimuli.

In one variation, Landman and colleagues did the same experiment as before but changed the size of the rectangles rather than the orientation, and then in a final experiment, changed either the size or the orientation. The interesting result is that subjects were no worse at

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

detecting changes in either orientation or size than they were at detecting changes in size alone. That suggests that the subjects have a representation of the rectangles that combines size and orientation from which either one can be recovered with no loss due to the dual task, again supporting the subjects' reports.

There is some reason to think that the longest lasting visual representations of this sort come with practice and when subjects learn to "see" (and not "look"). Sligte et al. (2008) found long persistence, up to 4 seconds in a paradigm similar to that of Landman with lots of practice. Others (Long 1980; Yang 1999) have noted that practice in partial report paradigms makes a big difference in subjects' ability to make the visual experience last. These experiments are hard for the subjects and there are large differences among subjects in the ability to do the experiment (Long 1985, Yang 1999); in some cases experimenters have dismissed subjects who simply could not perform the tasks (Treisman et al. 1975). Snodgrass and Shevrin (2006) also find a difference (in a different paradigm) between "poppers" who like to relax and just see and "lookers" who are more active visual seekers.

The main upshot of the Landman et al. and the Sligte et al. experiments (at least on the surface – debunking explanations will be considered later) is along the same lines as that of the Sperling experiment: The subject has persisting experiences as of more specific shapes than can be brought under the concepts required to report or compare those specific shapes with others. They can all be brought under the concept "rectangle" in the Landman et al. experiment or "letter" in the Sperling experiment, but not the specific orientation-concepts which would be required to make the comparisons in Landman et al. or to report the letters in Sperling.

Why are subjects able to gain access to so few of the items they see in the first condition of the Landman et al. experiment (i.e., as described in [a] of Figure 2) and in the Sperling phenomenon without the tones? I am suggesting that the explanation is that the "capacity" of phenomenology, or at least the visual phenomenal memory system, is greater than that of the working memory buffer that governs reporting. The capacity of visual phenomenal memory could be said to be at least 8 to 32 objects – at any rate for stimuli of the sort used in the described experiments. This is suggested by subjects' reports that they can see all or almost all of the 8 to 12 items in the presented arrays, and by the experimental manipulations just mentioned in which subjects can give reports which exhibit the subjects apprehension of all or almost all of the items. In contrast, there are many lines of evidence that suggest that the "working memory" system – the "global workspace" – has a capacity of about 4 items (or less) in adult humans and monkeys and 3 (or less) in infants.

When some phenomenal items are accessed, something about the process erases or overwrites others, so in that sense the identities of the items are not all accessible. However, any one of the phenomenal items is accessible if properly cued, and so in that sense all are accessible. Another sense in which they are all accessible is that the subject knows that he sees them all (or almost all). The upshot is that there is phenomenology without accessibility (Block 1995a), in one sense of the term but not another (Chalmers 1997; Kobes 1995). Of course, there is no point in arguing about which sense of the word "accessibility" to use.

The argument of this target article is thus importantly different from that of earlier works (Block 1995b; 1997), in which I claimed that the Sperling experiment directly shows the existence of phenomenal states that are not cognitively accessible – a conclusion that is not argued for here. In this article, I use the fact of overflow to argue for the conclusion that the machinery of phenomenology is at least somewhat different from the machinery of cognitive accessibility. I then argue that there is a neural realization of the fact of phenomenological overflow – if we assume that the neural basis of phenomenology does not include the neural basis of cognitive access to it as a constituent, and that is a reason to accept that assumption. The neural argument suggests that the machinery of cognitive access is not included in the machinery of phenomenology.

What does it mean to speak of the representational capacity of a system as a certain number of objects? Working memory capacity is often understood in terms of "slots" that are set by the cognitive architecture. One capacity measure that is relevant to phenomenology is the one mentioned above in connection with the Landman et al. (2003) and Sligte et al. (2008) experiments – Cowan's (and Pashler's) K, which I will not discuss further. Another capacity measure relevant to phenomenology is what subjects say about seeing all of a number of items.

There are two significant remaining issues:

1. How do we know that the Sperling, Landman et al., and Sligte et al. effects are not retinal or otherwise pre-phenomenal?
2. How do we know we can believe subjects' reports to the effect that they experience all or almost all of the objects in the Sperling and the Landman et al. experiments? Perhaps subjects confuse potential phenomenology with actual phenomenology just as someone may feel that the refrigerator light is always on because it is on when he looks.

10. Is the effect retinal or otherwise pre-phenomenal?

The persistence of phenomenology is based in the persistence of neural signals. But some neural persistence may *feed* phenomenology rather than *constitute* it, and that creates a difficulty for the view that the capacity of phenomenology is larger than the capacity of working memory. It is no news that the "representational capacity" of the retina is greater than 4! Activations of the retina are certainly not part of the minimal neural basis of visual phenomenology. Rather, activations of the retina are part of the causal process by which that minimal supervenience base is activated. A minimal neural basis is a necessary part of a neural sufficient condition for conscious experience. We know that the retina is not part of a minimal core neural basis, because, for one thing, retinal activation stays the same even though the percept in the binocular rivalry experiments mentioned earlier shift back and forth. It may be said that phenomenological capacity is no greater than that of working memory, the appearance to the contrary deriving from tapping large capacity preconscious representations.

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

The experimental literature points to activations at all levels of the visual system during phenomenal persistence (Coltheart 1980; Di Lollo 1980). However, there are clear differences between the effects of retinal persistence and persistence higher in the cortex.

One of the neatest dissections of phenomenal persistence in low level-vision from that in high-level vision comes from an experiment by G. Engel (1970). (The discussion of this experiment in Coltheart [1980] is especially useful.) Engel used as stimuli a pair of "random dot stereograms" of the sort invented by Bela Julesz in 1959. I can't show you an example that would allow you to experience the effect because the kind Engel used requires a stereo viewer. I can best explain what they are by telling you how they are made.

You start with a grid of, say, 100 by 200 tiny squares. You place a dot in a random location within each square in the grid. The result is something that looks a bit like snow on an old black and white TV – like one of the rectangles in Figure 3. Then you copy the grid dot by dot, but you make a certain change in the copy. You pick some region of it and move every dot in that region slightly to the left, leaving the rest of the dots alone. The right rectangle in the picture is the result of moving a square-shaped region in the middle horizontally. The resulting figure looks to the untrained naked eye just like the first rectangle, but since the visual system is very sensitive to slight disparities, if each eye is presented with one of the two rectangles in a stereo viewer (or if you can "free fuse" the images), the viewer sees a protruding square.

The illusion of depth requires the two rectangles to be presented to the two eyes, but if the presentations to the two eyes are at different times, there will be no loss of the experience of depth so long as the two presentations are not separated by too much time. Suppose the left stereogram is presented to the left eye for 10 msecs, and then the right stereogram is presented to the right eye 50 msecs later. No problem: there is an experience of depth. Indeed, one can present the second stereogram up to 80 msecs later and still get the experience of depth. Monocular persistence, persistence of the signal to a single eye, lasts 80 msecs. Think about the left eye getting its stereogram and the right eye then getting its stereogram 50 msecs later. If there is no independent stereo persistence, the depth experience would expire in another 30 msecs, when the monocular persistence to the left eye runs out. But that does not happen. Depth experience goes on much longer. Engel considered the question of how long one can wait to present another pair of stereograms before the subject loses the experience of depth. He presented sequences of the left stimulus, then the right, then the left, then the right, and so on. If the initial left was followed by a right within 80 msecs, he found that the next left had to come within 300 msecs in order for the subject's experience of depth to be continuous. That is, the experience of depth lasts 300 msecs. The retina is of course completely monocular: each retinal activation depends on input to just one eye. Indeed, substantial numbers of binocular cells are not found in early vision. The conclusion: Depth requires processing in areas of the visual system higher than early vision.



Figure 3. Random-dot stereograms. (Thanks to Júlio M. Otuyama.)

So this experiment shows two kinds of persistence, monocular persistence lasting 80 msecs and binocular persistence lasting 300 msecs; and the binocular persistence is clearly phenomenal since it is a matter of the subject continuing to see depth.

Here is another item of evidence for the same conclusion. There is phenomenal persistence for visual motion, which cannot be due merely to persistence of neural signals in the retina or in early visual areas. Anne Treisman used a display of 6 dots, each of which moved in a circular pathway, either clockwise or counterclockwise (Treisman et al. 1975). Subjects were asked to report whether the motion was clockwise or counterclockwise. Treisman et al. found a partial report advantage much like that in Sperling's experiment. (See also Demkiw & Michaels 1976.)

Why can't this phenomenon be accounted for by neural persistence in the retina or early vision? The point can be understood by imagining a moving dot that goes from left to right on a TV screen. Suppose the screen is phosphorescent so that the images leave a ghost that lasts 1.5 seconds (inspired by the Landman et al. experiment) and suppose that the initial moving dot moves across the screen in 100 msecs. What the viewer will then see is a dot on the left that expands into a line towards the right over a 100 msec period. The line will remain for 1,300 msecs and then it will shrink towards the right to a dot on the right over another 100 msec. The idea of the analogy is to point to the fact that retinal persistence of the signals of a moving object cannot be expected to create phenomenal persistence of the experience of that moving object. Phenomenal persistence has to be based in neural persistence that is a good deal higher in the visual system. As will be discussed later, there is an area in the visual system (V5) that is an excellent candidate for the neural basis of the visual experience of motion.

Perhaps the strongest evidence for cortical persistence comes from the Sligte et al. paper mentioned earlier. There is evidence that the persisting visual experience can be divided into two phases. In the first phase, it is indistinguishable from visual perception. This phase typically lasts at most a few hundred msecs, and often less than 100 msecs (unless subjects are dark-adapted, in which case it lasts longer). The persistence of the experience can be tested by many methods, for example, asking subjects to adjust the timing of a second visual stimulus so that what the subject experiences is a seamless uninterrupted visual stimulus. (See Coltheart [1980] for a description of a number of converging experimental paradigms that

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

measure visible persistence.) In the second phase, the subject has a fading but still distinctly visual experience. The first two phases are of high capacity and disturbed if the test stimulus is moved slightly, and easily "masked" (Phillips 1974, Sligte et al. 2008) by stimuli that overlap in contours with the original stimulus. (Such a mask, if presented at the right lag, makes the stimulus hard to see.)

Sligte et al. (2008) used dark adaptation to increase the strength of the first phase, producing what could be described as a positive afterimage. They also introduced a further variable, two kinds of stimuli: a black/white stimulus and a red/gray isoluminant stimulus in which the foreground and background have the same level of luminance. The idea was to exploit two well-known differences between rods and cones in the retina. Rods are color blind and also have an extended response to stimulation, whereas cones have a brief burst of activity. Rods react to isoluminant stimuli as to a uniform field. The black and white stimulus in dark adaptation will however maximize rod stimulation, producing longer visible persistence without affecting the later working memory representation (Adelson 1978). Sligte et al. found, not surprisingly, that the black and white stimuli produced very strong visible persistences, much stronger than the isoluminant red and gray stimuli when the cue was given just after the array of figures was turned off. (In arrays with 16 items, the subjects had a capacity of 15 for the black and white stimuli but only 11 for the red and gray stimuli.) Here is the very significant result for the issue of retinal persistence versus cortical persistence. A brief flash of light just after the array wiped out this difference. However, when the flash of light was given later after about 1,000 msecs after the array stimulus, it had no effect. Further, a pattern mask did have a huge effect at 1,000 msecs, lowering the capacity to the level of working memory. The flash of light right after the stimulus interferes with retinal persistence, whereas the pattern mask after 1,000 msecs interfered with cortical persistence.

As I mentioned, Sligte et al. used as many as 32 items instead of the 8 of Landman et al. The capacity for the black/white stimulus was close to 32 for the early cue, the capacity of the red/gray stimulus was about 17 and both fell to about 16 for the cue late in the blank space. And both fell further to somewhat over 4 – as in Landman et al. – once the new stimulus came on. If the cue was presented 10 msecs after the first stimulus (the analog of [c] in Figure 2), the black/white stimulus produced greater retention, but if the cue was presented late in the blank period (or once the new stimulus came on as in [a]), the black/white and red/grey stimuli were equivalent. The upshot is that the first phase is very high capacity and is over by 1,000 msecs; the second phase is high capacity and lasts up to 4 seconds; and the third phase has a similar capacity to the working memory phase in Sperling and in Landman et al.

The results mentioned earlier in connection with the Sperling and the Landman et al. experiments are likely to be based in central parts of the visual system, and so are not due to something analogous to "looking again" as in the imaginary dialog presented earlier. However, the question of exactly which central neural activations constitute phenomenology, as opposed to constituting input to phenomenology, is just the question of what phenomenology is in the brain; and of course the topic of this article is whether that can be empirically investigated. So it may seem that I have unwittingly shown the opposite of what I am trying to show, namely, that every attempt to give an empirical answer ends up presupposing an answer. So how can my argument avoid begging the question?

I have three responses. First, the evidence suggests neural persistence at all levels in the visual system. There is no reason to think the phenomenal level is an exception. Second, it appears as if the activations of lower-level vision are relatively brief as compared with the activations of higher-level vision. Third, as mentioned earlier, there is evidence to come that a certain kind of activation of V5 is the core neural basis of the experience of motion. We can see how experimental evidence from phenomenal persistence could dovetail with the evidence outside of memory for V5 as the neural basis for the visual experience of motion. If some version of Treisman's experiment were done in a scanner, my point of view would predict persisting V5 activations of the aforementioned kind. So the issue is not beyond investigation.

11. The Refrigerator Light Illusion

The argument of this article depends on the claim that subjects in the Sperling and the Landman et al. experiments have phenomenal experiences of all or almost all of the shapes in the presented array. One objection is that subjects' judgments to that effect are the result of an illusion in which they confuse potential phenomenology with actual phenomenology. In order to explain this allegation and defend against it, I will first have to say more about cognitive accessibility.

The dominant model of cognitive accessibility in discussions of consciousness – and one that is assumed both in this target article and by Stan Dehaene and his colleagues, the critics who I will be talking about in this section – is a model of broadcasting in a global workspace that started with the work of Bernard Baars (1988; 1997) The idea is closely related to my notion of access consciousness and Dan Dennett's (1993; 2001) notion of "cerebral celebrity" or fame in the brain.[4] Think of perceptual mechanisms as suppliers of representations to consuming mechanisms which include mechanisms of reporting, reasoning, evaluating, deciding, and remembering. There is empirical evidence that it is reasonable to think of perceptual systems as sending representations to a global active storage system, which is closely connected to the consuming systems. Those representations are available to all cognitive mechanisms without further processing. (That's why blindsight "guesses" don't count as cognitively accessible in this sense; further processing in the form of guessing is required to access the representations.) This workspace is also called "working" memory – the word "memory" being a bit misleading because, after all, one can report an experience while it is happening without having to remember it in any ordinary sense of the term.

Dehaene and colleagues (Dehaene et al. 1998; Dehaene & Nacchache 2001) have given impressive evidence that our ability to report our phenomenal states hinges on such a global workspace and that the connection between perception and the workspace lies in long-range neurons in sensory areas in the back of the head which feed forward

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

to the workspace areas in the front of the head. In past publications, I argued for phenomenology without cognitive accessibility (Block 1995a; 1995b; 2001) on the basis of the Sperling experiment. Dehaene and Naccache (2001) replied, making use of the global workspace model:

> Some information encoded in the nervous system is permanently inaccessible (set I_1). Other information is in contact with the workspace and could be consciously amplified if it was attended to (set I_2). However, at any given time, only a subset of the latter is mobilized into the workspace (set I_3). We wonder whether these distinctions may suffice to capture the intuitions behind Ned Block's (Block 1995b; see also Block 2001) definitions of phenomenal (P) and access (A) consciousness. What Block sees as a difference in essence could merely be a qualitative difference due to the discrepancy between the size of the potentially accessible information (I_2) and the paucity of information that can actually be reported at any given time (I_3). Think, for instance, of Sperling's experiment in which a large visual array of letters seems to be fully visible, yet only a very small subset can be reported. The former may give rise to the intuition of a rich phenomenological world – Block's P-consciousness – while the latter corresponds to what can be selected, amplified, and passed on to other processes (A-consciousness). Both, however, would be facets of the same underlying phenomenon. (Dehaene & Naccache 2001, p. 30)

The distinction between I_1, I_2, and I_3 is certainly useful, but its import depends on which of these categories is supposed to be phenomenal. One option is that representations in both categories I_2 (potentially in the workspace) and I_3 (in the workspace) are phenomenal. That is not what Dehaene and Naccache have in mind. Their view (see especially section 3.3.1 of their paper) is that only the representations in I_3 are phenomenal. They think that representations in the middle category (I_2) of potentially in the workspace *seem* to the subject to be phenomenal but that this is an illusion. The only phenomenal representations are those that are *actually* in the workspace. But in circumstances in which the merely potential workspace representations can be accessed at will, they *seem* to us to be phenomenal. That is, the subjects allegedly mistake merely *potential* phenomenology for actual phenomenology.

Importantly, the workspace model exposes a misleading aspect of talk of cognitive accessibility. What it is for representations to be in the workspace (I_3) involves both actuality (sent to the workspace) and potential (can be accessed by consuming mechanisms without further processing). The representations that are actually in the workspace are in active contact with the consuming systems, and the consuming systems can (potentially do) make use of those representations. We might speak of the representations in I_3 (in the workspace) as cognitively accessible in the narrow sense (in which consuming mechanisms make use of what is already there), and representations in the union of I_3 and I_2 as cognitively accessible in the broad sense. It is narrow cognitive accessibility that Dehaene et al. identify with phenomenology. When I speak of phenomenology overflowing cognitive accessibility, I mean that the capacity of phenomenology is greater than that of the workspace – so it is narrow accessibility that is at issue. In the rest of this article, I will be using "cognitive accessibility" in the narrow sense. The thesis of this article is that phenomenology does not include cognitive accessibility in the narrow sense. Here we see that as theory, driven by experiment, advances, important distinctions come to light among what appeared at first to be unified phenomena (See Block & Dworkin 1974, on temperature; Churchland 1986; 1994; 2002, on life and fire).

But what is wrong with the broad sense? Answer: The broad sense encompasses *too much*, at least if a necessary and sufficient condition of phenomenology is at stake. Representations in I_2 can be "amplified if . . . attended to", but of course uncontroversially *unconscious* representations can be amplified too, if one shifts attention to what they represent (Carrasco 2007). So including everything in I_2 in consciousness would be a mistake, a point I made (Block 1997) in response to the claim that consciousness correlates with a certain functional role by Chalmers (1997). No doubt a functional notion that is intermediate between narrow and broad could be framed, but the challenge for the framer would be to avoid ad hoc postulation.

An experimental demonstration that shifting attention affects phenomenology to a degree sufficient to change a sub-threshold stimulus into a supra-threshold stimulus is to be found in a series of papers by Marisa Carrasco (Carrasco et al. 2004) in which she asked subjects to report the orientation of one of a pair of gratings which had the higher contrast. Carrasco presented an attention-attracting dot on one side of the screen or the others that subject was supposed to ignore, slightly before the pair of gratings. She showed that attention to the left made a grating on the left higher in contrast than it would otherwise have been. In subsequent work (Carrasco 2007), Carrasco has been able to show precisely measurable effects of attentional shifts on contrast and color saturation, but not on hue. This alleged conflation of potential phenomenology with actual phenomenology could be called the Refrigerator Light Illusion[5] (Block 2001), the idea being that just as someone might think the refrigerator light is always on, confusing its potential to be on with its actually being on, so subjects in the Sperling and the Landman et al. experiments might think that all the items register phenomenally because they can see any one that they attend to. In the rest of this section, I argue against this allegation.

Let us begin by mentioning some types of illusions. There are neurological syndromes in which cognition about one's own experience is systematically wrong; for example, subjects with anosognosia can complain bitterly about one neural deficit while denying another. And cognitive illusions can be produced reliably in normals (Piattelli-Palmarini 1994). To take a famous example, doctors are more reluctant to recommend surgical intervention if they are told that a disease has a mortality rate of 7% than if they are told it has a survival rate of 93%. Moving to a cognitive illusion that has a more perceptual aspect, much of vision is serial but subjects take the serial processes to be simultaneous and parallel (Nakayama 1990). For example, G. W. McConkie and colleagues (McConkie & Rayner 1975; McConkie & Zola 1979) created an eye-tracking setup in which subjects are reading from a screen of text but only the small area of text surrounding the fixation point (a few letters to the left and 15 to the right) is normal – the rest is perturbed. Subjects have the mistaken impression that the whole page contains normal text. Subjects suppose that the impression

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

of all the items on a page is a result of a single glance, not realizing that building up a representation of a whole page is a serial process. These illusions all have a strong cognitive element.

Are the results from experiments like those of Sperling and Landman et al. the result of cognitive illusions? One reason to think they are not is that the phenomena which the Sperling and the Landman et al. experiments depend on do not require that subjects be asked to access any of the items. It is a simple matter to show subjects arrays and ask them what they see without asking them to report any specific items (as was done first in Gill & Dallenbach 1926). This suggests that the analysis of subjects' judgments in the partial report paradigms as based on cognition – of noticing the easy availability of the items – is wrong. A second point is that cognitive illusions are often, maybe always, curable. For example, the framing illusion mentioned above is certainly curable. However, I doubt that the Sperling and the Landman et al. phenomenology is any different for advocates of the Dehaene view. Third, the sense that in these experiments so much of the perceptual content slips away before one can grab hold of it cognitively does not seem any kind of a cognition but rather is percept-like.

Recall, that in the Sperling experiment, the results are the same whether the stimulus is on for 50 msecs or 500 msecs. Steve Schmidt has kindly made a 320 msec demo that is available on my web site at: http://*www.nyu.edu/gsas/dept/philo/faculty/block/demos/Sperling320msec.mov*.

See for yourself.

The suggestion that the putative illusion has a perceptual or quasi-perceptual nature comports with the way Dan Dennett and Kevin O'Regan describe the sparse representations allegedly revealed by change "blindness" (Dennett 1991; O'Regan 1992).[6] Their idea is that the way it *seems that it seems* is – supposedly – not the way it *actually seems*. They allege not a mismatch between appearance and external reality as in standard visual illusions but rather a mismatch between an appearance and an appearance of an appearance. We could call this alleged kind of illusion in which the introspective phenomenology does not reflect the phenomenology of the state being introspected a *hyper-illusion*.

But are there any clear cases of hyper-illusions? I don't know of any. One candidate is the claim, often made, that although the "self" is really a discontinuous stream of experiences, we have the illusion that it is a continuous existent (Strawson 2003). But this alleged hyper-illusion is suspect, being perhaps more a matter of failing to experience the gappiness rather than actually experiencing non-gappiness. Further, subjects' introspective judgments led to the prediction investigated by Sperling, Landman et al., and Sligte et al. One should have an empirical reason to judge that this experimentally confirmed introspective judgment is wrong.

Subjects in the Landman et al. experiment are looking right at the rectangles for half a second, a long exposure, and it is not hard to see the orientations clearly. It does not appear to them as if something vaguely rectangularish is coming into view, as if from a distance. In (c) of Landman et al., they see all the rectangle orientations for up to 1.5 seconds in the Landman et al. version and up to 4 seconds in the Sligte et al. version. It is hard to believe that people are wrong about the appearances for such a long period.

Dehaene et al. (2006) revisit this issue. Here are the relevant paragraphs (references are theirs):

> The philosopher Ned Block, however, has suggested that the reportability criterion underestimates conscious contents (Block 2005). When we view a complex visual scene, we experience a richness of content that seems to go beyond what we can report. This intuition led Block to propose a distinct state of "phenomenal consciousness" prior to global access. This proposal receives an apparent confirmation in Sperling's iconic memory paradigm. When an array of letters is flashed, viewers claim to see the whole array, although they can later report only one subsequently cued row or column. One might conclude that the initial processing of the array, prior to attentional selection of a row or column is already phenomenally conscious. (Block 2005, Lamme 2003)

> However, those intuitions are questionable, because viewers are known to be over-confident and to suffer from an "illusion of seeing". (O'Regan & Noë 2001). The change blindness paradigm demonstrates this "discrepancy between what we see and what we think we see" (Simons & Ambinder 2005). In this paradigm, viewers who claim to perceive an entire visual scene fail to notice when an important element of the scene changes. This suggests that, at any given time, very little of the scene is actually consciously processed. Interestingly, changes that attract attention or occur at an attended location are immediately detected. Thus, the illusion of seeing may arise because viewers know that they can, at will, orient attention to any location and obtain conscious information from it. (Dehaene et al. 2006, p. 210)

Dehaene and his colleagues propose to use the change "blindness" results to back up their view of the Sperling result. But the issues in these two paradigms are pretty much the same – our view of one is conditioned by our view of the other. Further, as I mentioned earlier, the first form of the Landman et al. experiment (see Fig. 2, Part [a]) is itself an experiment in the same vein as the standard change "blindness" experiments. The subject sees 8 things clearly but has the capacity (in the sense of Cowan's K) to make comparisons for only 4 of them. And so the Landman et al. experiment – since it gives evidence that the subject really does see all or almost all the rectangles – argues against the interpretation of the change "blindness" experiments given by Dehaene and his colleagues.

Dehaene et al. (2006) say, "The change blindness paradigm demonstrates this discrepancy between what we see and what we think we see." But this claim is hotly contested in the experimental community, including by one of the authors that they cite. As I mentioned earlier (see Note 2), many psychologists would agree that initial interpretations of change "blindness" went overboard and that, rather than seeing the phenomenon as a form of inattentional blindness, one might see it as a form of inattentional inaccessibility (Block 2001). That is, the subject takes in the relevant detail of each of the presented items, but they are not conceptualized at a level that allows the subject to make a comparison. As Fred Dretske (2004) has noted, the difference between the two stimuli in a change blindness experiment can be one object that appears or disappears, and one can be aware of that object that constitutes the difference without noticing that there is a difference.

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

Compare Figure 1 with Figure 4. It can be hard for subjects to see the difference between Figure 1 and Figure 4, even when they are looking right at the feature that changes. The idea that one cannot see the feature that changes strains credulity.

Two of the originators of the change "blindness" experiments, Dan Simons and Ron Rensink (see Simons & Rensink 2005a) have since acknowledged that the "blindness" interpretations are not well supported by the "change blindness" experiments. In a discussion of a response by Alva Noë (2005), they summarize (Simons and Rensink 2005b):

> We and others found the "sparse representations" view appealing (and still do), and initially made the overly strong claim that change blindness supports the conclusion of sparse representations (Rensink et al. 1997; Simons 1997). We wrote our article because change blindness continues to be taken as evidence for sparse – or even absent – representations, and we used O'Regan and Noë's influential paper (O'Regan & Noe 2001) as an example. However, as has been noted for some time . . . this conclusion is logically flawed. (Simons & Rensink 2005b, p. 219)

I have been appealing to what the subjects say in Sperling-like experiments about seeing all or almost all the items. However, there is some experimental confirmation of what the subjects say in a different paradigm. Geoffrey Loftus and his colleagues (Loftus & Irwin 1998) used a task devised by Vincent Di Lollo (1980) and his colleagues using a 5 by 5 grid in which all but one square is filled with a dot. Loftus et al. divided the dots into two groups of 12, showing subjects first one group of 12 briefly, then a pause, then the other group of 12 briefly. The subjects always were given partial grids, never whole grids. Subjects were asked to report the location of the missing dot – something that is easy to do if you have a visual impression of the whole grid. In a separate test with no missing dots, subjects were asked to judge on a scale of 1 to 4 how temporally integrated the matrix appeared to be. A "4" meant that one complete matrix appeared to have been presented, whereas a "1" meant that two separate displays had been presented. The numerical ratings are judgments that reflect phenomenology: how complete the grids looked. The length of the first exposure and the time between exposures was varied. The Loftus et al. experiment probes persistence of phenomenology without using the partial report technique that leads Dehaene and his colleagues (2001; 2006) to suggest the Refrigerator Light illusion. The result is that subjects' ability to judge which dot was missing correlated nearly perfectly with their phenomenological judgments of whether there appeared to be a whole matrix as opposed to two separate partial matrices. That is, the subjects reported the experience of seeing a whole matrix if and only if they could pick out the missing dot, thus confirming the subjects' phenomenological reports.

To sum up: (1) the subjects' introspective judgments in the experiments mentioned are that they see all or almost all of the items. Dehaene and his colleagues (2001; 2006) seem to agree since that is entailed by the claim that the introspective judgments are illusory. (2) This introspective judgment is not contingent on subjects' being asked to report items as would be expected on the illusion hypothesis. (3) This introspective judgment leads to the prediction of partial report superiority, a prediction that is borne out. (4) The accuracy of the subjects' judgments is suggested by the fact that subjects are able to recover both size and orientation information with no loss. (5) These results cohere with a completely different paradigm – the Loftus paradigm just mentioned. (6) Dehaene and his colleagues offer no empirical support other than the corresponding theory of the change "blindness" results which raise exactly the same issues.

The conclusion of this line of argument is, as mentioned before, that phenomenology overflows cognitive accessibility and so phenomenology and cognitive access are based at least partly in different systems with different properties. I will be moving to the promised argument that appeals to the mesh between psychology and neuroscience after I fill in some of the missing premises in the argument, the first of which is the evidence for a capacity of visual working memory of roughly four or less.



Figure 4. Compare this with Figure 1 without looking at the two figures side by side. There is a difference that can be hard to see.

12. Visual working memory

At a neural level, we can distinguish between memory that is coded in the active firing of neurons – and ceases when that neuronal firing ceases – and structural memory that depends on changes in the neural hardware itself, for example, change in strength of synapses. The active memory – which is active in the sense that it has to be actively maintained – is sometimes described as "short term" – a misdescription since it lasts as long as active firing lasts, which need not be a short time if the subject is actively rehearsing. In this target article, the active memory buffer is called "working memory".

You may have heard of a famous paper by George Miller called "The magical number seven, plus or minus two: Some limits on our capacity for processing information" (Miller 1956). Although Miller was more circumspect, this paper has been widely cited as a manifesto for the view that there is a single active memory system in the brain that has a capacity of seven plus or minus two "items." What is an item? There are some experimental

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

results that fill-in this notion a bit. For example, Huntley-Fenner et al. (2002) showed that infants' visual object tracking system – which, there is some reason to believe, makes use of working memory representations – does not track piles of sand that are poured, but does track them if they are rigid. One constraint on what an item might be comes from some experiments that show that although we can remember only about four items, we can also remember up to four features of each one. Luck and Vogel asked subjects to detect changes in a task somewhat similar to the Landman et al. task already mentioned. They found that subjects could detect changes in four features (color, orientation, size, and the presence or absence of a gap in a figure) without being significantly less accurate than if they were asked to detect only one feature (Luck & Vogel 1997; Vogel et al. 2001).

In the 50 years since Miller's paper, reasons have emerged to question whether there really is a single active memory system as opposed to a small number of linked systems connected to separate modalities and perhaps separate modules – for example, language. Some brain injuries damage verbal working memory but not spatial working memory (Basso et al. 1982), and others have the opposite effect (Hanley et al. 1991). And evidence has accumulated that the capacity of these working memories – especially visual working memory – is actually lower than seven items (Cowan 2001, Cowan et al. 2006).

The suggestion of seven items was originally made plausible by experiments demonstrating that people, if read lists of digits, words or letters, can repeat back about seven of them. Of course, they can repeat more items if the items can be "chunked." Few Americans will have trouble holding the nine letters "FBICIAIRS" in mind, because the letters can be chunked into 3 acronyms.

More relevant to our discussion, visual working memory experiments also come up with capacities in the vicinity of four, or fewer than four, items. (For work that suggests fewer than four, see McElree 2006). Whether there is one working memory system that is used in all modalities or overlapping systems that differ to some extent between modalities, this result is what is relevant to the experiments discussed above. Indeed, you have seen three examples in this target article itself: the Sperling, Landman et al., and Sligte et al. experiments. I will briefly mention a few other quite different paradigms that have come up with the same number. One such paradigm involves the number of items that people – and monkeys – can effortlessly keep track of. For example, at a rhesus macaque monkey colony on a small island off of Puerto Rico, Marc Hauser and his colleagues (Hauser et al. 2000) did the following experiment: Two experimenters find a monkey relaxing on its own. Each experimenter has a small bucket and a pocket full of apple slices. The experimenters put down the buckets and, one at a time, they conspicuously place a small number of slices in each bucket. Then they withdraw and check which bucket the monkey goes to in order to get the apple slices. The result is that for numbers of slices equal to or smaller than four, the monkeys overwhelmingly choose the bucket with more slices. But if either bucket has more than four, the monkeys choose at random. In particular, monkeys chose the greater number in comparison of one versus two, two versus three, and three versus four, but they chose at random in cases of four versus five, four versus six, four versus eight, and, amazingly, three versus eight. The comparison of the three versus four case (where monkeys chose more) and the three versus eight case (where they chose at random) is especially telling. The eight apple slices simply overflowed working memory storage.

Infant humans show similar results, although typically with a limit more in the vicinity of three rather than four (Feigenson et al. 2002). Using graham crackers instead of apple slices, Feigenson et al. found that infants would crawl to the bucket with more crackers in the cases of one versus two and two versus three, but were at chance in the case of one versus four. Again, four crackers overflows working memory. In one interesting variant, infants are shown a closed container into which the experimenter – again conspicuously – inserts a small number of desirable objects (e.g., M&Ms). If the number of M&Ms is one, two, or three, the infant continues to reach into the container until all are removed, but if the number is more than three, infants reach into the container just once (Feigenson & Carey 2003).

I mentioned above that some studies have shown that people can recall about four items including a number of features of each one. However, other studies (Xu 2002) have suggested smaller working memory capacities for more complex items. Xu and Chun (2006) have perhaps resolved this controversy by showing that there are two different systems with somewhat different brain bases. One of these systems has a capacity of about four spatial locations, or objects at four different spatial locations, independent of complexity; the other has a smaller capacity depending on the complexity of the representation. The upshot for our purposes is that neither visual working memory system has a capacity higher than four.

This section is intended to back up the claim made earlier about the capacity of working memory – at least visual working memory. I move now to a quick rundown on working memory and phenomenology in the brain with an eye to giving more evidence that we are dealing with at least partially distinct systems with different properties.

13. Working memory and phenomenology in the brain

Correlationism in its metaphysical form (which, you may recall, regards cognitive accessibility as part of phenomenology) would have predicted that the machinery underlying cognitive access and underlying phenomenal character would be inextricably entwined in the brain. But the facts so far can be seen to point in the opposite direction, or so I argue.

In many of the experiments mentioned so far, a brief stimulus is presented, then there is a delay before a response is required. What happens in the brain during the delay period? In experiments on monkeys using this paradigm, it has been found that neurons in the upper sides of the prefrontal cortex (dorsolateral prefrontal cortex) fire during the delay period. And errors are correlated with decreased firing in this period (Fuster 1973; Goldman-Rakic 1987). Further, damage to neurons in

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

this area has been found to impair delayed performance, but not simultaneous performance, and damage to other memory systems does not interfere with delayed performance (except possibly damage to parahippocampal regions in the case of novel stimuli; Hasselmo & Stern 2006). Infant monkeys (1.5 months old) are as impaired as adult monkeys with this area ablated, and if the infant area is ablated, the infants do not develop working memory capacity. It appears that this prefrontal area does not itself store sensory signals, but rather, is the main factor in maintaining representations in sensory, sensorimotor, and spatial centers in the back of the head. As Curtis and D'Esposito (2003, p. 415) note, the evidence suggests that this frontal area "aids in the maintenance of information by directing attention to internal representations of sensory stimuli and motor plans that are stored in more posterior regions." That is, the frontal area is coupled to and maintains sensory representations in the back of the head that represent, for example, color, shape, and motion. (See Supèr et al. [2001a] for an exploration of the effect of this control on the posterior regions.) The main point is that, as the main control area for working memory, this prefrontal area is the main bottleneck in working memory, the limited capacity system that makes the capacity of working memory what it is.

So the first half of my brain-oriented point is that the control of working memory is in the front of the head. The second half is that, arguably, the core neural basis of visual phenomenology is in the back of the head. I will illustrate this point with the example of one kind of visual experience of motion (typified by optic flow). But first a caution: No doubt the neural details presented here are wrong, or at least highly incomplete. We are still in early days. My point is that the evidence does point in a certain direction, and more important, we can see how the issues I have been talking about could be resolved empirically.

Here is a brief summary of *some* of the vast array of evidence that the core neural basis of one kind of visual experience of motion is activation of a certain sort in a region in the back of the head centered on the area known as V5[7]. The evidence includes:

Activation of V5 occurs during motion perception (Heeger et al. 1999).

Microstimulation to monkey V5 while the monkey viewed moving dots influenced the monkey's motion judgments, depending on the directionality of the cortical column stimulated (Britten et al. 1992).

Bilateral (both sides of the brain) damage to a region that is likely to include V5 in humans causes akinetopsia, the inability to perceive – and to have visual experiences as of motion. (Akinetopsic subjects see motion as a series of stills.) (Rees et al. 2002a, Zihl et al. 1983).

The motion aftereffect – a moving afterimage – occurs when subjects adapt to a moving pattern and then look at a stationary pattern. (This occurs, for example, in the famous "waterfall illusion.") These moving afterimages also activate V5 (Huk et al. 2001).

Transcranial magnetic stimulation (TMS[8]) applied to V5 disrupts these moving afterimages (Theoret et al. 2002).

V5 is activated even when subjects view "implied motion" in still photographs, for example, of a discus thrower in mid-throw (Kourtzi & Kanwisher 2000).

TMS applied to visual cortex in the right circumstances causes stationary phosphenes[9] – brief flashes of light and color. (Kammer 1999) When TMS is applied to V5, it causes subjects to experience moving phosphenes (Cowey & Walsh 2000).

However, mere activation over a certain threshold in V5 is not enough for the experience as of motion: the activation probably has to be part of a recurrent feedback loop to lower areas (Kamitani & Tong 2005; Lamme 2003; Pollen 2003; Supèr et al. 2001a). Pascual-Leone and Walsh (2001) applied TMS to both V5 and V1 in human subjects, with the TMS coils placed so that the stationary phosphenes determined by the pulses to V1 and the moving phosphenes from pulses to V5 overlapped in visual space. When the pulse to V1 was applied roughly 50 msecs later than to V5, subjects said that their phosphenes were mostly stationary instead of moving. The delays are consonant with the time for feedback between V5 and V1, which suggests that experiencing moving phosphenes depends not only on activation of V5, but also on a recurrent feedback loop in which signals go back to V1 and then forward to V5. Silvanto and colleagues (Silvanto et al. 2005a; 2005b) showed subjects a brief presentation of an array of moving dots. The experimenters pinpointed the precise time – call it t – at which zapping V5 with TMS would disrupt the perception of movement. Then they determined that zapping V1 either 50 msecs before t or 50 msecs after t would also interfere with the perception of the moving dots. But zapping V5 a further 50 msecs after that (i.e., 100 msecs after t) had no effect. Silvanto et al. argue that in zapping V1 50 msecs before t, they are intercepting the visual signal on its way to V5, and in zapping V1 50 msecs after t, they are interfering with the recurrent loop. These results suggest that one V1-V5-V1 loop is the core neural basis for at least one kind of visual experience as of motion (and also necessary for that kind of experience in humans).

Recurrent loops also seem to be core neural bases for other types of contents of experience (Supèr et al. 2001a). The overall conclusion is that there are different core neural bases for different phenomenal characters. (Zeki and his colleagues have argued for a similar conclusion, using Zeki's notion of micro-consciousness [Pins & ffytche 2003; Zeki 2001]).[10]

14. Neural coalitions

But there is a problem in the reasoning of the last section. Whenever a subject reports such phenomenology, that can only be via the activation of the frontal neural basis of global access. And how do we know whether those frontal activations are required for – indeed, are part of – the neural basis of the phenomenology? Metaphysical correlationists say that they are; epistemic correlationists say we can't know. This section draws together strands that have been presented to argue that both kinds of correlationism are wrong because we have empirical reason to suppose that activation of working memory circuits are not part of the neural basis of phenomenology (not part of either the core or total neural basis).

A preliminary point: Pollen (in press) summarizes evidence that prefrontal lobotomies on both sides and other

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

frontal lesions do not appear to decrease basic perceptual content such as luminance or color. Frontal damage impairs access but it doesn't dim the bulb (Heath et al. 1949). Still, it could be said that *some* degree of frontal activation, even if minimal, is part of the background required for phenomenal consciousness, and, epistemic correlationists would allege, once there is so much frontal damage that the subject cannot report anything at all, there is no saying whether the person has any phenomenal consciousness at all.

In the rest of this section, I give my argument against this view, the one that the second half of the article has been leading up to: If we suppose that the neural basis of the phenomenology does *not* include the workspace activations, we can appreciate a neural mechanism by which phenomenology can overflow cognitive accessibility.

There is convincing evidence that the neural processes underlying perceptual experience can be thought of in terms of neural network models (see Koch 2004, Ch. 2, pp. 19, 20). In visual perception, coalitions of activation arise from sensory stimulation and compete for dominance in the back of the head, one factor being feedback from the front of the head that differentially advantages some coalitions in the back. Dominant coalitions in the back of the head trigger coalitions in the front of the head that themselves compete for dominance, the result being linked front and back winning coalitions. Support for this sort of model comes from, among other sources, computerized network models that have confirmed predictive consequences (see Dehaene & Nacchache 2001; Dehaene et al. 1998; 2006). Furthermore, some recent experiments (Sergent & Dehaene 2004) provide another line of evidence for this conclusion that is particularly relevant to the themes of this article.

This line of evidence depends on a phenomenon known as the "attentional blink." The subject is asked to focus on a fixation point on a screen and told that there will be a rapid sequence of stimuli. Most of the stimuli are "distractors," which in the case of the Sergent and Dehaene version are black nonsense letter strings. The subject is asked to look for two "targets": a white string of letters, either XOOX or OXXO, and a black name of a number, for example, "five." One or both targets may be present or absent in any given trial. At the end of the series of stimuli, the subjects have to indicate what targets they saw. In the *standard* attentional blink, subjects simply indicate the identity of the target or targets. The standard finding is that if the subject saw the first target (e.g., "XOOX"), and if the timing of the second target is right, the second target (e.g., "five") is unlikely to be reported at certain delays (so long as it is followed by distractors that overwrite the phenomenal persisting representation, as in Figure 5). In this setup, a delay of about 300 msecs makes for maximum likelihood for the second target to be "blinked." Sergent and Dehaene used a slight modification of this procedure in which subjects were asked to manipulate a joystick to indicate just how visible the number name was. One end of the continuum was labeled maximum visibility and the other was total invisibility. (See Fig. 5.)

The interesting result was that subjects tended to indicate that target 2 was either maximally visible or maximally invisible: intermediate values were rarely chosen. This fact

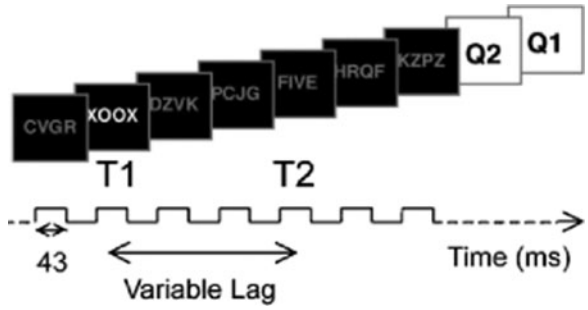


Figure 5. The Attentional Blink. A sequence of visual stimuli in which the first target is a white string of letters, either XOOX or OXXO and the second target is the name of a number. At the end of the series the subject is asked to indicate how visible target 2 was and whether target 1 was present, and if so, in which form.

suggests a competition among coalitions of neurons in the front of the head with winners and losers and little in between. Usually, the coalition representing the number word either wins, in which case the subject reports maximum visibility, or it loses, in which case subjects report no second target and there is no cognitive access to it at all.

I have guessed (Block 2005) that there can be coalitions in the back of the head which lose by a small amount and thus do not trigger winning coalitions in the front, but which are nonetheless almost as strong as the back of the head coalitions that do trigger global broadcasting in the front. The subject sees many things, but only some of those things are attended to the extent that they trigger global broadcasting. A recent study (Kouider et al. 2007) suggests that indeed there are strong representations in the back of the head that do not benefit from attention and so do not trigger frontal activations. (See also Tse et al. 2005 for convergent results.)

Kouider et al. contrasted a subliminal and supraliminal presentation of a stimulus, a lower-case word. In the subliminal case, the stimulus was preceded and succeeded by masks, which have the effect of decreasing the visibility of the stimulus (and, not incidentally, decreasing recurrent neural activation – see Supèr et al. 2001b). In the supraliminal case, the masks closest to the stimulus were omitted. The supraliminal but not the subliminal stimulus could be identified by the subjects when given a forced choice. In the key manipulation, the subject was told to look for an upper-case word, ignoring everything else. In those conditions of distraction, subjects claimed that they were aware of the lower-case stimuli in the supraliminal case but that they could hardly identify them because they were busy performing the distracting task on the upper-case stimulus (which came later). The difference between the supraliminal and subliminal stimuli in conditions of distraction was almost entirely in the back of the head (in occipito-temporal areas). Supraliminal stimuli activated visual areas in the back of the head strongly (subliminal stimuli did not) but did not activate frontal coalitions.[11] The strong activations in the back of the head did, however, *modulate* frontal activity.

Kouider et al. (2007) and Dehaene et al. (2006) acknowledge that there are highly activated losing

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

coalitions in the back of the head. They argue that such losing coalitions are the neural basis of "preconscious" states – because they cannot be reported. But the claim that they are not conscious *on the sole ground of unreportability* simply assumes metaphysical correlationism. A better way of proceeding would be to ask whether a phenomenal state might be present even when it loses out in the competition to trigger a winning frontal coalition.

Here is the argument that the second half of this article has been building up to: If we assume that the strong but still losing coalitions in the back of the head are the neural basis of phenomenal states (so long as they involve recurrent activity), then we have a neural mechanism which explains why phenomenology has a higher capacity than the global workspace. If, on the contrary, we assume that the neural basis of phenomenology includes workspace activation, then we do not have such a mechanism. That gives us reason to make the former assumption. If we make the former assumption – that workspace activation is not part of the neural basis of phenomenology – we have a mesh between the psychological result that phenomenology overflows cognitive accessibility and the neurological result that perceptual representations that do not benefit from attention can nonetheless be almost as strong (and probably recurrent) as perceptual representations that do benefit from attention. The psychological argument from overflow showed that the machinery of phenomenology is at least to some extent different from that of cognitive accessibility, since something not in cognitive accessibility has to account for the greater capacity of phenomenology. What the mesh argument adds is that the machinery of phenomenology does not include the machinery of cognitive accessibility.

Of course, my conclusion that the neural machinery of cognitive access is not partially constitutive of phenomenology leaves room for *causal* influence in both directions. And it may be that top-down causal influence is almost always involved in making the phenomenal activations strong enough. But that is compatible with the possibility of the relevant amplification happening another way, for example, by recurrent loops confined to the back of the head or even by stimulation by electrodes in the brain, and that is enough to show that top-down amplification is not *constitutively* necessary.

My first conclusion then is that the overlap of the neural machinery of cognitive access and the neural machinery of phenomenology can be empirically investigated. Second, there is evidence that the latter does not include the former. These points are sufficient to refute the correlationism of the sort advocated by Dehaene and his colleagues and to answer the question posed at the beginning of this article.

Further, this theoretical picture leads to predictions. One prediction is that in the Sperling, Landman et al., and Sligte et al. experiments the representations of the unaccessed items will prove to involve recurrent loops. Another upshot is that if the activations of the fusiform face area mentioned earlier in the patient G.K. turn out to be recurrent activations, we would have evidence for phenomenal experience that the subject not only does not know about, but in these circumstances cannot know about. The fact that the fusiform face activations produced in G.K. by the faces he says he doesn't see are almost as strong as the activations corresponding to faces he does see, suggests that top-down amplification is not necessary to achieve strong activations.

The mesh argument suggests that workspace activation is not a constitutive part of phenomenology. And given that actual workspace activation is not a constitutive part of phenomenology, it is hard to see how anyone could argue that potential workspace activation is a constitutive part. Further, as noted a few paragraphs back, it is doubtful that potential workspace activation is even causally necessary to phenomenology.

15. Conclusion

If we want to find out about the phenomenological status of representations inside a Fodorian module, we should find the neural basis of phenomenology in clear cases and apply it to neural realizers inside Fodorian modules. But that requires already having decided whether the machinery of access should be included in the neural kinds in clear cases, so it seems that the inquiry leads in a circle. This target article has been about breaking out of that circle. The abstract idea of the solution is that all the questions have to be answered simultaneously, tentative answers to some informing answers to others. The key empirical move in this article has been to give meshing answers to psychological and neural considerations about overflow.

ACKNOWLEDGMENTS

Earlier versions of this paper were presented at the following venues: the 2005 NYU Colloquium in Mind and Language; Georgia State University; MIT; Harvard Medical School; Dan Pollen's course at Harvard College; Dartmouth College; NYU Medical Center; Aarhus University, Denmark; Stockholm University, Sweden; Göteborg University, Sweden; Umeå University, Sweden; Indiana University, Brown University; the Pacific APA 2006; University of Western Ontario; Cal State, LA; UCLA; the Association for the Scientific Study of Consciousness; the Institut Jean Nicod; the University of Toronto; and at Katalin Balog's class at Yale. I am grateful to Thomas Nagel for his role as chief inquisitor when this material was defended at the NYU Mind and Language Colloquium and for some of his suggestions; and I am grateful to Katalin Balog, Alex Byrne, Tyler Burge, Susan Carey, Max Coltheart, Leonard Katz, Sean Kelly, Victor Lamme, Mohan Matthen, Alva Noë, Dan Pollen, David Rosenthal and Susanna Siegel for comments on a previous draft.

NOTES

1. See Siegel (2006a; 2006b) for discussion of what kind of thing the content of a phenomenal state is.

2. Phenomenal persistence and persistence of accessible information should be distinguished from what Koch (2004, Ch. 9) calls gist perception. We have specialized detectors for certain kinds of scenes, and in learning to read, we develop similar detectors for words. These detections can take place in 100 msecs and seem to require no attention. (See Potter 1993; Rousselet et al. 2002.)

3. The inattentional blindness view can be found in Rensink (1997), Simons (1997), Noë (2004), and O'Regan & Noë (2001). Views more closely related to the inattentional inaccessiblity view can be found in articles by philosophers – (e.g., Block 2001; Cohen 2002; Dretske 2004) – and by psychologists – (e.g., Simons & Rensink 2005a; 2005b; Wolfe 1999).

4. The "cerebral celebrity" view of consciousness is not the view in Dennett's *Consciousness Explained* (Dennett 1991),

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

but was introduced a few years after that, I think first in Dennett (1993). I argued for a distinct notion of "access-consciousness" in Block (1990; 1992).

5. I used this term in Block (2001), but I discovered years later that Nigel Thomas published pretty much the same idea first, deriving it from Marvin Minsky's "Immanence Illusion." Minsky's (1986, sect. 15.5) Immanence Illusion is this: "Whenever you can answer a question without a noticeable delay, it seems as though that answer were already active in your mind." At least in what I have read, Minsky does not focus on the idea that potential phenomenology is supposed to be confused with actual phenomenology. Thomas does focus on phenomenology, arguing for a view similar to that of O'Regan and Noë mentioned earlier:

> The seeming immediate presence of the visual world to consciousness does not arise because we have built a detailed internal representation of it, rather it is (like the ever shining fridge light) a product of the "immanence illusion" (Minsky 1986). For the most part, the visual perceptual instruments ask and answer their questions so quickly and effortlessly that it seems as though all the answers are already, and contemporaneously, in our minds. (Thomas 1999, p. 219)

6. Noë (2002; 2004) suggests an even more pervasive form of such an illusion – that all experience is a matter of potentiality, but precisely because it is so pervasive, he does not regard the view as one that postulates an illusion. (See Cohen 2002.)

7. The first classical "visual" cortical area is V1; later classic "visual" areas include V2, V3, V4, and V5. The latter, V5, has two names, "MT" and "V5" because it was identified and named by two groups. I put "visual" in scare quotes because there is some debate as to whether some of the classic "visual" areas are best thought of as multimodal and spatial rather than as visual per se. The motion area I am talking about in the text is actually a complex including MT/V5 and surrounding areas, and is often referred to as hMT+. (See Kriegeskorte et al. 2003).

8. TMS delivers an electromagnetic jolt to brain areas when placed appropriately on the scalp. The effect is to disrupt organized signals but also to create a signal in a quiescent area. Thus TMS can both disrupt moving afterimages and create phosphenes. A comparison is to hitting a radio: the static caused might interrupt good reception going on but also cause a noise when there is no reception. (I am indebted here to Nancy Kanwisher and Vincent Walsh, personal communication.)

9. To experiences phosphenes for yourself, close your eyes and exert pressure on your eye from the side with your finger. Or if you prefer not to put your eyeball at risk, look at the following website for an artist's rendition: http://www.reflectingskin.net/phosphenes.html

10. TMS stimulation directed to V1 may also stimulate V2 (Pollen 2003). Perhaps V2 or other lower visual areas can substitute for V1 as the lower site in a recurrent loop. Blindsight patients who have had blindsight for many years can acquire some kinds of vision in their blind fields despite lacking V1 for those areas. One subject describes his experience as like a black thing moving on a black background (Zeki & ffytche 1998). Afterimages in the blind field have been reported (Weiskrantz et al. 2002). Stoerig (2001) notes that blindsight patients are subject to visual hallucinations in their blind fields even immediately after the surgery removing parts of V1; however, this may be due to a high level of excitation that spreads to other higher cortical areas that have their own feedback loops to other areas of V1 or to other areas of early vision such as V2. (See also Pollen 1999.)

11. In a different paradigm (de Fockert et al. 2001), working memory load can increase the processing of distractors.

Open Peer Commentary

Access to phenomenality: A necessary condition of phenomenality?

doi: 10.1017/S0140525X07002798

Katalin Balog

Department of Philosophy, Yale University, New Haven, CT 06520-8306.
katalin.balog@yale.edu **http://pantheon.yale.edu/~kb237**

Abstract: Block argues that relevant data in psychology and neuroscience show that access consciousness is not constitutively necessary for phenomenality. However, a phenomenal state can be access conscious in two radically different ways. Its *content* can be access conscious, or its *phenomenality* can be access conscious. I argue that while Block's thesis is right when it is formulated in terms of the first notion of access consciousness, there is an alternative hypothesis about the relationship between phenomenality and access in terms of the second notion that is not touched by Block's argument.

Ned Block (in Block 1990; 1995b; 2002; and the present target article) has made a *conceptual* distinction between the *phenomenality* of a mental state (a.k.a. its *phenomenal character* or the *quale* it instantiates) and *access consciousness* of the same state. There is a view – Block calls it "epistemic correlationism" – according to which the *metaphysical* relationship between these two is not scientifically tractable. While cognitive accessibility is intrinsic to our *knowledge* of phenomenology, it might not be *constitutive* of the phenomenal facts themselves. According to the epistemic correlationist, there is no possible empirical evidence that could tell us one way or another. This view is Block's main target. His thesis is that the issue of the relationship between phenomenal and access consciousness is an empirical one; and that moreover, the issue is approachable by the same empirical methods we employ in science in general. Block's aim is to show that by looking at the relevant data, and employing the method of inference to the best explanation, we can mount an argument for the specific thesis that access consciousness is not constitutively necessary for phenomenality. If this is so, he has given reasons to reject "metaphysical correlationism" as well – that is, the view that the cognitive access relations that underlie reportability are *constitutive* of phenomenology.

Block's thesis needs further clarification. A phenomenal state can be access conscious in two radically different ways. Its *content* (or part of its content) can be access conscious, or its *phenomenal character* can be access conscious. Say, I am having a visual experience of a red circle in an orange background. In this case, both the content and the phenomenal character of this experience can be access conscious. I can be aware that I am seeing a red circle in an orange background, and I can also be aware that my experience has such and such a phenomenal character. However, in the experiments that Block discusses in his article, the two kinds of access come apart. As a result, although Block's thesis holds if understood as involving the first notion of access, there is an alternative thesis involving the second sense of access that is untouched by Block's arguments. After clarifying Block's thesis, I will briefly sketch this alternative hypothesis.

Consider the following kind of experiment, which provides crucial support for Block's thesis. Following Sperling's (1960) famous experiments, Landman et al. (2003) showed subjects eight rectangles in different orientations for half a second. The resulting experience e has a – presumably non-conceptual (pictorial or iconic) – representational content r, and, according to the introspective reports of subjects, a phenomenal character p.[1] Given the model of access consciousness assumed in Block's paper as broadcasting of *conceptual* representations

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

in a *global workspace* (Baars 1988; 1997), Block takes *e* to be access conscious if and only if conceptual representations of *e*'s content are present in the global workspace. In other words, *e* is access conscious if and only if there are conceptual representations in the global workspace that extract the content of *e* (e.g., "There were rectangles of the following orientations. . .").

That typically we are not access conscious, in the sense described above, of *all* aspects of a phenomenal experience's content is convincingly shown by the Landman et al. (2003) experiments. After the brief exposure, subjects are only able to report on the precise orientation of up to four of these rectangles. These experiments show, to my mind conclusively, that access consciousness of this sort – that is, the existence of conceptual representations in global workspace that extract all the relevant content of *e* – is not constitutively necessary for the phenomenality of the experience. This finding is further supported by the neurophysiological data Block cites, which show the neural implementation of sensory representations and the neural implementation of global access to be physically separate and independent from each other.

However, these experiments – which comprise the bulk of Block's supporting evidence – do not show that *no* access is constitutively necessary for phenomenality. Notice that the aforementioned interpretation of these experiments crucially relies on the subjects' introspective report of the phenomenality of their *entire visual experience*, including those aspects of the experience whose content is not access conscious. Introspective access to the phenomenality of the entire experience was part of the evidence in the Sperling and the Landman et al. experiments for why access to the *conceptualized content* of the experience is not necessary for phenomenality. But these data leave room open for the hypothesis that access to the *phenomenality* of the experience *is* constitutively necessary for that phenomenality. How exactly should we think about access to the phenomenality of the experience if it is not access to its conceptualized content?

Notice that the representations in the global workspace that are not constitutively necessary for phenomenality are *separate from* the representations whose phenomenality is in question. Phenomenal experience quite plausibly involves non-conceptual representation; representations that enter the global workspace, on the other hand, are conceptual representations. There are different representations involved. What about access to the *phenomenality* of the experience itself? It seems plausible that the relationship between phenomenality and the representation of it that is in the global workspace is more intimate. Here is an idea: Perhaps phenomenality requires that a *conceptual* representation of the *phenomenal character* of the experience, more precisely, a judgment to the effect that the relevant phenomenal experience occurs, itself is in the global workspace. Plausibly, this would not involve any old conceptual representation of the phenomenality of the experience, but a *phenomenal* representation involving phenomenal concepts. There is a plausible account of phenomenal concepts, the *constitutional account* (see, e.g., Papineau 2002), according to which phenomenal concepts – in their canonical, first person, present tense applications relevant to these experiments – are partly constituted by the *experience* they refer to. That is, the first-person, present-tense judgment that *e* has phenomenal character *p* is partly constituted by *e* itself. Notice that here the experience whose phenomenality is at issue and the state in the global workspace that constitutes access to it are not separate and independent. The conceptual representation in the global workspace involves *e* itself and this adds to the plausibility of the idea that this kind of access is intrinsic to phenomenality.

Unlike the thesis Block is criticizing (let's call it the Access_c thesis), this thesis (let's call it the Access_p thesis) seems to be a viable hypothesis. None of the data discussed by Block rule it out, or even make it implausible. But if the Access_p thesis is true, then some interesting consequences follow – for example, that despite suggestions to the contrary by Block, activations in the "fusiform face area" of "visuo-spatial extinction" patients, or any other early visual state that is not access_p conscious, could not be phenomenal.

ACKNOWLEDGMENT

I would like to thank Georges Rey for useful suggestions.

NOTE

1. I want to sidestep the issue of representationalism about qualia here. All I assume is that *e* has *r* and *p*; I won't discuss the relationship between *r* and *p*.

Psychology supports independence of phenomenal consciousness

doi: 10.1017/S0140525X07002804

Tyler Burge

Department of Philosophy, University of California, Los Angeles, Los Angeles, CA 90024.

burge@ucla.edu

Abstract: Inference-to-best-explanation from psychological evidence supports the view that phenomenal consciousness in perceptual exposures occurs before limited aspects of that consciousness are retained in working memory. Independently of specific neurological theory, psychological considerations indicate that machinery producing phenomenal consciousness is independent of machinery producing working memory, hence independent of access to higher cognitive capacities.

Ned Block argues, "the machinery of cognitive access is not included in the machinery of phenomenology" (target article, sect. 9, para. 13). His argument is plausible, but I think psychological considerations support his conclusion, independently of neurological conjecture.

The view that all consciousness must be available to higher cognitive faculties is motivated by worry that without "reportability," consciousness cannot be studied scientifically. Either the view tries to rule apriori on empirical matters – how could it be apriori that animals that lack propositional attitudes cannot feel pain? – or it envisions too narrow a range of possible empirical evidence. Block has widened the range. Here, I think he underplays psychological considerations.

In the Sperling (1960), Landman et al. (2003), and Sligte et al. (2008) experiments, subjects, using iconic memory, take themselves to have seen a relatively specific number of items, experimentally investigated to have been in the 8 to 32 range. The items are displayed long enough for normal perceptual processing to be completed. In any given trial, using working memory, subjects can make use of information on only four instances of *specific types* of items, say, specific alphanumeric characters. They can do this for specific types of *any* 4 of 8 to 32, if cued *after* presentation.

I believe these experiments support two conclusions: (a) In any given trial, there are phenomenally conscious perceptions of specific types of items not accessed by working memory; and (b) causal machinery produces specific types of phenomenal consciousness on given occasions, but on those occasions does not register those types in working memory.

Block accepts, but does not highlight, conclusions (a) and (b). He argues for a stronger conclusion: (c) The machinery of working memory *does not overlap* the machinery of phenomenal consciousness. He reaches this third conclusion in three steps. First, he holds that if one accepts the Sperling-type experiments at face value, the minimum concessions required of someone who believes working memory (and through it, "reportability") is constitutive to phenomenal consciousness are: (d) "the 'capacity' of . . . the visual

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

phenomenal memory system, is greater than that of the working memory buffer that governs reporting" (sect. 9, para. 11), and (e) "the machinery of phenomenology is at least somewhat different from the machinery of cognitive accessibility" (sect. 9, para. 13). (That is, (a)-(b) entails (d)-(e), but not vice-versa.) Second, he argues for accepting Sperling-type experiments at face value. Third, he invokes neurological conjecture to support (c).

One might accept (a)-(b) and (d)-(e), but insist that working memory and "reportability" are constitutive to phenomenal consciousness. One might hold that although *some* specific phenomenally conscious items do not appear in working memory, *all* phenomenal consciousness depends constitutively on some items' being accessible to working memory. Block marshals neurological considerations for (c) against such a position.

In his second stage, Block opposes Dehaene's attempts to avoid taking Sperling-type experiments at face value. Block effectively criticizes postulating what he calls a refrigerator-light illusion, and points out that it is question-begging to invoke "change-blindness" to support the position that the subjects in Sperling-type experiments are under an illusion that they had phenomenal experiences of items that do not appear in working memory. The two cases are disanalogous in a way that Block does not note. On Dehaene's view that Sperling-type subjects are phenomenally conscious only of items actually in working memory, the subjects cannot have had, before the cue that selects those items retained in working memory, a phenomenologically conscious perception of *any* of the specific 8–32 items that they seem to have experienced. On that view, subjects' sense of having consciously perceived even specific retained items before they appear in working memory is illusion. No one postulates analogous total illusion in "change-blindness" cases. Even proponents of the (I think mistaken) view that items that change unnoticed are not consciously seen do not claim that *nothing* is consciously seen.

I believe that Sperling-type experiments support (c), not just (a), (b), (d), (e). I argue by dilemma. If retained and unretained items are held *not* to be conscious before any items are retained in working memory, what is the evidence that memory of their having been conscious is total illusion? Exposure is long enough for perceptual processing to be complete. Why should phenomenology, even of specific retained items, have been missing? We have independent evidence about working memory. It does have constructive functions: making consciousness more vivid, rehearsing to facilitate retention and reproduction of imagery (Andrade 2001; Pearson 2001). But its primary function is to preserve perception already formed. Holding that its preservations convey systematic illusion is ad hoc. The fact that subjects seem to remember having seen all items, and could be cued to retain any item specifically, supports believing that even specific unretained items are phenomenally conscious. Now suppose that all, or at least the retained, specific items are held to be conscious before being preserved in working memory. What is the evidence that mere accessibility to working memory is constitutive to their being occurrently conscious beforehand? Such a view labors under heavy empirical burden. Consciousness is an occurrent, not a dispositional, condition. We have no good idea how mere dispositional accessibility to working memory could be causally necessary to *occurrence* of consciousness *before* working memory operates. Why should the door's being open matter to the occurrence of something that does not use the door until after it already occurs? Such a view would require very special evidence and explanation. In the absence of specific empirical support, the idea is not a serious contender. The best explanation of current evidence is that conscious perception of the specifics of items later retained, indeed of all 8–32 items, occurs independently of working memory. The machinery of phenomenal consciousness appears to be independent of the machinery of working memory. Conclusion (c) is supported independently of Block's neurological conjecture.

Further evidence for (c) may lie in the formation speed of at least *generic* phenomenally conscious aspects of visual perception. Some super-ordinate object categorization occurs in less than 150 msec – before a signal even reaches working memory (VanRullen & Thorpe 2001; Rousselet et al. 2004a; 2004b). Such considerations are tentative. But it is important not to be *so* fixed on neurological matters that one underrates the force of psychological considerations in supporting psychological conclusions.

Do we see more than we can access?

doi: 10.1017/S0140525X07002816

Alex Byrne[a], David R. Hilbert[b], and Susanna Siegel[c]

[a]*Department of Linguistics and Philosophy, Massachusetts Institute of Technology, Cambridge, MA 02319;* [b]*Department of Philosophy, University of Illinois at Chicago, MC 267, Chicago, IL 60607;* [c]*Department of Philosophy, Harvard University, Emerson Hall, Cambridge, MA 02138.*

abyrne@mit.edu **http://web.mit.edu/abyrne/www**
hilbert@uic.edu **www.uic.edu/~hilbert**
ssiegel@fas.harvard.edu **http://www.people.fas.harvard.edu/~ssiegel**

Abstract: One of Block's conclusions, motivated by partial-report superiority experiments, is that there is phenomenally conscious information that is not cognitively accessible. We argue that this conclusion is not supported by the data.

Block's overall argument appeals to the lemma that "in a certain sense phenomenal consciousness overflows cognitive accessibility" (target article, Abstract), which Block takes to be supported by Landman et al. (2003) and Sligte et al. (2008). (For reasons of space we will ignore the latter.) Block summarizes his discussion of these two papers as follows:

> The main upshot of the Landman et al. and the Sligte et al. experiments (at least on the surface – debunking explanations will be considered later) is along the same lines as that of the Sperling experiment: The subject has persisting experiences as of more specific shapes than can be brought under the concepts required to report or compare those specific shapes with others. (sect. 9, para. 10)

Thus, in the first condition of the Landman et al. experiment, Block holds that the subjects have persisting experiences as of [a circle of] eight rectangles, with the horizontal/vertical orientation of each rectangle specified. And if that is so, then, as Block says, the subject's experiences are not completely accessible, because the subjects can report the orientation of only four (or so) rectangles.

Although most of Block's discussion is couched in terms of "phenomenal consciousness" and the like, for present purposes we can talk instead (as Block himself sometimes does) of what the subjects see. Put this way, Block's claim is that the subjects continue to see each rectangle as oriented horizontally or vertically after the stimulus has been replaced with a gray screen. In the terminology of Coltheart (1980), this is an example of *visible persistence*.

Coltheart distinguishes *visible* persistence from *informational* persistence. The latter is defined not in terms of seeing, or phenomenal consciousness, but in terms of the persistence of rich visual information about a stimulus after it has been replaced. Sperling-type experiments show that stimulus information is held in a high-capacity but transient memory, and thus that there is informational persistence. One might hold that there is informational persistence simply because there is visible persistence; that is, stimulus information continues to be available because the subject continues to see the stimulus. Coltheart argues, however, that the phenomena are not connected so intimately. One consideration is that informational persistence lasts longer than a few hundred ms, the duration of visible persistence. (As Block notes, the duration of informational persistence found by Landman et al. is about 1,500 msecs.)[1]

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

With this distinction in hand, consider Landman et al.'s conclusion:

> The present data agree with the presence of two parallel types of short term memory. . . . Almost all items enter the first type of memory. It is like iconic memory, because it has a high capacity and it is maskable. . . . The second type of memory is one that resists interference by new stimuli. When new items enter the visual system, they replace the old items, except the ones that have entered the second type of representation. . . . The cue-advantage arises because the subjects selectively transfer the cued item from iconic memory to the more durable working memory. . . . (Landman et al. 2003, p.162)

Landman et al. are, then, concerned with informational persistence, not visible persistence. Their paper contains no data concerning visible persistence. Since informational persistence is consistent with no visible persistence at all, Block's appeal to Landman et al. must be somewhat indirect.

And indeed it is. Block's argument for visible persistence is based on subjects' reports: "[subjects say they are] continuing to maintain a visual representation of the whole array" (sect. 9, para. 6).

We have three points about this. First, Block needs only the weaker claim that the subjects in the Landman et al. experiment saw each rectangle as oriented horizontally or vertically, not the stronger claim that the subjects *remain* in this state after the stimulus has been replaced. The weaker claim implies Block's conclusion about inaccessibility for the same reason that the stronger one does.

We do not dispute that information about the orientation of each rectangle persists and is not as a whole accessible; we do dispute Block's claim that this inaccessible information characterizes what the subjects see. Our second point is that it is unclear that subjects' reports unequivocally support Block. Block needs subjects to agree that they saw *each* rectangle as oriented horizontally or vertically (even if they can't report which orientation each rectangle has). More precisely: for each rectangle x, either they saw x as horizontal, or they saw x as vertical. If the subjects merely say that they saw eight rectangles, some horizontal and some vertical, or that "they can see all or almost all the 8 to 12 items in the presented arrays" (sect. 9, para. 11), this is insufficient.

According to Landman et al., selected stimulus information is transferred from the transient iconic memory to the more durable working memory. Working memory therefore contains less information about the stimulus than iconic memory. If that is all that working memory contains, and if working memory governs subjects' reports about what they see (as Block supposes), then subjects would simply say that they saw a circle of rectangles and saw some of them as oriented horizontally/vertically. They would not, then, agree that they saw details, some of which they can't report. So our third point is this: Block must deny that the contents of working memory are simply a subset of the contents of iconic memory, which is to go beyond the results of Landman et al. If Block is right and subjects report (correctly) that they saw each rectangle as oriented horizontally or vertically, then the contents of working memory should include, not just certain information about the stimulus transferred from iconic memory, but also the meta-information that some information was *not* transferred. We are not saying that this proposal about the contents of working memory is wrong, but only that the Landman experiment does not address it.

NOTE

1. The question of the exact relationship between visible and informational persistence remains open. Loftus and Irwin (1998) argue that the many measures of visible and informational persistence pick out the same underlying process. Nevertheless, the distinction is still useful and our discussion does not rely on the assumption that it marks a real difference.

Experience and agency: Slipping the mesh

doi: 10.1017/S0140525X07002828

Andy Clark and Julian Kiverstein

School of Philosophy, Psychology and Language Sciences, George Square, Edinburgh EH8 9JX, Scotland, United Kingdom.

andy.clark@ed.ac.uk **s9903600@sms.ed.ac.uk**

http://www.philosophy.ed.ac.uk/staff/clark.html

Abstract: Can we really make sense of the idea (implied by Block's treatment) that there can be isolated islets of experience that are not even potentially available as fodder for a creature's conscious choices and decisions? The links between experience and the availability of information to guide conscious choice and inform reasoned action may be deeper than the considerations concerning (mere) reportability suggest.

In this elegant and tightly argued target article, Ned Block seeks to persuade us that phenomenal consciousness routinely "overflows" cognitive accessibility. By this he means that we can (and do) have experiences even in cases where we lack the kind of access that would yield some form of report that such and such an experience had occurred. The case Block makes for such an apparently hard-to-support judgment rests on a "mesh" between psychological results and work in neuroscience. The psychological data seem to show that subjects can see much more than working memory enables them to report. Thus, in the Landman et al. (2003) experiments, for instance, subjects show a capacity to identify the orientation of only four rectangles from a group of eight. Yet they typically report having seen the specific orientation of all eight rectangles. Working memory here seems to set a limit on the number of items available for conceptualization and hence report.

Work in neuroscience then suggests that unattended representations, forming parts of strong-but-still-losing clusters of activation in the back of the head, can be almost as strong as the clusters that win, are attended, and hence get to trigger the kinds of frontal activity involved in general broadcasting (broadcasting to the "global workspace"). But whereas Dehaene et al. (2006) treat the contents of such close-seconds as preconscious, because even in principle (given their de facto isolation from winning frontal coalitions) they are unreportable, Block urges us to treat them as phenomenally conscious, arguing that "the claim that they are not conscious *on the sole ground of unreportability* simply assumes metaphysical correlationism" (sect. 14, para. 9; italics in original). That is to say, it simply assumes what Block seeks to question – that is, that the kind of functional poise that grounds actual or potential report is part of what constitutes phenomenology. Contrary to this way of thinking, Block argues that by treating the just-losing coalitions as supporting phenomenally conscious (but in principle unreportable) experiences, we explain the psychological results in a way that meshes with the neuroscience.

The argument from mesh (which is a form of inference to the best explanation) thus takes as its starting point the assertion that the *only grounds* we have for treating the just-losing back-of-the-head coalitions as non-conscious is the unreportability of the putative experiences. But this strikes us as false, or at least premature. For underlying the appeal to reportability is, we suspect, a deeper and perhaps more compelling access-oriented concern. It is the concern that any putative conscious experience should be the experience *of an agent*. The thought here is that we cannot make sense of the image of free-floating experiences, of little isolated islets of experience that are not even potentially available as fodder for a creatures rational choices and considered actions. Evans (1982) rather famously rejects the very idea of such informationally isolated islands of experience. According to Evans, an informational state may underpin a conscious experience only if it (the informational state) is in some sense input to a reasoning subject. To count as a conscious experience an informational state must:

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

> [serve] as the input to a thinking, concept-applying and reasoning system: so that the subject's thoughts, plans, and deliberations are also systematically dependent on the informational properties of the input. When there is such a link we can say that the person, rather than some part of his or her brain, receives and processes the information. (Evans 1982, p.158)

The real point here is (or should be) independent of Evan's appeal to conceptualization. What matters, rather, is that the information must be available to the agent qua "reasoning subject," where this may be unpacked in many different ways, not all of them requiring full-blown concept-use on the part of the agent (see, e.g., Bermúdez & Macpherson 1998; Hurley 1997). Evans' insight is that the notions of conscious experience and reasoned agency (here very broadly construed) are deeply intertwined: that there are non-negotiable links between what is given in conscious awareness and the enabled sweep of deliberate actions and choices available to a reasoning subject. Such a story opens up a different way of interpreting the Sperling (1960) and the Landman et al. (2003) results. In these cases (we suggest) subjects report phenomenally registering all the items because information concerning each item was, at that moment, available to be deployed in the service of deliberate, reasoned, goal-directed action. Such momentary potentiality is not undermined by the (interesting and important) fact that the selection of a few items to actually play that role then precludes the selection of the rest.

Contrariwise, Block argues that a subject like G.K. can be having an experience of a face and yet it be impossible for him to know anything of this experience. Block takes G.K.'s phenomenal experience to be constituted by recurrent processing in the fusiform face area. We believe that G.K. can be consciously experiencing a face only if this experience is at least momentarily poised for use in reasoning, planning, and the deliberate selection of types of action. Recurrent processing in the fusiform area will no doubt prove to be among the conditions necessary for realizing a state that plays this causal role.

The contents of conscious phenomenal experience, if all this is on track, must be at least potentially available for use in the planning and selection of deliberate, stored-knowledge–exploiting, and goal-reflecting and goal-responsive, actions. Block's just-losing coalitions fail to trigger winning frontal coalitions and hence fail to be in a position to contribute their contents in this manner to the full sweep of the agent's deliberate acts and choices. It is this fact (rather than the more superficial indicator of unreportability) that should motivate our treating the contents of the just-losing coalitions as non-conscious. If this is correct, then the staging post for the argument from mesh is called into question. Until the considerations concerning links between experience and rational agency are more fully addressed, it remains unclear whether the kind of "fit" to which Block appeals can really favor his conclusion over our own.

ACKNOWLEDGMENTS

Thanks to Matthew Nudds, Tillmann Vierkant, and the participants in the Edinburgh Philosophy, Psychology, and Informatics Reading Group for stimulating discussion of the target paper. This commentary was prepared thanks to support from the AHRC, under the ESF Eurocores CNCC scheme, for the CONTACT (Consciousness in Interaction) project, AH/E511139/1.

Why babies are more conscious than we are

doi: 10.1017/S0140525X0700283X

Alison Gopnik

Department of Psychology, University of California at Berkeley, Berkeley, CA 94720.

gopnik@berkeley.edu

ihd.berkeley.edu/gopnik.htm

Abstract: Block argues for a method and a substantive thesis – that consciousness overflows accessibility. The method can help answer the question of what it is like to be a baby. Substantively, infant consciousness may be accessible in some ways but not others. But development itself can also add important methodological tools and substantive insights to the study of consciousness.

Infants and young children cannot report their phenomenology. This has led some philosophers to argue that babies' consciousness must be limited. Even if we think that babies are conscious it might seem impossible to recover the particular character of their experience. A version of Block's abductive method can at least partially solve this problem. We can consider a wide range of functional and neural correlates of conscious experience in adults and then look to see similarities and differences in babies.

In adults attention is highly correlated with vivid consciousness. This consciousness has a particular subjective quality – "the spotlight" – with a defined – "brightly lit" – focus and surrounding darkness. Both phenomenologically and functionally, attention to one event seems to inhibit consciousness of other events – as in inattentional blindness.

There is a distinction between exogenous and endogenous attention. Exogenous attention is driven by information-rich external events. These events may be intrinsically salient. But exogenous attention may also be driven by subtle unexpectedness. Exogenous attention is marked by characteristic event-related potential (ERP) signatures, eye movements, decelerating heart rate, and parietal activation. Exogenous attention and vivid consciousness characteristically fade as information is obtained, a process of habituation.

Endogenous attention is the sort of top-down "paying attention" that is motivated by specific goals rather than by intrinsic interest. It has been the focus of the adult literature for methodological reasons, and endogenous and exogenous attention are often not distinguished in discussions of consciousness. Frontal activation seems to be particularly important for endogenous attention.

These functional features of attention correlate well with neurological patterns. In adult animals, endogenous attention leads to the release of cholinergic transmitters to some parts of the brain and inhibitory transmitters to other parts. Attention increases both the efficiency of a particular part of the brain and its plasticity; and it inhibits activation and plasticity in other brain areas.

So for adults there is an elegant if undoubtedly oversimplified story about how consciousness works. In the canonical case, the goal-directed control systems in frontal cortex indicate that a particular kind of event is important. The perceptual system, guided by attention, zooms in on just that event and the brain extracts information about the event and modifies itself, that is, learns accordingly. Significantly, though, this whole process is highly focused; other parts of the brain may actually be shut down in the process. Vivid spotlight consciousness is the phenomenological result.

Even very young infants have extensive exogenous attention capacities. When they are presented with even highly subtle and conceptually unexpected novel events, they immediately focus their gaze on these events, and show similar heart rate deceleration and ERP signatures to those of adults. Indeed, this is the basis for the habituation technique that is our principal source of information about infant's minds. However, infants develop endogenous attention much later, and it is still developing during the preschool years. Moreover, and probably correlated with this fact, infants and young children appear to have less focused attention than older children – for example, they show better incidental memory.

The neurology suggests a similar picture. The parietal and sensory systems involved in exogenous attention are on line at an early age. The top-down frontal regions and connections that control endogenous attention only mature later. Young animals' brains are far more plastic than adult brains and this plasticity is much less focused and attention-dependent. Cholinergic transmitters are in place early, while inhibitory transmitters emerge only later.

So again, an undoubtedly oversimplified but suggestive picture emerges. When infants and young children process information

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

there is much less top-down control and inhibition. Infants are sensitive to any information-rich stimuli, not just those stimuli that are relevant to their goals and plans. And because they have much less experience, more stimuli will be information-rich for them than for us. With less top-down inhibition their processing and plasticity will be more distributed and less focused than those of adults.

These pictures also make sense from an evolutionary point of view. Childhood itself, that paradoxical period of helpless immaturity, is more protracted in species that rely heavily on learning – and most protracted in humans. Childhood reflects an evolutionary division of labor. Children are very good at learning about everything and are not very good at doing anything – adults use what they have learned in childhood to design effective actions.

As Block suggests, we can use this mesh of functional, neurological, and evolutionary facts to make a hypothesis about phenomenology. Babies are more conscious than we are. They are ceaselessly and broadly engaged in the kind of information-processing and learning that adults direct only at limited, relevant events. And babies are less subject to the processes that actively cause unconsciousness in adults – inhibition and habituation. When adults are placed in a situation that is functionally similar to babies, such as traveling in a strange country or meditating in certain ways, we experience a similarly vivid but distributed phenomenology – consciousness becomes a lantern instead of a spotlight.

For babies, consciousness may also overflow accessibility. But babies make the ambiguity of "accessibility" vivid. Obviously, information in infant consciousness is not accessible for reporting – it may also not be accessible for goal-directed planning nor be the subject of the endogenous attention that accompanies such planning. But it may indeed be accessible for purposes of induction, prediction, generalization, and intuitive theory-formation – the principal psychological activities of babies. Does that constitute overflow or not?

This leads to a broader methodological point, and to the greatest advantage of developmental data. Human adults are our chief source of information about consciousness. But in adults, many very different functional and neurological processes are highly correlated. Attention is focused, inhibitory, and top-down, *and* it leads to plasticity and learning, *and* it is reportable, *and* it is in the service of goals, *and* it is connected to a sense of self, *and* so on, *and* so forth. In development, this apparently unified picture breaks apart in unexpected and informative ways: Babies may have access for induction but not planning, have exogenous but not endogenous attention, and have plasticity and facilitation without inhibition. Rather than a single story of capital-c Consciousness, there may be many very varied relations between phenomenology and function.

A plug for generic phenomenology

doi: 10.1017/S0140525X07002841

Rick Grush

Department of Philosophy, University of California San Diego, La Jolla, CA 92093.

rick@mind.ucsd.edu **http://mind.ucsd.edu**

Abstract: I briefly sketch a notion of *generic phenomenology*, and what I call the *wave-collapse illusion* – a less radical cousin of the refrigerator light illusion – to the effect that transitions from generic to detailed phenomenology are not noticed as phenomenal changes. Change blindness and inattentional blindness can be analyzed as cases where certain things are phenomenally present, but generically so.

A number of researchers have used mismatches between what subjects think they are aware of and what subjects can report on, to argue that we are subject to a grand illusion – that the idea that we are presented with a rich, detailed visual scene is false. Block shows that with some of these phenomena (e.g., the phenomena in the Sperling [1960] and the Landman et al. [2003] experiments), the better analysis is that the items subjects think are phenomenally present are in fact present, but inaccessible in a certain sort of way. But there are other kinds of cases – change blindness and inattentional blindness – to which Block's analysis doesn't seem to apply, and so the threat of this illusion remains. I want to assist Block in dulling the teeth of the illusion argument.

The first topic I will discuss involves a sketch of the notion of *generic phenomenology*, the idea that some of what is phenomenally present is present only generically. Roughly the idea is that, for example, when reading a page of text, a word on the page in the visual periphery is phenomenally present, but not as the specific word it is, just as the generic *text word*. If I am right, then much mental representation, including phenomenology, is sui generis, and not well described as either picture-like or sentence-like; and in particular, generic phenomenology is neither a description, nor a blurry image. Nor is it an *interpreted image* (a combo-metaphor of both picture and sentence). The second part suggests how the problematic phenomena, change blindness and inattentional blindness, can be given plausible, and less alarmist, interpretations using the notion of generic phenomenology.

While there is no space for a full positive characterization of this sui generis format, I sketch three of its features here: First, it is possible that incorrect genera are employed in some cases. Second, a generic phenomenal presentation can present its content as *clear* without presenting the details. In this respect it differs from an image. This is, I currently believe, an affordance issue: a clear phenomenal presentation presents its content as affording answers to queries of detail. Images only offer this affordance if they actually have the detail. Third and relatedly, this representational format supports and is subject to a less radical version of the refrigerator light illusion, what I will call (with a great deal of trepidation, since I don't want to encourage quantum-mechanical approaches to the study of consciousness) the *wave collapse illusion*: If a given generic phenomenal presentation is *clear* because it presents itself as affording detail queries of a certain kind, and the phenomenal presentation changes from generic to detailed as a result of such a query, then this transition is not noticed as a phenomenal transition. This is actually a simplification, as I think there are two kinds of clarity: being presented (1) as affording answers to queries effected by modulation of sensory apparatus (e.g., moving the eyes), and (2) as affording answers to queries effected by attention modulation alone without modulation of sensory apparatus; both contrast with being presented as query-resistant, for example as occluded or blurry. For purposes of this commentary I speak here of a univocal *clarity*, though if pressed I would hypothesize that the wave-collapse illusion is present, or strongest, only in the latter sort of case. (As I understand it, the refrigerator light illusion is more radical, because it claims that when something can be expected to be present upon query, then the transition from *nothing* to the presence of that thing is not noticed as a phenomenal change. It is not obvious to me that we are really subject to this illusion.)

So, most of the words on the page are generically phenomenally present; some at the small region of foveation are present in a phenomenally detailed way, and some near this area, but just outside it, are presented as clear but only as the generic *text word*. When the eyes are moved to foveate the word, the added sensory detail effects a transition from the phenomenally generic *text word* to a phenomenally specific and detailed presentation of that text word. Now, because the genera employed can be wrong, it is possible to arrange for a situation where all the "words" in this penumbral area are switched around when an eye movement to that area is detected. The subject will notice nothing because (1) the genus *text word* was used for the clear generic; (2) that genus remained applicable when more detail was available; and (3) the wave collapse illusion was in effect.

When being asked to count ball passes between people clothed in white, the perceptual system employs the ad hoc genera

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

person-clothed-in-white and *person-clothed-in-black*. A person in a gorilla suit walking through the middle of the scene is phenomenally presented as clear and under the genus *person-clothed-in-black*. Many instances of change blindness would seem to be cases where some corresponding regions or objects in two scenes or images are phenomenally present as clear but generic, and under the same genus – even though either that one genus is applicable to both but the details differ, or in one case the genus is actually wrong.

If this admittedly sketchily presented hypothesis is on the right track, then we *are* under an illusion regarding the phenomenal content we entertain – but not a terribly radical illusion. We seem to have a full, richly detailed, phenomenal representation of the visual scene, though in fact what we have is, albeit full and clear (in the specified sense), only actually detailed in some places and in some respects, and in other places and other respects it is clear but *generic*. This might seem like no large departure from what some radicals have defended, but in fact it provides traction at exactly the point where the slide from "no fully detailed rich internal representation" to "no internal representation" occurs.

What is cognitively accessed?

doi: 10.1017/S0140525X07002853

Gilbert Harman

Department of Philosophy, Princeton University, Princeton, NJ 08544-1006.
harman@princeton.edu
http://www.princeton.edu/~harman/

Abstract: Is Block's issue about accessing an experience or its object? Having certain "flow" experiences appears to be incompatible with accessing the experience itself. And any experience of an object accesses that object. Such access either counts as cognitive or does not. Either way, Block's issue seems resolvable without appeal to the scientific considerations he describes.

Block is concerned with the question of whether there are cases of *phenomenology* in the absence of cognitive access. I assume that, more precisely, the question is whether there are cases in which a subject *S* has a phenomenological experience *E* to which *S* does not have *direct* cognitive access? (*S* might have indirect cognitive access to *E* through scientific reasoning. I take it that this is not the sort of cognitive access in question.)

It is somewhat unclear – in at least two ways – what Block means by "cognitive access." First, it is unclear what cognitive access is supposed to be access *to*. Second, it is unclear what makes access *cognitive*.

Let me begin my discussion with the first question, about what cognitive access is supposed to be access *to*. Suppose first that *E* is not an experience of or awareness of a phenomenal or intentional object *X*. Then it would seem that the relevant cognitive access can only be to *E* itself, that is to *S*'s having *E*. On the other hand, suppose that the relevant phenomenological experience *E* involves being aware of something *X*, the phenomenal or intentional object of *E*, what *S* is aware of in having experience *E*. In that case, would the relevant cognitive access be access to *X* or to *E*?

It may be that *S*'s experience of *X* is compatible with and perhaps even sufficient for *S* to cognitively access X, although *S*'s trying cognitively to access *E* is incompatible with *S*'s having the experience *E*. In such a case it would seem that *S* does not have (direct) cognitive access to *E*. For example, *S*'s being completely engaged in what *S* is doing – as in optimal "flow" experience (Csikszentmihalyi 1990) – is compatible with (and even sufficient for) *S*'s cognitively accessing what *S* is doing while at the same time at least sometimes being incompatible with *S*'s being aware of being so engaged. In such cases it would seem that *S* does not have direct cognitive access to having such flow experiences.

So, if the relevant cognitive access is access to *E*, there seem to be clear cases of phenomenology without cognitive access, and no need for the sort of investigation Block describes.

Alternatively, the relevant cognitive access might be to the (intentional or phenomenal) object *X* of *S*'s experience *E*. This seems to be the sort of cognitive access Block has in mind. Consider his discussion of the subjects in Sperling's (1960) experiment who reported being aware of all the items in a briefly displayed grid even though they could identify only some of the items. The items in question are the *objects* of the subject's perceptual experiences.

But can one have a phenomenal experience of *X* without having cognitive access to *X*? Indeed, can one have a phenomenal experience of *X* without that phenomenal experience of *X* being a cognitive experience of *X* that constitutes cognitive access to *X*? Or, to put the question the other way, can there be a phenomenal experience of *X* that is not itself a cognitive experience of *X*? What could possibly distinguish a cognitive experience of *X* from a noncognitive experience of *X*? Clearly, it depends on what is meant by "cognitive" experience.

Suppose that a necessary condition of *E*'s being a "cognitive" experience of *X* is that *E* should have a certain sort of "intentional content" – an experience of *X*'s being *F*, for some relevant *F*. Given this supposition, if *S*'s having *E* is a cognitive experience, *E* consists at least in part in *X*'s seeming (appearing, looking, etc.) to *S* to be *F* for some relevant *F*. Furthermore, it might be suggested that, *S*'s having a *phenomenal* experience *E* of *X* need not (and maybe never does) consist even in part in *X*'s seeming to *S* to be *F*. If so, then given the aforementioned supposition, it would seem that there could be a phenomenal experience *E* that is not cognitively accessible.

So it may seem highly relevant whether some *S* could have phenomenal experiences of *X* that did not consist at least in part in *X*'s seeming to *S* to be *F* for some relevant *F*. In Harman (1990) I argue (in effect) that the answer is "No, this is not possible." In fact, I argue that phenomenal content is the same thing as intentional content, a conclusion that is widely (but not universally) accepted in recent philosophical discussion.

In this (controversial) view, *S* must have cognitive access to the object of a phenomenal experience *E* because any phenomenal experience of *X* is itself a cognitive experience of *X*. Not only are the objects of phenomenal experience cognitive accessible, they are ipso facto cognitively accessed. If Block assumes that phenomenal content is not always intentional content, so that the phenomenal content of experience is not always cognitively accessed, this assumption by itself appears to guarantee a negative answer to Block's question without any appeal to the scientific considerations he mentions.

The "mesh" as evidence – model comparison and alternative interpretations of feedback

doi: 10.1017/S0140525X07002865

Oliver J. Hulme[a] and Louise Whiteley[b]

[a]*Anatomy Department, Wellcome Trust Centre for Neuroimaging, University College London, London WC1E6BT, England, United Kingdom;* [b]*Gatsby Computational Neuroscience Unit, University College London, London WC1N 3AR, England, United Kingdom.*
o.hulme@ucl.ac.uk **l.whiteley@ucl.ac.uk**
http://www.gatsby.ucl.ac.uk/~louisew/

Abstract: We agree that the relationship between phenomenology and accessibility can be fruitfully investigated via meshing, but we want to emphasise the importance of proper comparison between meshes, as well as considerations that make comparison especially difficult in this

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

domain. We also argue that Block's interpretation of the neural data in his exemplar mesh is incorrect, and propose an alternative.

What Block calls a "mesh" can also be considered a "model" of the interrelationship between neural and psychological data, and of the relationship of both sources of data to the concepts of phenomenology and accessibility. This kind of model fitting is massively underdetermined, and there is a real danger of overfitting – massaging the data to create meshes that can support any assumption. In a Bayesian framework, competing quantitative models are evaluated by comparing their *marginal likelihoods* – how likely the observed data is under each model. The marginal likelihood incorporates an Occam's razor–like penalty, penalising complex models that have more parameters, and are thus *more* likely to overfit the data (Mackay 2004). The kind of "mesh" at issue here is clearly not yet a quantitative model for which a marginal likelihood can be computed, but the principles of model comparison should still be followed as closely as possible.

Block presents a single mesh whose explanatory power gives us "reason to believe" (target article, sect. 1, para. 2) the assumption embedded in it. We would like to emphasise the necessity of comparing fully developed competing meshes before we can have more reason to believe Block's assumption than its counterpart – a single entrant in a race will always be the winner. We will now use Block's example to demonstrate some of the issues inherent in evaluating whether a particular mesh really does provide the "best explanation" in this domain.

In his example, Block claims that without making the assumption that accessibility is constitutive of phenomenology it is not possible to build a mesh that provides a mechanism for overflow. The starting point is an argument from behavioural data that phenomenology overflows accessibility, which is then taken as a given for the mesh to explain. This conclusion is itself a kind of inference to the best explanation, which ultimately contains an assumption about the answer to the methodological puzzle. The importance of the puzzle is diminished by embedding an assumption about the answer within a larger explanatory structure. However, how much the structure relies upon such assumptions will affect the strength of the evidence it provides, and it is therefore essential to be extremely precise about which assumptions are made, and where.

Another issue that makes model comparison particularly difficult in this domain is the imprecise terminology. In his example, Block uses unitary concepts of accessibility and reportability to refer to different phenomena, in different contexts. This can lead to confusion. For instance, whatever is currently part of access-consciousness is reportable in the sense that it can be reported at any time, without any reorienting of attention or the sense organs. Objects in a Sperling task are reportable only if attended, but no terminological distinction is made between these two types of reportability even though the distinction is critical to Block's disentangling of phenomenology and accessibility. We propose that a more precise taxonomy of the different types of accessibility and reportability would be easy to develop and would resolve much of this confusion.

A significant source of difficulty in building and comparing meshes is how we should describe the neural data, and how it maps to psychological states. Block acknowledges that the neural data he invokes may be "wrong or at least highly incomplete" (sect. 13, para. 3). Therefore, great care should be taken at this stage because an error here can cascade through the mesh, severely limiting the strength of the evidence it provides. In demonstration of this point we suggest an alternative mapping between neural states and phenomenology which, in our opinion, would strengthen the mesh by better accounting for the data, or at least cast doubt over Block's version.

Block proposes that cortico-cortical (CC) recurrent loops from higher to lower areas are the core neural basis for phenomenology – for example, that feedback from V5 to V1 is constitutive of phenomenal motion. We propose that motion phenomenology should instead be viewed as being composed of different component phenomenal characters that correspond to the activity of independent but interacting cortical areas, and that recurrent feedback is then necessary for *binding* these phenomenal characters together. As an illustrative oversimplification, the phenomenal experience of a simple moving grating could be characterised as being composed of a phenomenal experience of a retinotopic array of contrast elements, and a phenomenal experience of motion within a particular retinotopic region. It has been argued elsewhere that the best candidate for the core neural basis of the phenomenal experience of contrast elements is V1 (Pollen 1999) and of motion is V5 (Zeki & Bartels 1999). Under normal conditions the two attributes are bound to each other and one directly experiences the surface qualities of the thing that is moving – a *modal* experience. If the two are not bound, one will have an *amodal* experience of motion.

This hypothesis can be tested by looking at what happens when feedback from V5 to V1 is prevented. Block claims that V5 activity over a certain threshold in the absence of feedback to lower areas is not sufficient for the experience of motion (sect. 13). We would claim that it is in fact sufficient for an *amodal* phenomenal experience, supported by the observation (cited by Block in his Note 10) that patient G.K. can experience fast motion in his "blind field," described as looking like "black moving on black" (Zeki & ffytche 1998), despite having no possibility of recurrent feedback from V5 into V1/V2 in his damaged hemisphere (Semir Zeki, personal communication).

The same logic can be applied to the transcranial magnetic stimulation (TMS) evidence that disrupting feedback from V5 to V1 prevents the perception of a moving phospene (Pascual-Leone & Walsh 2001). Subjects reported that the phosphene was "present but stationary," which is equally consistent with the role of CC feedback in binding phenomenal contents. A true test would be whether subjects can distinguish between sham and real V5 stimulation in the context of disrupted feedback to V1, and whether this is correlated with an experiential report of amodal motion. Adopting this alternative mapping would potentially strengthen Block's mesh by better fitting the neuroscientific evidence.

Block's proposed approach compares meshes as a proxy for a direct comparison of competing hypotheses about the relationship between accessibility and phenomenology. To make this comparison meaningful, we must be very precise about the necessity of the hypothesis to the explanatory power of the mesh. From this perspective the explanatory gap may prove to be more like a crack in the pavement than a gaping hole, but we must still take great care in stepping over it.

Many ways to awareness: A developmental perspective on cognitive access

doi: 10.1017/S0140525X07002877

Carroll E. Izard, Paul C. Quinn, and Steven B. Most
Department of Psychology, University of Delaware, Newark, DE 19716.
izard@udel.edu pquinn@udel.edu most@udel.edu

Abstract: Block's target article makes a significant contribution toward sorting the neural bases of phenomenal consciousness from the neural systems that underlie cognitive access to it. However, data from developmental science suggest that *cognitive* access may be only one of several ways to access phenomenology. These data may also have implications for the visual-cognitive phenomena that Block uses to support his case.

Prelingual infants perform mental operations that signal levels of awareness beyond phenomenal consciousness and that may be observed or indexed by processes other than those involved in cognitive accessibility. These include: (1) using action selection processes to group stimuli into perceptual categories, thereby

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

laying groundwork for subsequent concept formation; (2) recognizing and responding appropriately to faces and facial expressions of basic emotions and making stable emotion-cognition connections; and (3) exhibiting movements indicative of intentionality, goal-directed behavior, and problem-solving (Slater & Lewis 2007; cf. Merker 2007). These behaviors may reflect the development of different levels or complexities of awareness and offer possibilities of extending current conceptualizations of ways to access phenomenal experience.

Mechanisms of accessibility and levels of awareness. Developmental scientists routinely observe evidence suggesting that prelinguistic infants not only experience objects and events phenomenologically, they respond to them in meaningful ways. They discriminate between animate and inanimate faces (Ellsworth et al. 1993), familiar and strange persons (Bushnell et al. 1989), and among a wide variety of objects sufficiently well to place them in categories (Quinn 2006). As well, they respond differentially to others' emotion expressions (Walker-Andrews 1998), and execute movements to influence persons, objects, and events, and to solve problems (Sommerville & Woodward 2005).

Foundations of concept formation. Through visual tracking and eye movements, young infants have shown that they can parse visual experiences into perceptual groups that subsequently attain conceptual significance. For example, 3-to-4–month-olds presented with visual images of realistic photographs of cats generalize their looking-time responsiveness to subsequently presented novel cats, but use comparison and selection processes to display visual preferences for exemplars from novel categories including birds, dogs, and horses (Eimas & Quinn 1994; Oakes & Ribar 2005; Quinn et al. 1993). Studies demonstrating categorization abilities early in life indicate that infants do not experience objects in the world as undifferentiated, but as separate groups that fall into distinct representations (Quinn & Eimas 1996). These representations may then serve as placeholders for the acquisition of the more abstract and non-obvious information that occurs beyond infancy, through language and more formal learning of semantic categories (Quinn & Eimas 1997; 2000). Thus, over time, the perceptual placeholder representation for cats will come to include the information that cats eat tuna, hunt mice, give birth to kittens, have cat DNA, and are labeled as "cats."

Emotion expression. In mother-infant face-to-face play, young infants display emotion expressions in synchrony with the mother but also periodically take the lead, indicating behavior other than imitation (Stern 1974; Tronick 1989). Their behavior in the "still-face" procedure is typically interpreted as an effort to re-engage the mother (Muir & Lee 2003). 2.5-to-9–month-old infants respond differentially and predictably to the mother's discrete emotion expressions (Izard et al. 1995; Montague & Walker-Andrews 2001). Through their emotion expressions, children without a cerebral cortex also show evidence of access to phenomenal experience (Merker 2007).

Body movement/action. Kinematic data have shown that 10-month-old infants reach more rapidly for a ball that they have been encouraged to throw into a basket than for one they have been encouraged to fit into a plastic tube (Claxton et al. 2003). From ages 9 to 19 months, infants show clear developmental changes in target selection and movements that ultimately lead to problem solving (McCarty et al. 1999).

The three foregoing types of data suggest that prelingual infants have emotion- and action-systems that mediate access to contents of phenomenal experience independently, or largely independently, from those involved in cognitive access. Thus young infants (like hydranencephalic children; cf. Merker 2007) appear to possess forms of accessibility that may lie outside the pale of Block's "cognitive" criteria.

Independence and interdependence of access modes. Evidence suggestive of accessibility via emotion- and action-systems (in the absence of cognitive accessibility) can be found in adults as well as in children. However, due to maturation and resultantly enriched connections among neural systems, examples of functionally independent access pathways may become rarer with age. In adults, observations of such independence might often require artificially constrained manipulations or incidental brain lesions. Milner and Goodale's (1995) work with patient D.F., for example, suggests that dorsal and ventral visual pathways output separately to action and cognitive systems respectively (also see James et al. 2003). When asked to indicate explicitly the orientation of a slot, D.F. (who had a compromised ability to process information via her ventral visual pathway) was unable to do so. However, when asked to insert a card into the slot as if posting a letter, her action was immediate and correct. Examples also abound of situations where emotion processes seem to glean aspects of experience that lie beyond reach of cognitive accessibility. These include emotion's role in decision making (Bechara et al. 2000) and in the power of emotional stimuli to guide attention even when people cannot report them (Jiang et al. 2006).

Recent evidence suggests that contents accessible to the neural systems of emotion can be made available to the systems of cognitive access (e.g., particularly emotional task-irrelevant stimuli appear to gain access to explicit report mechanisms at the expense of non-emotional target stimuli; Arnell et al. 2007; Most et al. 2005a). However, the quality and function of the experience change after the emotion-cognition connection (Izard, in press). For example, when people label emotions, there follows an observable activation decrease in neural areas associated with emotional reactivity (Lieberman et al. 2007). Emotion-cognition-action connections and interactions have played a critical role in the evolution and functioning of consciousness and continue to influence the development of higher levels of awareness in ontogeny.

Implications for visual cognition. Evidence of multiple routes for accessing phenomenology might help reframe findings from the visual cognition literature, several of which Block described in making his argument. Rather than casting phenomena such as the attentional blink (Chun & Potter 1995; Raymond et al. 1992), inattentional blindness (Mack & Rock 1998; Most et al. 2001; 2005b; Simons & Chabris 1999), and object substitution masking (Di Lollo et al. 2000; Reiss & Hoffman 2006; Woodman & Luck 2003) in terms of a conscious versus unconscious divide, it may be more fruitful to regard them as failures of cognitive access but not necessarily as failures of other types of access. The limitations that constrain cognitive accessibility might not generalize to other forms of accessibility. For example, evidence suggests that whereas people often fail to detect neutral targets during an attentional blink, emotion stimuli are much less susceptible to this effect (except in cases of bilateral amygdala damage; Anderson & Phelps 2001).

A different look at accessibility and reportability of levels of awareness. Developmental research potentially allows us to examine a time before connections between cognition-, emotion-, and action-systems are fully formed, thereby providing insights that might not be as readily gleaned from, but which may affect interpretation of, the adult literature (e.g., evidence of separable access systems). The developmental data also have implications for Block's endeavor to pinpoint neural bases of phenomenal experience unadulterated by access mechanisms. His suggestions for ruling out the machinery of cognitive accessibility are important, but ruling out mechanisms involved in accessibility via emotion and action processes may prove even more challenging. Block's ability to construct an empirically based proposition for how phenomenology might overflow cognitive accessibility gives testimony to the burgeoning of cognitive science investigations relevant to this enterprise. Empirical investigations of similar constraints on emotion- and action-accessibility have yet to catch up. We hope that Block's path-making endeavor will kick-start this process.

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

What is "cognitive accessibility" accessibility to?

doi: 10.1017/S0140525X07002889

Pierre Jacob

Institut Jean Nicod, UMR 8129, EHESS/ENS/CNRS, DEC, Ecole Normale Supérieure, Pavillon Jardin, 75005 Paris, France.

Pierre.Jacob@ehess.fr

http://www.institutnicod.org

Abstract: I first argue that some of Block's formulations may misleadingly suggest that the function of mechanisms of so-called cognitive accessibility is to make one aware, not of visible features of the visible world, but of one's own psychological life. I then ask whether Block's view of phenomenology in the present target article is consistent with his endorsement of non-representationalism elsewhere.

According to Block's diagnosis, the present scientific evidence is compatible with the hypothesis that the neural machinery underlying cognitive accessibility is not a constitutive part of the neural machinery underlying visual phenomenology. The evidence shows that locally recurrent activity within the occipital areas is not sufficient for cognitive accessibility: the latter requires broadcasting to a global workspace involving frontal and parietal areas by means of long-range projections from the occipital areas. But the evidence does not rule out the hypothesis that recurrent activity within occipital areas is sufficient for visual phenomenology. I find Block's diagnosis compelling and his hypothesis quite plausible. Here, I want to press him on a pair of conceptual issues respectively raised by his present conception of cognitive accessibility and by his present conception of visual phenomenology.

Block's hypothesis presupposes the acceptance of some such psychological distinction as the distinction between belief (or judgment) and phenomenal experience. One can visually experience (or see) an object that exemplifies the determinate property *F* without using (or even possessing) the concept of *F*. But one cannot believe that some object is *F* unless one possesses the concept of property *F*.

As Block recognizes, the belief–experience distinction faces a direct challenge from the behaviorist reliance on reportability. Since the content of one's report depends on what one believes, not on the content of one's visual experience, one basic reason for rejecting the distinction between experiencing and believing is the behaviorist suspicion that, unlike the content of one's belief, the content of one's experience runs the risk of being unreportable and thereby escaping the scope of scientific investigation.

If one accepts the belief–experience distinction, then the question may also arise: What does it take to be aware of one's own experiences? At one extreme is Dretske's (1993; 1994) view that one might have a conscious experience and not be conscious of having it. At another extreme is Dennett's (1991) first-person operationalist view, according to which one cannot be aware of a stimulus unless one believes that one is aware of it.

Block rejects first-personal operationalism. But instead of endorsing the view that one might have a conscious experience and not be conscious of it, he considers a deflationary view and the "same order" view. I wonder why. On the deflationary view, one is supposed to experience one's experience just as one dances one's dance or one smiles one's smile. On the same order view, a conscious experience is reflexive. The deflationary view sounds to me like a failed attempt at dissolving rather than solving the problem of self-awareness. Certainly, one is not having a visual experience of one's own visual experience of a red rose when one is having a visual experience of a red rose. Nor is it clear that the same order view can accommodate Block's explicit purpose of providing a unified account of experience: Is a mouse in a reflexive state of awareness when it sees a piece of cheese?

Furthermore, Block sometimes gives, I think, the misleading impression that the function he assigns to the mechanisms of cognitive accessibility is to make one aware, not of features of distal stimuli, but of one's own phenomenology or of features of one's own phenomenally conscious experiences. For example, early on he introduces cognitive accessibility in terms of Fodor's criterion of modularity according to which we do not have cognitive access to some of our own perceptual states and representations. This is slightly misleading because Block's view is that by being broadcasted to the global workspace for further processing, the content of a dominant visual representation makes one cognitively aware of visible features of one's non-mental environment, not of one's own psychological life (or computational architecture).

I now turn to Block's present conception of visual phenomenology. So far as I can see, none of Block's arguments in this paper presupposes his (2003; 2007) anti-representationalist view of the character of phenomenal visual experiences. Given my own representationalist bias, this independence makes his arguments for the view that the neural machinery of visual phenomenology does not include the neural machinery of cognitive accessibility more easily acceptable. In particular, I fully concur with the main line of Block's accounts of the Sperling experiment, change blindness experiments, and the Landman et al. experiment.

Consider Block's present account of the Sperling experiment. To say of subjects that they visually experience the whole array of 12 alphanumeric characters is to say that they are able to bring each character under the general concept "letter." The reason they fail to report more than 4–5 of such characters is that they fail to bring more than 4–5 of them under the concepts of their distinctive shapes. On this view, the content of a reportable representation seems to stand to the phenomenal character of a non-reportable experience of one and the same stimulus just as the concept of a determinate shape property (e.g., rectangle) stands to the concept of a determinable shape property (e.g., polygon). If so, then the question arises: Is the rejection of representationalism, which Block has endorsed elsewhere, consistent with his present account of visual phenomenology?

The question, I think, is made more pressing by the fact that Block is willing to draw a contrast between two memory systems with different storage capacities: The capacity of the so-called visual phenomenal memory system is said to overflow the capacity of working memory. But on the face of it, the argument for the view that the neural machinery for visual phenomenology does not contain the neural machinery for cognitive accessibility seems to presuppose that only contents that are cognitively accessible, not the contents of visual phenomenology, are available for further processing by such cognitive mechanisms as attention, memory, and reasoning. Would Block be willing to postulate a special phenomenal attention system and a special phenomenal reasoning system for visual phenomenology? If not, then why not? Why single out visual phenomenal memory among other cognitive mechanisms?

Incomplete stimulus representations and the loss of cognitive access in cerebral achromatopsia

doi: 10.1017/S0140525X07002890

Robert William Kentridge

Department of Psychology, University of Durham, Durham DH1 3LE, England, United Kingdom.

robert.kentridge@durham.ac.uk

Abstract: When processing of stimuli occurs without attention, phenomenal experience, as well as cognitive access, may be lost. Sensory representations are, however, constructed by neural machinery extending far beyond sensory receptors. In conditions such as cerebral achromatopsia incomplete sensory representations may still elicit

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

phenomenal experience but these representations might be too aberrant to be integrated into the wider cognitive workspace.

In the target article Block aims to identify circumstances in which stimuli might elicit phenomenal experience but not elicit a representation which is cognitively accessible. In the examples Block considers, cognitive accessibility is impeded through deficits in attention (the extinction example) or through limitations in processing time (the partial report examples). Both mechanisms depend upon some constraint in processing the stimulus. There must be a concern that this might affect sensory processing, and hence that phenomenal experience of the stimulus is affected along with cognitive access. We know, for example, that attention affects visual sensitivity (Solomon 2004). Changes in the response gain of neurons in sensory areas of cortex are as likely to be affecting phenomenal experience as cognitive access (Carrasco et al. 2004; Treue & Martìnez Trujillo 1999). Are there circumstances in which the object of sensory processing can be examined at leisure and can be fully attended, and yet cognitive access is lost when phenomenal experience survives?

What does it mean for a representation to be cognitively accessible? As Block notes, "mechanisms of reporting, reasoning, evaluating, deciding, and remembering" (sect. 11, para. 2) should be able to make use of such a representation – a key aspect of cognitive access in models such as Baars' Global Workspace Theory (Baars 1997) is that information about a stimulus becomes widely available. The representation must therefore take a form that permits interaction with memories or plans about other stimuli (in Piagetian terms, for a cognitive system to assimilate a representation, the former must be accommodated to the latter). If stimuli no longer engage representations that can be integrated in any sense with the rest of a person's representational framework, then surely they cannot be said to be cognitively accessible.

How can representations become isolated from cognition? One might erroneously assume that sensations are "raw" – that they do not need to be prepared in a fashion to make them accessible to cognition. Sensations do not, however, correspond simply to the activity of sensory receptors. In color vision, for example, color appearance is far more closely related to a distal property of surfaces (things in the world about which we have cognitions), their spectral reflectance, than to the activations of cone photoreceptors in the retina (Hofer et al. 2005). Color sensation is the endpoint of a complex process. If this process is prevented from running to its conclusion we may be left with an incomplete signal that gives rise to sensation but cannot be integrated with cognition. Stimuli may therefore potentially elicit phenomenal experiences and these experiences may be discriminable yet they remain cognitively isolated. There is no framework within which to reason about or evaluate their differences, nor can they be remembered (I am not sure I see mere discrimination or decision as an act of cognition, but neither does Block when he discusses "guessing" in blindsight).

Do such circumstances exist? I will argue there are neurological patients who have all the time in the world to process stimuli, who can attend to them, but who have sensations divorced from cognition.

Cerebral achromatopsia is a neurological condition in which color vision is lost as a consequence of damage to ventromedial occipital cortex, usually in the vicinity of the fusiform and lingual gyri (Meadows 1974). Unlike more usual forms of color blindness there are no deficits or abnormalities in the retinal cone photoreceptors which form the starting point of normal color perception. Patients with cerebral achromatopsia do not usually make spontaneous comments about color. Questions about color sensation seem quite alien to them. They are unable to name the colors of stimuli presented to them or to perform nonverbal tests of color perception such as sorting or odd-one-out tasks. Although they may remember some semantic associations of color words (e.g., that bananas are yellow), they appear to have no sensation of color or means of remembering any aspects of the nature of color sensation (see, e.g., Heywood & Kentridge [2003] for a recent review). It is, however, well known, that cerebral achromatopsics do see (quite consciously) the border formed between regions of equiluminant color (Heywood et al. 1998). If a red and a green that a patient cannot tell apart are used to construct a red square against a green background, then the patient will effortlessly see the square although they will be unable to explain how the square and its background differ. My colleagues and I (Kentridge et al. 2004) have shown that not only can these patients see such color borders but they can discriminate between borders formed from different colors (and that this discrimination cannot be accounted for on the basis of chromatic contrast salience) even though they cannot see the colors of the adjoining surfaces that form the borders. Again, the discrimination is conscious. The borders somehow look different from one another. The patients cannot, however, explain in any sense how or why the borders differ (they have no accessible representations of color) – they just know that the borders look different. The patient is surely having differing phenomenal experience of these borders, yet these experiences are not cognitively accessible. It is true that they know of the existence of these borders, but subjects in a partial report experiment know of the existence of items they cannot describe. I suggest that this situation, in which sensory representations simply cannot be integrated into the global workspace, provides a better example of phenomenology in the absence of cognitive access than cases in which the integration of representations into the workspace is possible but temporarily unachieved.

Phenomenology without conscious access is a form of consciousness without top-down attention

doi: 10.1017/S0140525X07002907

Christof Koch[a] and Naotsugu Tsuchiya[b]

[a]*Division of Biology, California Institute of Technology, Pasadena, CA 91125;* [b]*Division of Humanities and Social Sciences, Psychology and Neuroscience, California Institute of Technology, Pasadena, CA, 91125.*

koch.christof@gmail.com

http://klab.caltech.edu/~koch/naotsu@gmail.com

http://www.emotion.caltech.edu/~naotsu/Site/index.html

Abstract: We agree with Block's basic hypothesis postulating the existence of phenomenal consciousness without cognitive access. We explain such states in terms of consciousness without top-down, endogenous attention and speculate that their correlates may be a coalition of neurons that are consigned to the back of cortex, without access to working memory and planning in frontal cortex.

We agree with Block's hypothesis that phenomenally conscious states may sometimes not be cognitively accessible. Partial report and dual-tasks paradigms show that we have only limited access to some aspects of phenomenal experience. Or, to adopt Block's earlier language, phenomenal consciousness can occur without access consciousness, a revision of our earlier position (Crick & Koch 1998a; Koch 2004). We argue here that sensory psychology has a long-standing framework to consider such cases, involving attentional selection processes.

We recently (Koch & Tsuchiya 2007; Tsuchiya & Koch 2008) summarized the empirical evidence that consciousness and top-down, volitionally controlled endogenous attention are distinct neurobiological processes with distinct functions (see also, Iwasaki 1993; Lamme 2003). In particular, sensitive psychophysical techniques can dissociate these two. Indeed, a range of phenomena exists in which subjects are conscious of certain stimuli attributes without top-down attention. This list includes

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

the pop-out target in a visual search task, gist perception of a scene, categorization of peripheral targets when a very intensive cognitive-resource task has to be performed at fixation, and iconic memory.

Consider Sperling's original iconic memory experiment (Sperling 1960) or Landman et al.'s (2003) variant. Subjects report that they clearly, vividly, and consciously see a field of letters or a bunch of bars arranged on a circle. This is also what we experience when we look at such displays. However, it is well known that subjects have only very limited access to the detailed properties of the individual elements, unless top-down attention is directed to a subset of stimuli using appropriately timed cues. Our basic point is that phenomenology without conscious access is an example of consciousness without top-down attention processing, though the converse is not true; that is, not every example of conscious perception in the absence of top-down attention is cognitively non-accessible. For example, the gender of a briefly presented face can be accurately reported even if subjects are engaged in a highly demanding task at the fixation (Reddy et al. 2004).

So what is the story at the level of the brain? Decades of electrophysiological recordings in monkeys have proven that the spiking response of neurons in the ventral visual stream (e.g., in areas V4 and IT) representing attended stimuli is boosted at the expense of the response to non-attended items. According to Crick and Koch (1995), this enables these neurons to establish a reciprocal relationship with neurons in the dorsolateral prefrontal cortex and related regions that are involved in working memory and planning (and language in humans), leading to reverberatory neuronal activity that outlasts the initial stimulus duration. Critical to the formation of such a single and integrated coalition of neurons are the long-range axons of pyramidal neurons that project from the back to the front of cortex and their targets in the front that project back to the upper stages of the ventral pathway (possibly involving stages of the thalamus, such as the pulvinar [Crick & Koch 1998b], and the claustrum [Crick & Koch 2005]). The subject now consciously sees these stimuli and can report on their character (e.g., identify the letter [Sperling 1960] or the orientation of the square [Landman et al. 2003]). Furthermore, the subject also has a strong conscious sense of the entire scene ("I see an array of letters") that is likewise mediated by a loop that involves the inferior temporal cortex and the frontal lobes half-way across the brain.

But what happens to those stimuli that do not benefit from attentional boosting? Depending on the exact circumstances (visual clutter in the scene, contrast, stimulus duration) these stimuli may likewise establish coalitions of neurons, aided by local (i.e., within the cortical area) and semi-local feedback (i.e., feedback projections that remain consigned to visual cortex) loops. However, as these coalitions of neurons lack coordinated support from feedback axons from neurons situated in the prefrontal cortex, thalamus, and claustrum, their firing activity is less vigorous and may decay much more quickly. Yet, aided by the neuronal representation of the entire scene, these weaker and more local coalitions may still be sufficient for some phenomenal percepts.

Block cites functional magnetic resonance imaging (fMRI) studies of patients with visuospatial hemi-neglect (Driver & Vuilleumier 2001; Rees et al. 2000; 2002b) that offer evidence of enhanced hemodynamic activity in the fusiform face area contralateral to a face that the patient is not aware of. For Block, this raises the question of whether this is likewise an example of phenomenal consciousness without cognitive access. We answer this question clearly in the negative. First, one should trust the first-person perspective: That is, in the absence of compelling, empirical evidence to the contrary (such as Anton's blindness, also known as hysterical blindness; Sackeim et al. 1979), if the subject denies any phenomenal experience, this should be accepted as a brute fact. If we take the existence of mere recurrent, strong neuronal activation as evidence for consciousness, why not argue that the spinal cord or the enteric nervous system is conscious but is not telling me (Fearing 1970)? Second, the relationship between neuronal firing activity and the associated hemodynamic BOLD response is a very complex one. In particular, there are well-documented cases where a vigorous fMRI signal is observed in the absence of any spiking from the principal neurons in that area (Harrison et al. 2002; Logothetis 2003; Logothetis & Wandell 2004; Mathiesen et al. 1998). Synaptic activity is a much larger driver of hemodynamic activity than are action potentials. Therefore, a much more cautious reading of these studies is that they demonstrate synaptic input into the fusiform face area in these patients; however, whether or not this input is vigorous enough to establish a sustained coalition of neurons is totally up in the air and requires further investigations.

In conclusion, the quiddity of the neuronal correlates of conscious access are long-range loops between the back and the front of cortex and its associated satellites (thalamus, basal ganglia and claustrum), enabled by top-down attention. Without this amplification step, most coalitions in the back are fated to die; however, given the right conditions, a few may survive and may be consciously experienced by the subject. Yet, as the informational content of these coalitions are not accessible to working memory and planning circuits in the front, the subject cannot consciously access the detailed stimulus attributes. Our explanation provides a plausible account of how phenomenal consciousness can occur without cognitive access.

Partial awareness and the illusion of phenomenal consciousness

doi: 10.1017/S0140525X07002919

Sid Kouider, Vincent de Gardelle, and Emmanuel Dupoux

Laboratoire de Sciences Cognitives et Psycholinguistique, Ecole Normale Supérieure, 75005 Paris, France, and CNRS/EHESS/DEC-ENS, 75005 Paris, France.

sid.kouider@ens.fr **gardelle@ens.fr**
dupoux@lscp.echess.fr **www.lscp.net**

Abstract: The dissociation Block provides between phenomenal and access consciousness (P-consciousness and A-consciousness) captures much of our intuition about conscious experience. However, it raises a major methodological puzzle, and is not uniquely supported by the empirical evidence. We provide an alternative interpretation based on the notion of levels of representation and partial awareness.

In his target article, Ned Block is dealing with a difficult problem: how to empirically demonstrate that phenomenal consciousness (hereafter P-consciousness) is dissociable from access consciousness (hereafter A-consciousness). An a priori argument in favor of this dissociation is the common intuition that the representational content of phenomenal experience is much richer than the limited content we can access at a given time. In Block's words, "phenomenology overflows cognitive accessibility" (sect. 8, para. 6). This intuition is so strong that it appears very easy, at first glance, to show how much richer P-consciousness is, compared with A-consciousness.

However, providing an empirical demonstration of this dissociation leads to a major methodological difficulty: any measure of consciousness seems inevitably to require the involvement of A-consciousness. From there on, it seems impossible to show evidence for P-consciousness without A-consciousness. This methodological puzzle arises also in dissociating consciousness and top-down attention. Demonstrating consciousness without attention seems impossible for similar reasons: To assess consciousness of the stimulus, one needs to direct the subject's attention on the stimulus! Although there is converging evidence that attention can affect both conscious and unconscious

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

perception, the reverse dissociation involving the possibility of consciousness without attention remains highly debated (see Dehaene et al. 2006; Koch & Tsuchiya 2007). Block acknowledges this methodological issue and proposes to take the set of evidence at hand *as a whole* and see whether it points towards the researched dissociation. Using, among others, examples from perception of unattended objects (e.g., attentional blink, change blindness), and from partial report Sperling-like experiments, Block assumes that we should adopt the A- versus P-consciousness dissociation and look for its respective neural bases.

Unfortunately, the evidence Block is using cannot unequivocally prove his theory. Furthermore, we think that the empirical data reviewed in his target article do not provide more support for his accounts over alternative and crucially simpler explanations. In the empirical phenomena that Block is using, one can distinguish two types of situations: those involving partial access and those involving undetectable stimuli.

The first type of situation involves stimuli that are visible but unattended – and importantly, not even detected – implying the absence of any conscious access. This is usually the case during attentional blink and inattentional blindness experiments. Block uses the fact that the stimulus is supra-threshold (it can be reported when attention is drawn towards it) to argue for P-without A-consciousness. But there is no evidence for this claim, since subjects do not even detect the unattended stimulus. Moreover, Block's claim that participants forget their phenomenal episode appears impossible to test, because any probe would modify their attention and hence make the stimulus consciously accessible (Kouider et al. 2007).

In the second type of situation, by contrast, subjects are aware of "some" information: stimuli are presented very briefly and/or in a degraded fashion, such that they are not fully visible but not subliminal either. Block assumes that this situation gives rise to P- without A-consciousness. We propose instead that what happens here is rather a form of partial awareness in the absence of full awareness. Partial awareness reflects the situation where subjects have transient access to lower but not higher levels of representation. For instance, visual word recognition implies the processing of several hierarchically organized levels (e.g., fragments, letters, whole word). With degraded presentation conditions, lower levels can be accessed (e.g., fragments/letters) while higher levels cannot (whole word). Still, subjects can use such partial information in conjunction with context/expectations to make hypotheses about the representational content at higher levels of processing (Kouider & Dupoux 2004). Under this perspective, Block's richness of phenomenal experience can be reinterpreted as the transient activation of a large quantity of degraded low-level information. In such partial awareness conditions, the available information is quantitatively rich but qualitatively poor.

This hypothesis allows us to construe the Sperling phenomenon (see Sperling 1960) as resulting from partial awareness: subjects have a transient and degraded access to fragments of all the letters in the grid. As subjects are not expecting anything other than letters, fragments are used to reconstruct as many letters as possible. Due to mnemonic decay and attentional overflow, subjects are able to reconstruct at most about four letters. Crucially, the unreported items are never identified as letters per se and remain coded as unidentified letter fragments. A similar situation of partial without full access is the McConkie experiment (McConkie & Rayner 1975) in which subjects see "letter-like" fragments in the periphery and infer that these are real letters. In our previous work (Kouider & Dupoux 2004) we have extended this phenomenon to a dissociation between the letter and the word level. We have induced subjects to access some but not all letters of a real or false color word (GREEN or GENER). We found that both the real and false color words are identified and treated as real words, as assessed by both subjective reports and the magnitude of the Stroop effect.

An important question for future research will be to characterize whether such reconstruction processes imply metacognitive/inferential interpretations or rather more direct perceptual illusions. Block acknowledges that McConkie's experiments involve a reconstruction process (what Block labels "cognitive illusions"). However, Block assumes that subjects in Sperling-like experiments are not reconstructing the visual scene but genuinely experiencing the whole set of items. To justify this special treatment, Block argues that the Sperling phenomenon is somehow mandatory and does not require subjects to explicitly report the stimuli, implying that it is a perceptual rather than a metacognitive reinterpretation. Yet, these statements remain highly speculative, as none of them has been empirically demonstrated. For us, it is highly probable that Sperling-like paradigms also lead to the "experience" of letters even when the uncued items consist of false letters. Of course, disentangling this issue requires further empirical research.

All these remarks point towards the same direction: Including a typology in terms of levels of representation during conscious access, along with the associated notion of partial awareness, provides a unified description of the empirical evidence at hands. In particular, this account offers more explicit specifications of the functional mechanisms leading to conscious perception.

Sue Ned Block!: Making a better case for P-consciousness

doi: 10.1017/S0140525X07002920

Victor A. F. Lamme

Department of Psychology, University of Amsterdam, 1018 WB Amsterdam, The Netherlands.

The Netherlands Institute for Neuroscience, part of the Royal Academy of Arts and Sciences (KNAW), 1105 BA Amsterdam, The Netherlands

v.a.f.lamme@uva.nl

www.cognitiveneuroscience.nl

Abstract: Block makes a case for the existence of conscious experience without access. His case would have been much stronger, however, if he had woven fully unconscious processing into the "mesh argument," and considered arguments that are intrinsic to neuroscience.

Sometimes, science looks like a court of law. There is a scientific hypothesis – the defendant – and there are its advocates and opponents – the defense and prosecution. Here, the defendant is conscious experience. It stands accused of not existing in its own right. Conscious experience is what we *say* we see or hear, what we cognitively access and manipulate – so claim the prosecutors. Ned Block is leading the defense team, arguing that we should not equate conscious experience with cognitive access. Many psychological experiments show mental representations that have higher capacity than what is reported by the subject. These mental representations have phenomenal qualities and might just as well be conscious representations. There is sufficient evidence to cast *reasonable doubt* on the accusation. Should the defendant be satisfied with such a defense, or litigate for malpractice? I think the latter. Block has made only half a plea (his "mesh argument" lacks a key consideration), and left out all the forensic evidence (arguments *intrinsic* to neuroscience).

Neurophysiological studies in primates (Supèr et al. 2001b), as well as electroencephalography (EEG) (Sergent et al. 2005), functional magnetic resonance imaging (fMRI) (Haynes et al. 2005), and transcranial magnetic stimulation (TMS) (Silvanto et al. 2005b) studies in human subjects, show that recurrent or re-entrant processing between different regions of the brain is necessary for conscious experience. In the Global Workspace theory (GWT) (Baars 2005; Dehaene et al. 2006), the content of information processed in, say, visual areas, is broadcast and made available for global access by means of recurrent amplification. "Workspace neurons," in prefrontal cortex, are vital to this amplification, because they provide

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

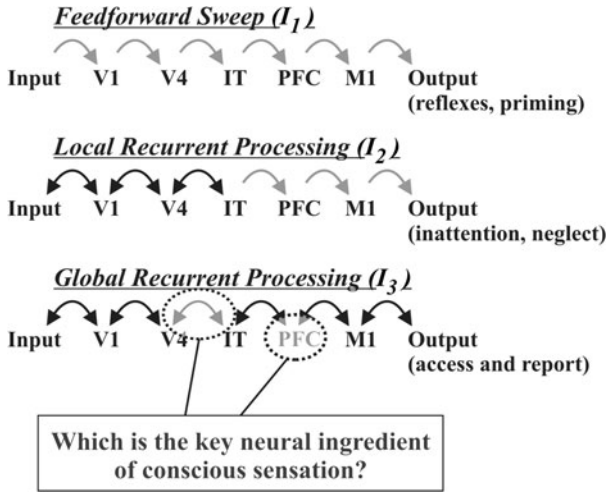


Figure 1 (Lamme). Three stages of visual processing: First, visual information is processed along the sensorimotor hierarchy (V1 to M1) by means of feedforward connections. This constitutes the feedforward sweep (I_1). Depending on attention, subsequent recurrent processing either remains localized to visual areas (V1, V4, IT; I_2) or extends towards areas involved in the planning and execution of movement (PFC, M1; I_3). Phenomenal sensation develops from I_1 (unconscious), via I_2 (P-conscious), to I_3 (A-conscious). See Lamme (2003).

long-range connections between sensory and motor cortices. With global recurrent amplification, conscious experience becomes accessible and reportable (Sergent et al. 2005). However, more localized recurrent interactions, restricted to visual areas and not involving specialized workspace neurons, are also possible and have been reported (Scholte et al. 2006). Advocates of GWT argue that in that case there is no conscious experience, only proto-consciousness, precisely because of the absence of global workspace (prefrontal cortex) activation.

Block, however, argues that such localized recurrent states correspond to phenomenality without access. He converges upon that view through the "mesh argument": If we assume that the neural basis of phenomenality (recurrent processing in visual cortex) does not include the neural basis of access (frontal cortex), we can understand why phenomenality overflows access, as is shown in the Sperling, Landman et al., and Sligte et al. experiments. Although I agree with the conclusion – localized recurrent processing *is* conscious processing – this is an argument that I suspect will convince only part of the jury.

The metaphysical correlationist can sketch a competing mesh argument, interpreting the iconic memory experiments as nonphenomenal, proto-consciousness overflowing "real" consciousness. And he would argue that this corresponds to the difference between processing with or without workspace neuron activation. No need for acquittal of the defendant. The epistemic correlationist would still find both options not scientifically distinguishable. Call for a mistrial.

What we need are independent arguments for attributing phenomenality to localized recurrent processing. The mesh argument should not only take the division between local and global recurrent processing (I_2 and I_3 in Dehaene et al.'s [2006] terms, Fig. 1) into account. The issue becomes much clearer when fully unconscious or inaccessible neural processing (I_1) is also considered. Since there is little disagreement about the absence of conscious experience in I_1, or about its presence in I_3, the question becomes whether I_2 is more like I_1 (i.e., unconscious) or like I_3 (conscious). This is an empirical issue. The question could be asked, whether properties we usually associate with conscious percepts (I_3) are also present in iconic memory (I_2), or in other alleged cases of inaccessible experience (attentional blink, neglect, split brain – probably all I_2). For example, unconscious processing (I_1 in the neural sense) is typically about feature extraction, whereas in conscious perception (I_3) features are combined into objects, backgrounds, and so on (Lamme 2004). Is there perceptual binding in iconic memory (Landman et al. 2003)? Do indirect effects (such as learning) of I_2 states operate along the dimensions of isolated features or of coherent percepts?

Similarly, it could be asked what the critical neural differences are between I_1, I_2, and I_3 states. The first 100 msecs of visual processing is dominated by feedforward activation of the brain. Information sweeps from visual to frontal areas, not accompanied by conscious experience – that is, fully inaccessibly (I_1) (Lamme & Roelfsema 2000). Subsequently, recurrent processing is instantiated by horizontal and feedback connections. With time, localized (I_2) recurrent cores may grow into more global ones (I_3), depending on bottom-up and top-down selection mechanisms (Lamme 2003). Where does the critical neural dichotomy lie? Between feedforward and recurrent processing – that is, between I_1 and I_2/I_3, as Block and I would argue (Block 2005; Lamme 2003) – or between I_1/I_2 and I_3, as GWT advocates try to let you believe (Dehaene et al. 2006)? Before you choose, please consider that also in fully unconscious feedforward activation (I_1), there is activation of workspace neurons, as is shown by masked stimuli activating prefrontal cortex (Lau & Passingham 2007; Thompson & Schall 1999). In addition, there are important differences in the properties of feedforward versus feedback synapses. It is likely that feedforward activation is not mediating synaptic plasticity and learning, while recurrent processing (of whatever extent) does do so (Singer 1995). Third, recurrent processing between visual areas has been shown to mediate perceptual organization, binding, and figure-ground organization, in cases of inattention and the absence of report, as well, whereas feedforward processing is typically about feature extraction and categorization (Lamme 2004). Finally, recurrent processing is suppressed by anesthesia, whereas feedforward is not (Lamme et al. 1998).

To the neuroscientist, it therefore seems pretty straightforward to draw a line between feedforward processing (I_1) on the one hand, and recurrent processing (I_2/I_3) on the other. Of course, the extent of these recurrent interactions matters: when frontal or motor areas are involved, a report is possible, otherwise not. But that also applies to feedforward processing. Unconscious behavioral effects (like priming) are possible only when the feedforward sweep penetrates deeply into the sensorimotor cascade. The key feature "causing" phenomenality in I_3 states therefore seems to be the recurrency, not the activation of workspace (frontal) neurons. Occam's razor thus obliges us to group I_2 with I_3, not with I_1, and to attribute phenomenality to both I_3 and I_2. The neuroscience angle brings that out immediately, and much more convincingly (Lamme 2004). The jury can now go out and deliberate.

Can we equate iconic memory with visual awareness?

doi: 10.1017/S0140525X07002932

Rogier Landman[a] and Ilja G. Sligte[b]

[a]*McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139;* [b]*Department of Psychology, University of Amsterdam, 1018 WB Amsterdam, The Netherlands.*

landmanr@gmail.com
i.g.sligte@uva.nl

Abstract: Every time we look around we can see a rich and detailed world surrounding us. Nevertheless, the majority of visual information seems to slip out of our thoughts instantly. Can we still say that this fleeting percept of the entire world was a conscious percept in the first place, as Block proposes?

Attention enables human observers to report and to remember visual information. However, in our experiments, the formation of the memory trace that gives rise to the partial report benefit

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

takes place without focal attention being directed to the items while they are in view. After the items have disappeared and the cue appears, focal attention shifts to the cued item in memory and enables observers to make a large amount of information available for report. This suggests that much more information is processed beyond the scope of focal attention.

The properties of the iconic representation indicate that the items in the scene have undergone more extensive processing than one might expect given that no focal attention was directed at them. We have shown that a cue allows for recovery of multiple features of the same object, indicating that feature binding has taken place, and objects are at least processed up to the level of figure-ground segregation (Landman et al. 2003; 2004); characteristics that historically were only attributed to attentional – that is, reportable – stages of visual perception. In addition, iconic representations can last up to several seconds, and they are not a mere after-effect (Sligte et al. 2008). Also, several older studies have already indicated that iconic memory has a spatiotopic component, suggesting that it has undergone more processing than if it were strictly retinotopic (e.g., Breitmeyer et al. 1982; McRae et al. 1987; for a review, see Cowan 1995).

In daily vision, and in most experiments, when new information enters the visual cortex, the processing of that information gets priority over keeping information about stimuli that are no longer there. In the iconic memory experiment, we limit the entry of new information by showing a blank screen after the stimuli presentation. This allows us to test how much information about the stimuli is available in the visual system, by cueing one of them. If there is a distinction between phenomenal awareness and access awareness, then iconic storage can be seen as a window into the contents of phenomenal awareness. This is difficult to achieve in another way. If we test while the stimuli are still in view, subjects can just get information by looking, or switching attention, whereas if we test once new stimuli have already appeared, the processing of those new stimuli interferes with information we were interested in in the first place.

Iconic memory may be so vulnerable to interference that even the subjects' own response interferes with it. Many classic iconic memory experiments required a verbal response and identification of items. It could be that while the subject was reporting items, the report itself interfered with memory. Our experiments, however, merely required a "yes/no" button press, and change detection instead of identification of the items. Thus, the "report" requirement was much smaller, while we could still get an estimate of the amount of information available. That could be a reason why cues in our experiments are effective up to several seconds after stimulus offset, longer than in a classic iconic memory experiment. Another major difference is that in standard partial report designs, subjects are shown a brief sample display just once before reporting about them, whereas we employed a match-to-sample design. Thus, in our design all items are shown twice; once during encoding and once during report. It seems that memory in general (including other forms of memory, such as long term memory) does better on recognition than on recall.

In favor of Block's proposal, studies so far indicate that iconic memory does not have the key properties of an unconscious process, while it does have properties typical of a conscious process. Unconscious processes (such as masked representations) typically remain inaccessible even when attention is focused on the specific object of interest, whereas iconic memory is accessible when cued. Also, in unconscious processing, even basic features may not be processed to the level of perception. Recent evidence shows that unconscious color-priming effects are more dependent on physical stimulus properties than on perceptual properties (Breitmeyer et al. 2004). In iconic memory, however, not only have features been resolved, but also figure-ground segregation and feature binding have already taken place, properties that are more associated with conscious processes.

Additional ways to find out whether iconic memory has properties in common with conscious processes include neurophysiological studies. In neurophysiology, it is possible to distinguish between feed-forward processing, and recurrent processing (Lamme & Roelfsema 2000). By selectively disrupting recurrent processing (RP), but leaving feed-forward processing intact, it has been observed that visual awareness does not arise. This was shown by backward masking (Lamme et al. 2002), by applying transcranial magnetic stimulation to the primary visual cortex (Jolij & Lamme 2005; Pascual-Leone & Walsh 2001), and by inactivating higher visual areas (Hupe et al. 1998; Lamme et al. 1998). Even when there are sudden lapses in awareness, it is observed that RP is absent, whereas feed-forward processing is intact (Supèr et al. 2001b). Many scholars agree that RP is likely to be involved in conscious perception. Current and future experiments are necessary to address what the exact mechanism behind iconic memory is, in which areas of the brain it occurs, and whether RP is involved. However, if there is RP, depending on how widespread it is, views will still differ on whether it looks more like a conscious process or more like an unconscious process. Some argue that RP within the visual cortices is sufficient for conscious perception to arise (Block 2005; Lamme 2003; 2006), while others maintain that consciousness requires more widespread recurrency, including areas involved in cognitive access and control, such as the prefrontal cortex (Dehaene et al. 2006).

It is important to test whether RP is involved in iconic memory. Preliminary data provides some indirect evidence for this, by showing a cortical origin and a long-lasting, reverberating nature. Still, more direct evidence is required to determine whether RP takes place and whether iconic processing is necessary and sufficient for visual awareness to occur.

Broken telephone in the brain: The need for metacognitive measures

doi: 10.1017/S0140525X07002944

Hakwan Lau and Navindra Persaud

Department of Experimental Psychology, University of Oxford, OX1 3UD Oxford, United Kingdom.

hakwan@gmail.com **http://hakwan.googlepages.com**
nav.persaud@utoronto.ca **http://navpersaud.googlepages.com**

Abstract: The fact that early visual processing has a larger capacity than later visual processing can be explained without positing distinct systems for phenomenology and cognitive accessibility. While phenomenology may overflow forced-choice reports, the later can also overestimate the former, as in the case of blindsight. Metacognitive measures of awareness offer a way to address the "methodological puzzle" of consciousness research.

Block claims that sometimes we see more than we can report because the neural system for phenomenology "overflows" the system for accessibility. He makes the additional claim that this implies there are distinct neural mechanisms for phenomenology and cognitive accessibility. We argue that the difference in capacity between phenomenology and accessibility can be explained by noise amplification without any need to posit distinct systems. We explain why we think Block's approach is unable to build upon empirical findings, and suggest that metacognitive approaches will be more fruitful.

As a message passes down the line in the "broken telephone" or "Chinese whisper" game, it becomes garbled and some of its elements are completely lost. That is, the quality of information tends to deteriorate. Noise propagation and amplification also limit late sensory processing in the brain. This is why early forms of vision that are brief and iconic have larger capacities than later verbal reports which require deeper information processing. Simply put, the retina has more visual information for a simple visual perceptual event than the motor cortex. Hence, a

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

difference in capacity is consistent with a single stream of serial processing and does not imply distinct processing systems, as Block claims. Note that our argument does not apply to situations when there are actually two streams of information, as in the dorsal-ventral distinction in visual processes (Goodale et al. 1991). In that case, the two systems are largely independent of each other. In Block's case though, the later cognitive access certainly depends on and receives its major inputs from the earlier brief processing, and for this reason the capacity difference is trivial.

Although we disagree with Block over the explanation of the differing capacities of phenomenology and cognitive accessibility, we agree that forced-choice reports can fail to capture what feels to be seen, especially when there is a lot going on in the visual presentation. This leads Block to propose "a neural mechanism by which phenomenology can overflow cognitive accessibility" (sect. 14, para. 3). But "overflow" is just one example of the failure of forced-choice reports. There are also cases in which forced-choice reports capture more than what is consciously seen. People with V1 lesions claim not to see anything in their affected visual field and yet make accurate visual discriminations; that is, they have blindsight (Weiskrantz 1986). There are other cases in which forced-choice reports made with different modalities (e.g., manual button press, eye blinks, verbal reports) yield inconsistent measures of phenomenology given the same stimulus (Marcel 1993). There are yet other cases in which a stimulus can cause people to make a forced-choice response that they do not want to make (Debner & Jacoby 1994; Persaud & McLeod 2007). So forced-choice reports are not ideal for measuring phenomenology. But this does not mean that we must associate phenomenology with a neural system that has a different processing capacity than that reflected by normal forced-choice reports. Nor does it mean that the capacity reflected by forced-choice reports under optimal cueing conditions (as in Sperling-style experiments) is the capacity for phenomenology. We never know, because forced-choice reports sometimes capture too much, sometimes capture too little, sometimes are inconsistent, and sometimes capture irrelevant information. We suggest that we must explore alternative measures, as it is vital to find reliable and valid ways of measuring phenomenology behaviourally before attempting to map it to a specific brain mechanism. We have been doing just this by employing metacognitive measures (Lau & Passingham 2006; Persaud et al. 2007); that is, we collect subjective reports, or judgements of performance, in addition to forced-choice reports regarding the stimuli.

Part of our motivation for using metacognitive measures is demonstrated by how Block's argument fails to find empirical support where he claims it does. Block claims that recurrent processing (feedback loops; for motion that is V1 → V5 → V1) within the visual cortices may support phenomenology. Given the above argument about capacities, it is clear that any processing stage prior to the stage that supports normal reportability would have a capacity large enough to "overflow" cognitive accesibililty, and thus be a good candidate for the supposed phenomenology. The retina, for instance, has all the visual information needed to support what is likely to be seen but not reported. Of course, Block does not think that the retina is a candidate. Presumably the reason is that the retina is not necessary for phenomenal vision: Electrical stimulation of the primary visual cortex can cause visual phenomenology without the retinal involvement. So being necessary for phenomenology is an important criterion. But pure feedforward processing (i.e., V1 → V5, without the feedback for the case of motion) may fit this criterion as well, and, if one follows the above argument about capacities and inheritance of noise, the information capacity of this processing would certainly overflow cognitive accessibility.

Block attempts to support his feedback hypothesis by pointing out that disrupting feedback processing is correlated with a lack of visual consciousness (Pascual-Leone & Walsh 2001). But if Block's argument that there can be stimuli which a person can see but not report is right, how do we know that in these cases there is a lack of phenomenology (and not just a lack of cognitive accessibility)? Block's argument backfires: If we allow for phenomenology without access, we would not be able to know when people do not see visual stimuli. For example, when magnetic fields disrupt feedback processing and people report not seeing stimuli they would otherwise see, how can we know that people do not actually see the stimuli? How can we know that feedback processing within the visual cortices does not just reflect cognitive access?

We believe that using alternative measures of phenomenology, such as metacognitive measures, may fill this gap. Although they may be imperfect, metacognitive measures are the best available method for determining when a person is aware of a stimulus. It is only after awareness can be properly measured that the neural substrates of consciousness can be found. Thus, metacognitive measures avoid the circularity inherent in Block's approach – that is, the very circularity in the "methodological puzzle of consciousness research" that Block attempts to address.

Two kinds of access

doi: 10.1017/S0140525X07002956

Joseph Levine

Department of Philosophy, University of Massachusetts, Amherst, MA 01003-9269.

jle@philos.umass.edu

http://www.umass.edu/philosophy/faculty/levine.htm

Abstract: I explore the implications of recognizing two forms of access that might be constitutively related to phenomenal consciousness. I argue, in support of Block, that we don't have good reason to think that the link to reporting mechanisms is the kind of access that distinguishes an experience from a mere state.

Block's original distinction between phenomenal consciousness and access consciousness has provoked a lot of critical comment in the last decade or so, my own included. I see two distinct sources of uneasiness with the distinction between phenomenal and access consciousness, and while these two sources have not usually been distinguished, some of Block's remarks in the present target article seem to acknowledge the need to do so. On the one hand, coming principally from a functionalist perspective, many philosophers and psychologists believe that there has to be a constitutive connection between what we are conscious of and what we can report. The idea is that consciousness is somehow reducible to this kind of access, and only if we can so reduce consciousness will it be amenable to scientific investigation. It is this view that is Block's target here.

On the other hand, some, like myself, have been uneasy with Block's distinction because the idea of phenomenal consciousness totally divorced from any access by the subject does not really seem like any kind of consciousness at all. As Block notes here, we have complained that the very phrase that serves to canonically express the notion of the phenomenal – "what it's like for *x* to . . ." – explicitly refers to the phenomenal state in question being "for" the subject. The way I would put it now is: Phenomenal states/properties are not merely instantiated in the subject, but are experienced by the subject. Experience is more than mere instantiation, and part of what that "more" involves is some kind of access.

So one possibility, suggested by Block's discussion here, is that for a state to be phenomenally conscious it must be accessible to the subject in some sense, but not necessarily to the mechanisms responsible for report. Block supports the view by appeal to the Sperling and the Landman et al. experiments, together with the associated neurological data, which, he argues, show that

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

phenomenal consciousness overflows the "working space" that is report-accessible. Is this appeal persuasive?

One thing to notice right away is that Block's own argument actually relies on building some notion of access into the notion of what is phenomenally conscious, which reinforces the argument mentioned above. After all, in the Sperling experiment, the evidence that is supposed to show that we are phenomenally conscious of more than we can report is that there is information concerning the identity of the letters in the array that we are aware of, though we lose it when reporting mechanisms are engaged. The evidence that it was there in phenomenal consciousness is primarily that we say we saw all the letters. My point is not about the saying, but about the fact that what we report is that we did have a kind of access to this information; and it is because the information is available phenomenally that, when prompted appropriately, we can report a portion of it.

But now one might object that we can't really tell from the experimental data that the identities of the letters were phenomenally conscious. One might claim that, instead, one was phenomenally aware of more generic information, and that the specific information manifested in the partial report was stored unconsciously. Block explicitly addresses the objection that perhaps the information persists retinally only, but one might claim, in response to his demonstration that it goes higher up, that it is still a matter of unconscious, or sub-personal storage. How can one ever rule that out, after all?

Of course, Block is not claiming to have a knock-down refutation of his opponents' interpretation of the data, just a more plausible version. I find myself largely sympathetic to his position, and so wonder about the source of the resistance to it. Why could it not be pretty much as he says? It seems to me that fueling this resistance is a nagging sense that access just *has* to be constitutive of conscious experience, that we just cannot understand what it would be to be conscious without it. But given the concession already made that some access is involved in phenomenal consciousness, why not go along with Block's interpretation of the Sperling and other data? Why insist on report-access in particular?

I cannot speak for Block's opponents, but I suspect that reasoning along the following lines underlies a good part of the resistance to his position- First, there is the relatively conceptual point that consciousness is constitutively related to subjective access – my aforementioned point about experience being more than mere instantiation. Once the notion of the subject – the "for whom" it is like what it's like – enters the picture, we then ask who or what this subject is. Well, the natural thought is that the subject is the person, and the person is the entity that both plays the highest executive role in deliberation and planning action, and reports to others (and to itself). Hence, the idea of a subjective access that is divorced from reporting mechanisms seems hard to swallow, as there doesn't seem to be a suitable candidate for the subject whose access is involved in the bit of phenomenal consciousness at issue.

If I am right about the source of the resistance, then I think there are two lines of reply. First, it might be that subjective experience does not entail the existence of a subject, at least not anything like what we normally take a subject to be. While what is consciously attended to might involve access by the sort of high-level executive we associate with the personal level, and we naturally think of as the subject, the person, it seems plausible that there are centers of experience more loosely connected in the mind and that do not involve an entity we would call a full-fledged subject. I take Block's argument here to show that this is plausible, given the data so far. The argument can be contested, of course, but I wonder what, besides a commitment to the conceptual binding of experience to a substantial subject, motivates the search for these alternative – and to my ear – strained, interpretations of the data.

Finally, one might retain the idea of a substantial subject as a necessary component of experience, and even retain the idea that the subject is that which reports. As Block noted in discussing Dehaene and Naccache's division of the global workspace into I_2 and I_3, the only dispute here is whether what's phenomenally conscious has to be in I_3. Block isn't denying that it must at least be in I_2. But if I_2 is understood as that pool of information from which I_3 draws, then what counts as the subject can be identified with the union of I_2 and I_3. The point is that whatever bottleneck exists due to the processing that gets an item from I_2 into I_3 shouldn't be taken to restrict what we count as the full-fledged subject of experience. Of course it might have been that way, and in the end it is an empirical question, as Block insists. But why think, as so many who insist on the constitutive connection between consciousness and reportability seem to, that it just *has* to be that way?

Phenomenality without access?

doi: 10.1017/S0140525X07002968

William G. Lycan

Department of Philosophy, University of North Carolina, Chapel Hill, NC 27599-3125.

ujanel@email.unc.edu **http://www.unc.edu/~ujanel**

Abstract: Block holds that there can be "phenomenology," "awareness," and even awareness of the phenomenology, without cognitive access by the subject. The subject may have an experience and be aware of the experience, yet neither notice it nor attend to it. How that is possible is far from clear. I invite Block to explain this very fine distinction.

I firmly endorse the idea that *qualitative* character overflows cognitive accessibility, because I believe we are often in sensory states that present qualitative features of which we are unaware. But Block makes a stronger claim based on a finer distinction. That distinction is unclear to me, and I invite him to explain it. He holds that there can be "phenomenology" and "awareness" without cognitive access by the subject.

First, I distinguish Block's distinction from the more familiar one between unconscious and conscious detection of a sensory quality. As noted, we often sense color or shape or sound or another environmental feature without being at all aware of doing so. But such nonconscious detection does not count as *phenomenal* for Block, since (a) as he uses that term, "when one has a phenomenally conscious experience, one is in some way aware of having it" (sect. 6, para. 1), and (b) for a mental state to be phenomenal, in Block's usual use of the term, there must be "something it is like" to be in it.

Yet according to him, such Awareness (his capital "A") does not require cognitive accessibility. Why not? At least in part, because a mental state that includes it need not be one that "the subject notices or attends to or perceives or thinks about" (sect. 6, para. 3).

But awareness of any sort is intentional, especially when "of" is explicitly appended. If one is aware *of* having an experience, in particular, the experience is the intentional object of the awareness. And I assume, without fear of disagreement from Block or any of his readers, that intentionality is representation. Therefore, a phenomenally conscious experience in his sense is one that is represented by one of the subject's own psychological states. In light of Block's allegiance to "same order" accounts of Awareness (sect. 6), I do not say that the experience is represented by *another* of the subject's states; Block's view is that the experience "consists in part in an awareness of itself" (sect. 6, para. 2), perhaps as if it were a conjunction along the lines of "There is some cheese over there and it is this very experience that tells me so."

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

But that "same order"–ness does not per se distinguish Awareness from noticing or (minimally) attending. As Block says, the awareness of itself is (only) *part* of the experience. So far as has been shown, it is a separable part, as is the foregoing cheese sentence's second conjunct: there is the first-order component of the experience and another part that represents the first-order component, as in Gennaro (1996). So how does that latter representational part differ from noticing or attending?[1]

The most obvious guess would be in terms of passive versus active introspection. "Higher order" theorist Armstrong (1981, p. 63) distinguishes between mere "reflex" introspective awareness and "scrutinizing" or actively exploratory introspection. The former is merely a "watching brief" and not really worth calling "introspection," while the latter is "introspection proper." Though Armstrong does not say so, I daresay this is a matter of a low and routine level of attention versus a high and active level of attention.

But that cannot be what Block means either. If the reflexive part of the experience were a matter of passive, routine "watching" and/or peripheral, low-level attention, it would still be watching and attention, which are what he is denying.[2]

Of course there is representation in the brain that does not constitute either noticing or attending. But the awareness Block is talking about is person-level; it is the whole subject who is supposed to be aware of her/his own experience.

What, then, is Awareness, and how does it differ from the various forms and degrees of cognitive accessibility?

NOTES

1. Kriegel (2005) faults Gennaro for so treating the first-order component and the self-referential part as separate and distinct; he maintains that the self-representation is somehow more "intrinsic" to the original state itself. But this is obscure and not explained.

2. Nor does help come from Block's (1995b) pneumatic drill example, designed to illustrate "phenomenal consciousness" absent "access-consciousness": "You were aware of the [drill] noise all along, but only at noon . . .[do you become access-]*consciously* aware of it" (p. 234; italics in the original). I can parse that in any of three ways: (1) You were detecting the noise all along, but only at noon do you become aware of the noise; (2) you were dimly aware of the noise all along, but only at noon do you become focally aware of it; (3) you were aware of the noise all along, but only at noon do you become aware of that awareness itself. Each of those makes sense, but I am pretty sure that none of them is what Block intended.

The measurement problem in consciousness research

doi: 10.1017/S0140525X0700297X

Rafael Malach

Department of Neurobiology, Weizmann Institute of Science, Rehovot, Israel 76100.

rafi.malach@weizmann.ac.il

http://www.weizmann.ac.il/neurobiology/labs/malach

Abstract: States of sensory absorption may offer a means to disentangle perception from report. Interestingly, such states lead to an *antagonistic* relationship between perceptual and cognitive-access networks, suggesting that perceptual awareness does not depend on a read-out by high order cognitive-access mechanisms. Rather, it may emerge internally, through a cooperative coding dynamics, whereby each neuron simultaneously represents and reads-out the perceptual awareness state.

Block forcefully illustrates a challenging methodological difficulty inherent in consciousness research: the measurement problem – that is, the fact that any exploration of consciousness depends on some kind of report, either external or through introspection. Consequently, one is invariably faced with the difficult task of disentangling the neuronal mechanisms associated with such reports from those underlying the phenomenal experience (e.g., sensory perception) itself. Human functional magnetic resonance imaging (fMRI), which can provide a highly detailed mapping of the conscious human brain, nevertheless suffers particularly severely from this methodological confound due to its sluggish temporal dynamics.

How then can one disentangle these closely related processes? Block provides an impressive survey of sophisticated experiments suggesting that reportability can be distinguished, and in fact is not necessary for phenomenal experience. Here I will present an alternative approach which nevertheless agrees with the notions proposed by Block – both point to the feasibility of phenomenal perception without a reporting perceiver.

The idea is quite straightforward: If the reporting/introspection stage can somehow be sufficiently segregated in time from the perceptual stage, one may be able to study in isolation the brain areas engaged during perception without confounding them with those involved in reporting and introspection. In such an experimental paradigm, the report/introspection of the percept is obtained only at a later stage, through recollection.

It could be argued that such a clean temporal separation of perception from reporting is simply not feasible. However, there are many instances in which it seems that perception occurs without any overt report or introspection. A striking example is the condition of sensory *absorption*, in which engagement with the perceptual stimuli is so intense that one gets the strong sense of "losing oneself in the act." Another, more common type of experience may happen when watching a highly engaging movie – again, one is clearly not in the business of reporting or self-introspection during such states.

Of course, there is no a priori reason to assume that such intuitive impressions of self-loss indeed reflect a true neuronal dissociation of perception from any self-related processes; for example, it could be that implicit activation of cognitive-access areas may occur even during highly absorbing moments. In particular, it has been suggested that neural processes underlying an implicit first-person perspective may be an essential element of any conscious awareness state. Here is where fMRI research may prove useful, because it allows the mapping of any neuronal activity, be it conscious as well as implicit or subconscious during such engaging moments.

So what does brain imaging during absorbing perceptual moments reveal? In an fMRI study of brain activation in subjects watching a highly engaging movie (Hasson et al. 2004), the results revealed a robust and wide-spread activation in the back (i.e., sensory part) of the cerebral cortex, in a system of areas we termed the "Extrinsic" system (i.e., cortical regions oriented towards the external environment). In contrast, the front part of the brain remained relatively unresponsive. Indeed, even in the back part we found several relatively unresponsive islands. This entire set of nonresponsive areas constituted a coherent system (largely overlapping with the default mode network; Raichle et al. 2001)) which we termed the "Intrinsic" system because of its complementary nature to the sensory-driven Extrinsic system. We hypothesized that the intrinsic/extrinsic divide may reflect a fundamental functional organization of the human cortex (Golland et al. 2007).

Our research, as well as that of many others, have indicated that the Intrinsic system deals with internally oriented functions – precisely the network one would assume is the most likely candidate for mediating cognitive access, introspection, and reportability, as well as the "first person perspective" (Baars et al. 2003). Critically though, the Intrinsic network in fact shows a strong *reduction* in activity precisely during moments of intense perceptual processing (Goldberg et al. 2006; and see Fig. 1 further on here). Thus, not only do we fail to find fMRI evidence for a synergistic activation of sensory representations in the back of the brain with self-related networks in the front,

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

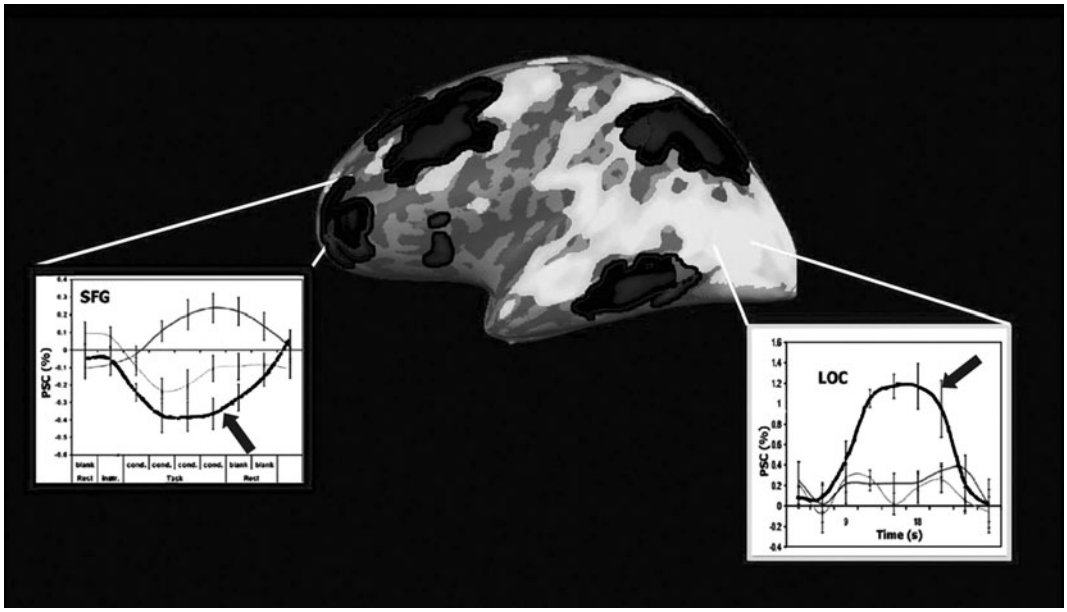


Figure 1 (Malach). Intrinsic and extrinsic systems. Antagonistic relationship between the extrinsic and intrinsic systems (light and dark patches, respectively) during an intense visual recognitions task. Arrows point to the high fMRI activation in the LOC region of the extrinsic system (right inset), as opposed to inhibition of activity in the SFG, a self-related part of the intrinsic system (left inset). Lateral view of an "inflated" left hemisphere. Back is to the right. [Modified from Goldberg et al. (2006) and Golland et al. (2007).]

but intense sensory engagement actually appears to shut off these cognitive access networks!

In summary, fMRI data during perceptual engagement suggests that perceptual awareness can emerge through internal activity in sensory systems, without a need to be "read out" by higher-order cognitive access mechanisms.

Finally, on a cautionary note, it should be emphasized that the present conclusions should be tempered by the methodological limitations of fMRI. Hence, substantial neuronal activity may go undetected by the fMRI method if the neuronal representations are too small (Avidan et al. 2002), or the signals are too rapid to affect the sluggish hemodynamic fMRI response. Furthermore, cortical regions which do not modulate their activity levels during sensory perception may go undetected using our standard fMRI methodology.

On the other hand, it is tempting to consider the more speculative implication of these recent findings: the notion that phenomenal experience may emerge through internal processing within sensory representations proper has far reaching consequences for neuronal theories of consciousness. To see why, let us consider, first, the concept of population coding, which is a widely accepted notion of sensory representations. In such coding schemes, the combined pattern of activity in a group of neurons (the population vector or state) represents a sensory percept, say a yellow color (see Fig. 1) generated by virtue of the fact that a "green" and a "red" neurons (R and G in Fig. 2) are active while the "blue" neuron (B) is not. Importantly, in conventional models, this network state is then read out by a higher-order station. Such coding allows a huge combinatorial power (e.g., Levy et al. 2004).

However, note that here we consider the possibility that phenomenal experience emerges within the bounds of the sensory representations themselves, without assuming a hierarchical flow into a high order read-out area. Instead, I would like to propose that the percept is coded cooperatively by the sensory network itself (see Fig. 2). Therefore, in such a scheme the phenomenal experience emerges when all relevant neurons in a network are informed about their own population state. Note that in such a cooperative coding model (somewhat

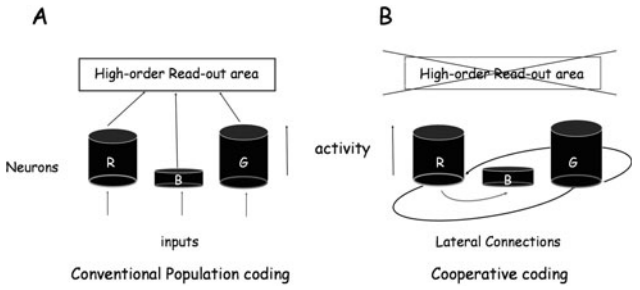


Figure 2 (Malach). Two alternative schemes for perceptual representations. In conventional population coding (A) the neuronal state is read out by high-order areas. In the "cooperative coding" scheme there is no hierarchical processing. The neuronal state is both generated and read out by the same neurons through their lateral connectivity. Here I propose that the latter dynamics leads to phemonenal experience.

analogous to a point attractor dynamics) there is no hierarchical processing – each active neuron simultaneously serves the roles of reading out and representing the perceptual state. Intriguingly, this single principle seems to successfully account for a large body of recent experimental data: for example, the high firing rates, relatively long durations, and dense local connectivity which appear to be critical for the emergence of conscious perceptual states.

To summarize, the field of consciousness research appears to be in the midst of an exciting period where the experimental jury is still out regarding such fundamental issues as the minimal spread of neuronal activity that is sufficient to elicit a conscious percept. However, in complete agreement with Block, I believe that these issues are experimentally tractable and will certainly lead to great advances in the construction of testable neuronal theories of conscious awareness.

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

Dodging the explanatory gap – or bridging it

doi: 10.1017/S0140525X07002981

Drew McDermott

Computer Science Department, Yale University, New Haven, CT 06520-8285.
drew.mcdermott@yale.edu
http://www.cs.yale.edu/homes/dvm

Abstract: Assuming our understanding of the brain continues to advance, we will at some point have a computational theory of how access consciousness works. Block's supposed additional kind of consciousness will not appear in this theory, and continued belief in it will be difficult to sustain. Appeals to "what it's like" to experience such-and-such will carry little weight when we cannot locate a subject for whom it might be like something.

Block says he wants to sidestep the "explanatory gap," but I'm afraid that it is too wide to walk around – you have to find a way to get across it.

Let us suppose we had a completely worked-out explanation of how information flows between the various chunks of brain discussed in the target article. Let's stipulate, with the author, that the explanations are roughly at the level of neural networks as described by Koch (2004). They will therefore be *computational*, in the sense that they would leave little room for distinctions between the actual brain and any device that passed equivalent coded messages between nodes standing for assemblies of neurons.

It seems likely that we will possess such explanations in the foreseeable future. At that point we will know how "access" in the author's sense will happen; that is, let us suppose, how information is distributed to a global workplace in Baars' (1997) framework. Will we also know how "phenomenology" happens? (I mean "phenomenology" in the sense Block uses it in sentences such as, "Sperling's clever idea was to test . . . whether the phenomenology persists after the stimulus was turned off by. . ."; target article, sect. 9, para. 2.) Well, no, we won't. The computational explanation of how information flows in order to enable subjects to report a row of Sperling's (1960) array after hearing a tone will of course not refer to anything like phenomenology, but only to neural structures playing the role of buffers and the like. This is what Chalmers (1996) calls the Paradox of Phenomenal Judgment, which is just another angle on the Explanatory Gap.

Block would like it to be the case that phenomena play a role in our psychological abilities. To extract a few exemplary phrases from section 9:

> [S]ubjects are able to deploy working memory so as to access only half the rectangles despite the fact that . . . subjects' reported phenomenology is of seeing all or almost all of the rectangles. . . . [Suppose] the subjects are continuing to maintain a visual representation of the whole array – as subjects say they are doing . . . (sect. 9, para. 6)

and

> The subject has persisting experiences as of more specific shapes than can be brought under the concepts required to report or compare those specific shapes with others. (sect. 9, para. 10)

Notice how terms like "working memory," "phenomenology," "visual representation," "experiences," and "concepts" are used as though they all belong at roughly the same level in the explanatory hierarchy; but, in fact, in our envisioned computational explanation the "phenomenology" and "experiences" will drop out, as they seemingly always do.

Our intuition that experiences are somehow inspected in making phenomenal judgments is,[1] as I argue at greater length in McDermott (2001), exactly backwards. The inspections are computational events that must *precede* the experiences, causally and temporally. If an experience is reported, or accessed in some weaker sense, then we can explain it by reference to how the brain thinks about itself.[2] But if an experience happens without being accessible, then it becomes a phantom.

Let me draw an analogy here. If we continue to suppose that cognitive science finds a satisfying computational explanation of visual processing and access consciousness, the unaccessed experiences will be in a position analogous to unobserved electrons in a two-slit experiment demonstrating interference of their wave functions. Some physicists, such as David Bohm (Bohm & Hiley 1993), found it hard to believe that an electron could shed its particlehood for all except brief flashes of time, and devised theories in which electrons always remained particles, whose positions ingeniously served as indices to wave functions. Few in the mainstream physics community found this theory attractive, for some technical reasons, but also for an easily grasped one: The "positions" in Bohm's theory, being in principle completely unobservable, were obviously vestigial ornamentation on a theory that was otherwise an inelegant variant of ordinary (nonrelativistic) quantum mechanics.

It seems clear that if we get a cognitive-scientific (i.e., a neuroscientific or a cognitive-psychological) explanation of access to experience, we are going to go through an intuition change at least as violent as that brought about by quantum physics. I think it is inevitable that our understanding of consciousness will change as radically as our understanding of, say, life, has since the seventeenth century. Will this be a new "definition" of consciousness? Block supposes (sect. 3) that we can always rephrase questions about phenomenology using the popular but at best vacuous formulation asking "what it is like" to have a particular experience. However, one has to ask to *whom* it might be "like something" for the fusiform face area of brain-lesion patient G.K. to light up when it is not like anything to G.K. himself?

I am actually in sympathy with the author's stance on methodology. I think the Occamish arguments he uses to justify his conclusions can be valid. I just don't think Block's methodological razor is a match for the scientific buzz saw that is cutting through this area of science.

NOTES

1. Note the unspoken premise in the quote from the target article above that the phenomenology of eight somethings must somehow be eight – what? Phenomenologies of some kind, I guess – only a few of which can be snatched before evanescing.

2. In other words, we can place it within the framework of a "higher-order thought" theory of phenomenal consciousness, although not the caricature of that hypothesis described by the target article in section 6; of course the theory does not posit that every conscious thought is conscious by virtue of another *conscious* thought targeted at it. If it did, the obvious infinite regress would make the theory unattractive to just about anyone aware of post-Renaissance developments in philosophy of mind.

Reportability and illusions of phenomenality in the light of the global neuronal workspace model

doi: 10.1017/S0140525X07002993

Lionel Naccache[a,b*] and Stanislas Dehaene[b]

[a]*Fédération de Neurophysiologie Clinique, Fédération de Neurologie, Hôpital de la Pitié-Salpêtrière, Paris, France;* [b]*INSERM Cognitive Neuro-Imaging Unit, CEA/SAC/DSV/DRM Neurospin Center, Gif/Yvette Cedex, France.*
lionel.naccache@normalesup.org
dehaene@cea.fr

Abstract: Can we ever experimentally disentangle phenomenal consciousness from the cognitive accessibility inherent to conscious reports? In this commentary, we suggest that (1) Block's notion of phenomenal consciousness remains intractably entangled with the need

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

to obtain subjective reports about it; and (2) many experimental paradigms suggest that the intuitive notion of a rich but non-reportable phenomenal world is, to a large extent illusory – in a sense that requires clarification.

In a very stimulating target article, Ned Block places phenomenal consciousness at the core of his conceptual framework for consciousness (Block 1995b; 2001; 2005). In the past, we have proposed that access of information to a prefronto-parietal global neuronal workspace, capable of broadcasting that information to many distant cortical areas and therefore making it available for verbal or nonverbal report, is the essential ingredient of a theory of conscious perception (Dehaene & Naccache 2001). For Block, however, conscious reports fail to capture the whole conscious experience (and this incompleteness is not imputable solely to spurious factors such as reporting biases or errors). Block defends the existence of a phenomenal consciousness irreducible to conscious reportability. The key problem that we have with his thesis is its lack of empirical testability. Can we ever experimentally disentangle phenomenal consciousness from the cognitive accessibility inherent to conscious reports? In this commentary, we suggest that (1) Block's notion of phenomenal consciousness remains intractably entangled with the need to obtain subjective reports about it; and (2) many experimental paradigms suggest that the intuitive notion of a rich but non-reportable phenomenal world is, to a large extent illusory – in a sense that requires clarification.

1. Reports, conscious access, and phenomenology. If one wants to define phenomenal consciousness differently than conscious reportability, then one should resist the temptation to make use of subjects' reports to credit the existence of phenomenal consciousness. If the only support for the existence of phenomenal consciousness comes from conscious reports, then we find no reason to accept a major distinction between these two concepts. Yet Block's argument makes recurrent use of subjects' reports to advocate the existence of phenomenal consciousness. For instance, Block states:

> When one has a phenomenally conscious experience, one is in some way aware of having it. (sect. 6, para. 1)

In this case, aren't we also able to report it? Later, when discussing Sperling's seminal "iconic memory" experiments, Block writes:

> I am taking what subjects say at face value. (sect. 9, para. 6)

and

> How do we know we can believe subjects' reports to the effect that they experience all or almost all of the objects in the Sperling and the Landman et al. experiments? (sect. 9, para. 15)

In these sentences, Block relies heavily on subjective reports. He suggests that they provide face-value support for a rich phenomenal experience, supposedly going way beyond the few details that can actually be reported. Yet this seems to prove our point: Availability for report remains the sole criterion for conscious perception. We are still waiting for a clear, empirically usable definition of "phenomenality" or "qualia" that goes beyond what can be done based on subjects' reports alone.

2. What is a report? It is, however, essential to better define what we mean by "report." A report is not a "cut and paste" copy of a visual scene, but rather a conscious comment on an inner mental representation. This representation can originate from perceptual systems at multiple levels, but ultimately it results from their redescription by evaluative and interpretative systems. At any given moment, it provides only a partial and possibly biased description of the perceptual scene. In our view, the fact that, on any given trial, subjects cannot report the whole scene is therefore non-essential – what is essential is that any consciously visible item is accessible for report.

Block's comments on split-brain patients, locked-in patients and aphasic patients, also prompt us to remind him that conscious reports can be nonverbal or even become entirely covert, due to motor system impairments (Gazzaniga et al. 1977; Laureys et al. 2005). The absence of overt verbal report in patients is not diagnostic of the absence of conscious access and of internal form of "self-report" (e.g., internal speech). Using neurophysiological tools such as functional magnetic resonance imaging (fMRI) and event-related potentials (ERPs), indirect conscious reports can be obtained from these patients. They can and should be further complemented by looking for other psychological properties which are thought to be exclusive to conscious processing, such as active maintenance of mental representations (Owen et al. 2006) or dynamic regulation of executive control. In this broader definition, the reportability criterion is far from being a "behavioristic ladder" (see target article, sect. 4, para. 5) – rather, it provides an excellent test of whether or not an information is conscious accessible (Naccache 2006b).

3. Sperling's experiments. Block interprets Sperling's iconic memory experiments (Sperling 1960) and their recent extension by Landman et al. (2003) as a clear proof of a rich though non-reportable phenomenal world. We disagree, and propose an alternative interpretation of these experiments in the global neuronal workspace framework.

What happens in the brain when a large array of letters is briefly flashed, then removed? Since no mask is present, there is no reason to doubt the subject's report that he or she is aware of the array. The conscious content presumably stems from both a dorsal map of approximate letter locations (perhaps in area lateral intraparietal [LIP]) and a ventral representation of the letters (perhaps in area V4; Dehaene et al. 2004). Both dorsal and ventral representations must be jointly accessed by prefrontal regions when subjects report that they see a spatially extended array of letters.

Sperling's experiments clearly demonstrate that all individual letters remain represented for ~1 second within a posterior, presumably retinotopic location-specific buffer (iconic memory). However, other studies indicate that when attention is distributed, temporal and prefrontal regions act as capacity-limited filters and cannot represent more than a few objects (Kastner et al. 1998). When subjects report seeing "all the letters," we suggest that they distribute their attention globally over the array, and thus are only able to determine its approximate numerosity and "letterhood"; our model predicts that only this approximate content, not the detailed letter identities, accesses a fronto-parietal global neuronal workspace. When subjects report individual letters, they change to a focused attention mode that allows them to enhance one location at a time, to let its specific content access the global workspace, and, therefore, to name it. Our model predicts that in this case, a local subregion of letter-sensitive retinotopic cortex (e.g., V4) becomes temporarily synchronized with anterior cortices and transmits a reportable visual content to prefrontal cortex (for related evidence, see Haynes et al. 2005; Sergent 2005).

The model we just sketched, although in need of further specification, seems capable of accounting for the subjects' reports within the global neuronal workspace without appealing to Block's hypothesis of non-reportable phenomenal states. The mismatch between the report of seeing the whole array and the reduced capacity to report the individual letters therefore fails to provide univocal support in favor of Block's phenomenality-without-report hypothesis (as noted above, it does not even speak to this issue, since it is a mismatch between two *reports*).

4. The illusion of seeing. Part of the confusion surrounding the Sperling paradigm may stem from our description of the report of seeing the whole array as an "illusion of seeing" (Dehaene et al. 2006). The term "illusion" applies imperfectly to Sperling's paradigm because, after all, the subjects' introspection turns out to be veridical: they *can* see the whole array, and they *can*, when probed, report any of its letters (though not all of them at once). However, the term "illusion" emphasizes that, until information is accessed by the fronto-parietal workspace, it is

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

illusory to think that subjects *necessarily* have veridical knowledge of it. Indeed, our model predicts that, if a single letter of the Sperling array was replaced by another letter or even a digit prior to the focused attention stage, subjects might not notice it and still maintain that they "see all the letters" – a clear illusion. This substitution paradigm would lead to many testable predictions. For instance, a subpart of area V4 should have veridical information about the symbol's identity, which could be decoded by fMRI (see Haynes & Rees 2006; Williams et al. 2007) – but this spot should be (temporarily) functionally disconnected from frontal decision areas, and its information should not be used in subject's reports.

Many other paradigms and neuropsychological syndromes (Naccache 2006a) that are not discussed by Block indicate that reports of a rich phenomenality cannot be taken at face value (though we agree with Block that they still have to be explained in all of their details). In the "moving window" paradigm, for instance, where a computerized display is changed in synchrony with eye movements, viewers claim that they see a normal page of text even when all parafoveal information is replaced by strings of X's (Rayner & Bertera 1979). Similarly, we all have the illusion of seeing a world in full color although color-sensitive cones are absent in the periphery of our retina. Such illusions suggest to us that building a theory of consciousness based on intuitions of phenomenality without reportability may be building on sand.

Phenomenal consciousness lite: No thanks!

doi: 10.1017/S0140525X07003007

J. Kevin O'Regan[a] and Erik Myin[b]

[a]*Laboratoire Psychologie de la Perception, CNRS – Université Paris Descartes, Centre Biomédical des Saints Pères, 75270 Paris cedex 06, France;* [b]*Centre for Philosophical Psychology, Department of Philosophy, Universiteit Antwerpen, 2000 Antwerpen, Belgium.*

jkevin.oregan@gmail.com **http://nivea.psycho.univ-paris5.fr**
Erik.Myin@ua.ac.be **http://www.ua.ac.be/erik.myin**

Abstract: The target article appeals to recent empirical data to support the idea that there is more to phenomenality than is available to access consciousness. However, this claim is based on an unwarranted assumption, namely, that some kind of cortical processing must be phenomenal. The article also considerably weakens Block's original distinction between a truly nonfunctional phenomenal consciousness and a functional access consciousness. The new form of phenomenal consciousness seems to be a poor-man's cognitive access.

A central piece of the argument presented in the target article is Ned Block's claim that "phenomenal consciousness overflows cognitive accessibility" (target article Abstract). Block's main motivation for this claim is his intuition, shared by most of us, that we see much more than we can report. The feeling is particularly clear when an image is flashed before us, and we feel we have seen "everything in it," yet generally will be unable to report more than 3 to 5 elements from the scene (the classic Sperling "whole report" paradigm).

The trouble is that an explanation for this feeling of being able to see more than we report might lie in the fact that the early stages of the visual system possess a certain degree of neural persistence. This neural persistence or "icon" might not itself be phenomenally conscious. Instead, it might, for a short while, be available for access by higher cortical mechanisms. Our impression of seeing "everything" might therefore derive not from actual phenomenality of the icon, but from the immediate availability for access of information in this persistent subcortical icon.

This is where Block musters new evidence from the recent experiments of Landman et al. (2003) and Sligte et al (2008). These experiments, contrary to the original Sperling experiments (Sperling 1960), show that the neural persistence that seems to be involved in giving us the impression of seeing everything can sometimes be of considerably longer duration than previously measured. Information about the orientation of eight rectangles, for example, can sometimes be recovered as long as one second after stimulus extinction. This long persistence suggests that the information cannot be subcortical but must be of cortical origin, and Block concludes that for this reason it is likely to be phenomenal.

So Block's argument contains this critical step: the step consisting in assuming that if something provides an impression of detail, and its substrate is cortical, then it is likely to be phenomenal. This makes sense to Block because he starts from an a priori assumption that consciousness is a "natural kind" and has some kind of "neural signature." Converging evidence, he says, suggests that if certain neural conditions are met (e.g., being cortical rather than subcortical), then visual information in the brain becomes conscious.

However, the neural signature hypothesis is merely speculative. It could indeed turn out that there is neural commonality in every case of consciousness, but why should one start out with this assumption? Moreover, even if it were the case that an identifiable neural signature for consciousness existed, what would its significance be? Block himself suggests it would leave us still struggling with a "hard problem" of consciousness, showing that this neural signature would give us little insight into phenomenal consciousness.

Leaving aside the criticism of Block's use of recent empirical data in defense of his hypothesis, we are surprised at another aspect of his argument.

Block once had a program of disproving philosophical functionalism. There were two kinds of consciousness: phenomenal consciousness and access consciousness. Phenomenal consciousness was pure and *nonfunctional.* Block thought the existence of phenomenal consciousness (for which he adduced empirical and theoretical material) proved there was more to consciousness than the functional (Block 1995b).

The problem, critics quickly pointed out, was that something purely nonfunctional, something that has no effects, and in principle can have no effects, is not only undiscoverable (any observation would be an effect), but epiphenomenal as well: it cannot even have any effects on the phenomenology of the subject in whose consciousness it is present. Nevertheless, if the goal of showing the plausibility of a nonfunctional phenomenal consciousness were attainable, this would be an important result.

But in the target article the opposition between the realm of the phenomenal and the realm of access is diluted. Today's version of phenomenal consciousness is no longer completely access-resistant. After all, Block musters evidence for its existence by noting that people report that they think they have seen everything in the scene even though they cannot always report all the details. So this is a report, just not a full report. Furthermore he claims further evidence by saying that they *can* report the details if cued early enough. This is also a report!

Thus, instead of the original strong opposition between functional access consciousness and nonfunctional phenomenal consciousness, there now seems actually to be a tight link between this new (avowedly less-than-) cognitive access, and phenomenal consciousness. Indeed, nothing in the present target article indicates phenomenal consciousness cannot be interpreted as a weaker variant of access consciousness. Importantly, and further encouraging this interpretation, phenomenal consciousness also no longer gets identified with the qualitative, often called "what-it-is-like" aspects of consciousness.

If phenomenal consciousness is no longer sharply distinguished from access consciousness, a significant issue on the Scientific Explanation of Consciousness is transformed into a rather minor quibble about the scope of the access subjects have to visual information in psychological experiments! Deciding whether subjects can describe four or eight masked items after a certain number of milliseconds is perhaps an important

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

project in psychology and neuroscience, but resolving the question does not seem like a big advance on the mind/body problem.

Reuniting (scene) phenomenology with (scene) access

doi: 10.1017/S0140525X07003019

David Papineau

Department of Philosophy, King's College London, Strand, London WC2R 2LS, United Kingdom.

david.papineau@kcl.ac.uk

http://www.kcl.ac.uk/kis/schools/hums/philosophy/staff/d_papineau.html

Abstract: Block shows that we can consciously see a scene without being able to identify all the individual items in it. But in itself this fails to drive a wedge between phenomenology and access. Once we distinguish scene phenomenology from item phenomenology, the link between phenomenology and access is restored.

Block describes how in Sperling's (1960) classic study subjects were briefly shown three rows of four alphanumeric characters. The subjects reported seeing the array, but when asked to identify the characters could only name four or five. Block invokes this study and an impressive number of further considerations to argue that "phenomenology overflows accessibility" (sect. 9) – you can be consciously aware of the alphanumeric array even though you cannot cognitively access the individual characters.

I agree that there is one sense in which this is true. But I'm not sure that it supports Block's larger program of distinguishing "phenomenal consciousness" from "access consciousness."

Suppose we distinguish "scene phenomenology" from "item phenomenology." In relation to the Sperling display, you have the former once you are phenomenally conscious of a 3×4 array, whereas you have the latter if you are phenomenally conscious of the individual characters. Now focus on scene phenomenology: Block's arguments seem to me to leave it quite open that scene phenomenology may constitutively require some kind of access. After all, his arguments depend crucially on the fact that normal subjects report that they are visually aware of an alphanumeric array – this is his basic reason for saying that there is more in phenomenology than the four or five characters they can name. At first pass, this certainly suggests that scene phenomenology requires at least "scene access" in some sense that allows normal subjects to report that they are aware of an array, even if they cannot identify all the characters.

No doubt Block would urge that scene phenomenology does not require even scene access. He maintains that the patient G.K. is sometimes consciously aware of a face even though G.K. denies seeing any such thing. Block also suggests that there can be phenomenology in Fodorian modules whose encapsulation presumably precludes cognitive access of any kind. Still, while Block makes these claims, it is not clear to me that he offers any arguments for scene phenomenology without scene access. Rather, his arguments all concern the possibility of scene phenomenology without item access, which is rather different. He shows that I can see the alphanumerical array without being able to report all the characters, but not that I can see it without even knowing that I am seeing an array. (Nor is it is entirely clear to me that Block is consistent in urging the possibility of scene phenomenology without scene access. Consider what he says in section 11 against the suggestion that a state is conscious if it is broadly cognitively accessible in the sense that it can be "amplified if attended to." Block's objection is that this would let in some "uncontroversially unconscious" states. But what makes a state "uncontroversially unconscious" if it is not that subjects tell us so? Once we allow that a state can be conscious even though normal subjects systematically deny this, then I'm not sure there will be any uncontroversially unconscious states.)

Having got this far, we might well wonder whether Block really does anything at all to separate phenomenology and access. He shows that we can have scene phenomenology without item access, but not that we can have scene phenomenology without scene access, or item phenomenology without item access.

Block argues that we need to let phenomenology without access explain the kind of data displayed in the Sperling experiment: in such cases we have phenomenology (in the back of the head) without cognitive access (in the front). But, for all he says, an equally good explanation would be that we have scene phenomenology/access (in the front of the head) without item phenomenology/access (which would also be in the front).

Does it make sense to posit scene phenomenology without item phenomenology? Can we be conscious of the array without being conscious of the individual characters? I don't see why not. If visual perception involved something like physical photographs in the brain, maybe this would be impossible. You can't photograph an array of characters without photographing the individual characters. But if conscious seeing is always seeing as – always a matter of bringing back-of-the brain activity under stored patterns – then I don't see why you can't consciously recognize something as an array without consciously recognizing the individual characters.

I have been suggesting that Block's arguments fail to drive a wedge between phenomenology and access. However, this is not because I am a "metaphysical correlationist," as Block suggests in his section 7. I don't think that there is any principled reason for insisting that consciousness must be tied to access. The passage that Block cites (from Papineau 2002, p. 187; see target article, sect. 7, para. 2) is my characterization of the "standard methodology" of consciousness research, not my own view. As is made clear in the later sections of the chapter from which Block quotes me, I am quite open to the possibility that there may be "hidden" conscious states of just the kind Block is interested in – states with phenomenology but to which the subject lacks cognitive access. Still, although I am open to this possibility, I don't think that Block's target article mounts an effective empirical case for such states, for the reasons given earlier.

In my book (Papineau 2002), I conjectured that this issue will always be incapable of resolution, and suggested that this is because the very concept of phenomenal consciousness is too indeterminate to allow serious empirical investigation of its boundaries. But now I am not so sure. Block's richly textured use of the empirical data may not amount to a conclusive case for phenomenology without access, but it certainly teaches us a great deal about both. It also makes me less pessimistic about the possibility that further empirical data may cast even more light on the boundaries of phenomenal consciousness.

Accessed, accessible, and inaccessible: Where to draw the phenomenal line

doi: 10.1017/S0140525X07003020

Jesse Prinz

Department of Philosophy, University of North Carolina, Chapel Hill, NC 27514.

jesse@subcortex.com

http://www.unc.edu/~prinz

Abstract: One can distinguish among perceptual states that have been accessed by working memory, states that are accessible, and states that are inaccessible. Block compellingly argues that phenomenology outstrips access but wrongly implies that phenomenology outstrips

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

accessibility. There is a subjective difference between Sperling cases and inattentional blindness, which suggests that phenomenology occurs under conditions of accessibility, and not inaccessibility.

For more than a decade Ned Block has been defending the thesis that phenomenal experience can occur without cognitive access. In his target article here, he compellingly argues for the more restricted view that phenomenal experience outstrips accessibility *in a narrow sense*. In Block's terminology, a perceptual state is narrowly accessible if it has been encoded in working memory (or encoded in a "global workspace" in frontal cortex). It would be more perspicuous to describe such states as "accessed," rather than "accessible." In contrast to these, there are states that are accessible but not yet accessed, and states that are totally inaccessible (Dehaene & Nacchache 2001). Therefore, we may ask whether consciousness outstrips mere accessibility in addition to access. Block is largely agnostic about this question here, but he implies that phenomenology outstrips accessibility; his methodological claims seem to entail this conclusion, and his phenomenal/access distinction may depend on it (Block 1995b). In this commentary, I argue that Block can grant the existence of inaccessible phenomenal states only on pain of multiplying the number of phenomenal states beyond plausibility.

Elsewhere I argue that phenomenal consciousness requires attention (Prinz 2005; 2007; forthcoming). I define attention as the process that allows representations in perceptual centers to gain access to working memory. Attended representations are accessible, but not necessarily accessed, as confirmed by studies of divided attention and passive movie watching (Goldberg et al. 2006; Kouider et al. 2007). If this account is right, the neural correlate of consciousness is the neurocomputational process that underwrites accessibility. That process might be a distinctive pattern of neural activation, or perhaps activation levels above a critical threshold. When the process occurs, the affected neurons can send afferent signals to working memory structures in frontal cortex, but task demands, lateral competition, and other factors determine which neurons in the window of attention propagate beyond perceptual centers and get encoded in working memory. Thus, there is a three-way distinction between unattended, attended, and encoded, or inaccessible, accessible, and accessed.

Block correctly concludes that phenomenology outstrips working memory encoding, but he underestimates the behavioral evidence for the view that accessibility is necessary for phenomenology. Research on inattentional blindness, attentional blink, extinction, and visual neglect strongly suggest that phenomenal experience disappears when attention is withdrawn. Block seems to disagree. He implies that phenomenology can occur without attention (consider his discussion of extinction and his enthusiasm for the view that reentry is the neural correlate of consciousness). Block implies that cases of inattention are comparable to Sperling cases (and standard change blindness experiments), in which participants report seeing a stimulus but cannot identify it. He may think that we have phenomenology in both cases. If so, this is a fatal flaw in his position. Sperling cases and cases of inattention are radically different. In cases of inattention, participants insist that they have seen nothing, and they are at chance in guessing whether there was an object presented in their unattended fields. There is no reason to postulate phenomenology under total inattention, and overwhelming reason to postulate phenomenology in Sperling cases (and cases of change blindness), where subjects attend to the stimuli and have vivid experiences, despite limitations of reportability. In my terminology, Sperling cases involve accessibility without (complete) encoding, and inattention cases render stimuli inaccessible. Therefore, consciousness depends on accessibility.

Block might counter with a methodological objection: The claim that phenomenology disappears under inattention is based on what participants report in the relevant studies, and taking reports as authoritative begs the question against those who believe that phenomenology outstrips accessibility. Against this, I have two replies. First, the evidence does not rely solely on reports. There are dramatic *subjective differences* between inattentional blindness and Sperling cases; anyone who has seen both kinds of stimuli can experience the contrast. There are presumably also psychometric and physiological differences that correlate with the subjective differences, and these can be used to test for phenomenology when reports are unavailable or unreliable (as with aphasia or anosagnosia). Under non-pathological conditions, every subject can subjectively confirm that when phenomenal experiences occur, we can at least report that *something* was experienced. That gives us reason to trust such reports in others. Reports can be subjectively confirmed.

Second, Block implicitly relies on reports in defending his own conclusions. For example, he dismisses the view that activations in low-level visual areas are conscious by appeal to the fact that there are low-level activations corresponding to unseen stimuli in binocular rivalry studies. But the assumption that participants in rivalry studies are not experiencing both concurrently presented stimuli is based on what they report. Block must either concede that reports are authoritative or deny that we can rule out the possibility of conscious states in V1 and earlier visual areas.

In response, Block should reject the conjecture that both stimuli are conscious in binocular rivalry on subjective grounds: When you participate in such a study, you feel like you are seeing just one stimulus. The subjective method of determining which states are conscious does not presuppose that consciousness depends on reportability; even if I don't trust the reports of other participants, I can participate in the studies and see for myself. If Block grants that subjective confirmation is authoritative, then he should concede that there is no experience of stimuli under conditions of inattention. If he denies that subjective confirmation is authoritative, then he will have no way of ruling out that we have conscious experiences of both stimuli in binocular rivalry – indeed, Block will have no way of ruling out the possibility that there are phenomenal states in the LGN and the retina. On the first horn, Block would have to concede that phenomenology requires accessibility, and, on the second horn, he would proliferate the number of phenomenal states beyond plausibility. In sum, I think Block can postulate inaccessible phenomenal states only on pain of rendering every visual response conscious. The dramatic subjective distinction between Sperling cases and cases of inattentional blindness undermines efforts to establish phenomenology without accessibility.

Phenomenological overflow and cognitive access

doi: 10.1017/S0140525X07003032

David M. Rosenthal

Program in Philosophy and Concentration in Cognitive Science, City University of New York Graduate Center, New York, NY 10016-4309.

davidrosenthal@nyu.edu

http://web.gc.cuny.edu/cogsci/dr.htm

Abstract: I argue that the partial-report results Block cites do not establish that phenomenology overflows cognitive accessibility, as Block maintains. So, without additional argument, the mesh he sees between psychology and neuroscience is unsupported. I argue further that there is reason to hold, contra Block, that phenomenology does always involve some cognitive access to the relevant experience.

Block's argument for the overflow of cognitive accessibility by phenomenology relies primarily on the partial-report results of Sperling (1960) and others. I argue first that those results do not by themselves support that overflow. I then raise concerns

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

about whether phenomenology could be altogether independent of cognitive access, as Block maintains.

Subjects in Sperling's experiment can identify only a few of the presented letters, but they report that they are all letters. Block convincingly argues that this limitation in identifying results from constraints imposed by visual working memory. He also urges that "phenomenology persists beyond the stimulus" (sect. 9, para. 4), since post-stimulus cuing enables subjects to focus on the subset of items they then identify. Because that cuing occurs after the stimulus has ceased, one reasonable hypothesis is that the identifying relies on persisting phenomenology.

Still, subjects identify all the items as alphanumeric independent of any cuing; as Block notes, subjects "can bring them [all] under a general concept like 'alphanumeric character'" (sect. 9, para. 2). So all the items are at least partially conceptualized independent of cuing, presumably before the stimulus ceases.

So it may well be that, even before the stimulus ceases, the items are all conceptualized in respect of their specific alphanumeric identity. Identification of specific letters might then rely solely on that conceptual information, rather than on persisting phenomenology. The working-memory bottleneck would still be operative; conceptual representations of the specific identities of all the letters would be available prior to cuing, but only about four could get through to figure in subjects' identifications.

On this model, only persisting conceptual representations figure in whatever identifications the bottleneck permits. As with the Global Workspace model of Dehaene and Naccache (2001), which Block considers, the limit on specific identifications results from a constraint on how much conceptual information can get through.

Perhaps phenomenology does persist, as Block urges. But some conceptual representations plainly persist as well, since subjects can identify all the items as alphanumeric long after phenomenology has clearly ceased. So if phenomenology does persist, it is accompanied by persisting conceptual representations, and those representations may by themselves be the basis of whatever specific identifications subjects can make.

The same holds for the related results of Landman et al. (2003). Block urges that these "[s]ubjects are apparently able to hold the visual experience for up to 1.5 seconds" (sect. 9, para. 7). But, as with the Sperling result, subjects' post-cuing ability to detect which rectangles have changed may be the result, not of persisting phenomenology, but rather, of persisting conceptual representations.

Block takes phenomenology to overflow cognitive access because subjects must have "persisting experiences as of more specific shapes than can be brought under the concepts required to report or compare those specific shapes with others" (sect. 9, para. 10). But the working-memory bottleneck may instead operate by restricting the number of fine-grained conceptual representations that can figure in reporting and comparing.

The upshot of the partial-report results, then, would not be "phenomenological overflow" (sect. 9, para. 13), but rather overflow of the working-memory bottleneck by fine-grained conceptual representations. Since the mesh Block argues for is "between the psychological result that phenomenology overflows cognitive accessibility and the neurological result that perceptual representations that do not benefit from attention" (sect. 14, para. 10), that mesh is in jeopardy. So it cannot help resolve Block's puzzle about "whether the cognitive access underlying reportability is a constitutive condition of phenomenal consciousness" (sect. 2, para. 9).

In any case, there is reason to think phenomenology does always involve cognitive access. Block notes that "when one has a phenomenally conscious experience [i.e., an experience with phenomenology], one is in some way aware of having it" (sect. 6, para. 1). This reflects the widespread intuition that phenomenology is a matter of how our mental lives appear to us. There being something it is like to have an experience is just a matter of such mental appearance. Block denies that being thus aware of an experience itself involves some cognitive access to it, but his reasons for that rejection are unconvincing.

One way to explain why phenomenology always involves awareness of the experience is to posit a distinct higher-order state about the experience (Rosenthal 2005). Block notes that phenomenology occurs even when we do not notice, attend to, or think about it. But that does not rule out having inattentive, nonconscious thoughts about one's phenomenology.

Block also notes that young infants are often in pain, though they may be incapable of such higher-order states. But it is possible that infants' pains are not conscious pains. A state can have qualitative character even if there is nothing it is like for one to be in that state (Rosenthal 2005, Ch. 7); it is question-begging here simply to assume otherwise. And because nonconscious pains have largely the same causal connections to behavior and to a strong desire for pain to cease, they are just about as bad. So infants' pains are compatible with the hypothesis that higher-order cognitive states are required for pains to be conscious – for there to be something it is like for one to have them.

Block also argues that a higher-order theory of consciousness "fits both science and common sense" (sect. 6, para. 3) less well than same-order theories, on which the awareness of an experience is internal to the experience itself. But according to standard same-order theories (Brentano 1874/1924; Kriegel 2005), that awareness is every bit as cognitive as on the higher-order-thought hypothesis (Rosenthal 2005).

Block's concession that phenomenology always involves being aware of it also raises a difficulty he does not address. Phenomenology may occur, he holds, when there is activation in the fusiform face area of the extinction patient G.K., even though G.K. "not only does not know about, but ... cannot know about" (sect. 14, para. 13) that phenomenology. But awareness of things standardly results in one's knowing about them and in being able to express that knowledge in verbal reports. If phenomenology always involves awareness of it, why would such awareness fail here to have those results?

Perhaps Block thinks that G.K. might in this case have only a weak phenomenology, like the nonconscious qualitative character that may occur in infants. But that is not ordinary phenomenology, which involves there being something it is like to have it.

Higher-order theories of consciousness and many same-order theories avoid this difficulty, since they hold that some cognitive awareness is constitutive of all phenomenology. By contrast, the global-workspace theories Block favors do not help explain why ordinary phenomenology always involves some awareness of it.

Block argues for the possibility of phenomenology wholly without cognitive access. But the awareness that always accompanies phenomenology and the compelling intuition that phenomenology is essentially a matter of mental appearance together suggest a close tie between phenomenology and some type of cognitive access.

Conscious access overflows overt report

doi: 10.1017/S0140525X07003044

Claire Sergent and Geraint Rees

Institute of Cognitive Neuroscience & Wellcome Department of Imaging Neuroscience, Institute of Neurology, University College London, London WC1N3AR, United Kingdom.

csergent@fil.ion.ucl.ac.uk

http:// www.icn.ucl.ac.uk/Research-Groups/awareness-group/

g.rees@fil.ion.ucl.ac.uk

http://www.fil.ion.ucl.ac.uk/~grees

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

Abstract: Block proposes that phenomenal experience overflows conscious access. In contrast, we propose that conscious access overflows overt report. We argue that a theory of phenomenal experience cannot discard subjective report and that Block's examples of phenomenal "overflow" relate to two different types of perception. We propose that conscious access is more than simply readout of a pre-existing phenomenal experience.

Visual presentation of a face can elicit strong activity in the fusiform gyrus even when a patient with neglect does not see it (Rees et al. 2000). Block's proposition that recurrent sensory activations constitute the core neural basis of phenomenal experience leads him to the following prediction:

> [I]f the activations of the fusiform area [. . .] in the patient G.K. turn out to be recurrent activations, we would have evidence for phenomenal experience that the subject not only does not know about, but in these circumstances cannot know about. (target article, sect. 14, para. 13)

In other words, the experimenter could detect phenomenal experience in a patient's brain *even if* the patient denies having that experience! This contradicts the notion of phenomenal experience itself, which is, by definition, a subjective, first-person experience. We suggest that any theory of phenomenal experience that discards subjective report as its fundamental measure is bound to lead to such contradictory predictions.

It is crucial to distinguish different forms of accessibility. In experiments using backward masking (Lau & Passingham 2006; Vorberg et al. 2003), when subjects are pressed to guess which stimulus has been presented they can perform better than chance, even when they claim they have not seen it. In this case, as in classical "blindsight" (Weiskrantz 1997), there is a clear distinction between *two types of report*: objective report (accessibility in a direct sense, probably via automatic stimulus-response routes) and subjective report (accessibility as defined in the term "conscious access"). We propose that this latter form of accessibility is intrinsic to what Block calls phenomenal consciousness.

In our view, the real challenge lies in designing tools to allow an objective, scientific measure of subjective report (Dennett 1992). Indeed, this is already an active area of study (Dienes & Scott 2005; Macmillan & Creelman 1991; Persaud et al. 2007; Sergent et al. 2005; Sergent & Dehaene 2004; Weiskrantz 1997) and could in principle produce biomarkers of consciousness that allow report of phenomenal experience in the absence of *overt* report, via external brain imaging devices (Naccache 2006b; Nachev & Husain 2007; Owen et al. 2006).

Block makes extensive use of partial-report paradigms (Landman et al. 2003; Sperling 1960) to support the argument that phenomenal experience overflows conscious accessibility. While an observer viewing such a letter array appears to "see it all," uncued report of the individual letters a short time later is poor. Nevertheless, a cue occurring after the array has been presented can lead to excellent recall of letter identity for the cued subset of the array. This excellent cued recall does not entail – as Block argues – that the earlier "seeing it all" experience therefore reflects phenomenal experience of all the letters *at the same level of detail*. Instead, we suggest that cued recall not only reflects readout of a fading information buffer, but it is a *modification of perception through attention* that does not differ much from classical sensory modulation by spatial attention (Ruff et al., in press). Although only the cued letters receive this additional attentional enhancement, the claim of "seeing it all" might still be based on access to lower-level attributes that are not explicitly probed by the experimenter – for example, presence of the uncued stimuli, rough recognition as "alphanumeric characters," or "Gestalt" properties of the scene gist. The apparent intuition lying behind phenomenal consciousness might therefore simply refer to access itself being much greater in some situations than the specific report required by the experimenter.

Strong sensory processing is consistently observed even when observers deny *any form* of experience of the stimulus. For example, the "attentional blink" describes impairment in reporting the second of two visual targets when they are separated by a short interval. Sergent and colleagues asked subjects to evaluate the visibility of a target word during the attentional blink on a continuous scale going from 0 to 100% visibility. If there were any form of experience of the target, even feeble or partial, they could acknowledge that by using intermediate visibility levels. However, missing a target corresponded to a subjective rating of zero, just as when no word was presented (Sergent & Dehaene 2004). Nevertheless, these "zero-visibility" words elicited strong, long-lasting brain potentials up to 300 msecs after they had been presented (Sergent et al. 2005), a time scale at which local recurrent loops ought to have been established (Lamme & Roelfsema 2000). Therefore, strong and elaborate processing within sensory areas is not sufficient for subjective experience.

In summary, we believe that what Block calls "phenomenal consciousness" actually reflects two different situations:

1. *Genuinely unconscious* stimuli that nevertheless receive strong sensory processing: this type of stimuli can influence our conscious mental life (Dehaene et al. 1998; Gaillard et al. 2006; Greenwald et al. 2003), but are not directly conscious.
2. Conscious stimuli that do not receive full attention and for which conscious access is partial to only certain attributes of the stimuli.

Finally, Block argues that activity in fronto-parietal areas does not reflect the *core* neural basis of consciousness and instead constitutes "read out" of the perceptual state represented in occipito-temporal areas. But the empirical evidence does not support a clear distinction between "perceptual" and "report" areas. For example, enhanced connectivity between fronto-parietal and visual cortices is associated with fluctuations in conscious contents during binocular rivalry (Lumer et al. 1998; Srinivasan et al. 1999) even when observers are not required to make any covert or overt reports of their experience (Lumer & Rees 1999). Therefore, the evidence points to an active interaction among frontal, parietal, and sensory areas that altogether constitute conscious perception (Dehaene et al. 2003; 2006; Rees et al. 2002a).

In conclusion, we are far from understanding the complexity of the neural mechanisms that underlie our rich subjective experience of the external world, and Block's article provides a stimulating reminder of the path that lies ahead. But, in our view, the study of conscious access through careful assessment of subjective reports is the most positive and promising movement in that direction so far.

ACKNOWLEDGMENTS

This work is supported by the Wellcome Trust (for Geraint Rees), and the Fyssen Foundation and the Marie Curie European Program (for Claire Sergent).

Global workspace theory emerges unscathed

doi: 10.1017/S0140525X07003056

Murray Shanahan[a] and Bernard Baars[b]

[a]*Department of Computing, Imperial College London, London SW7 2AZ, United Kingdom;* [b]*The Neurosciences Institute, San Diego, CA 92121.*

m.shanahan@imperial.ac.uk **http://www.doc.ic.ac.uk/~mpsha**

baars@nsi.edu **http://nsi.edu/users/baars**

Abstract: Our aim in this reply is to defend Global Workspace theory (GWT) from the challenge of Block's article. We argue that Block's article relies on an outdated and imprecise concept of access, and perpetuates a common misunderstanding of GWT that conflates the

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

global workspace with working memory. In the light of the relevant clarifications, Block's conclusion turns out to be unwarranted, and the basic tenets of GWT emerge unscathed.

Block's threefold strategy is as follows: (1) He seeks to separate phenomenology from what he calls "access" on the basis of empirical psychology. (2) Shifting to neuroscience, he locates phenomenology in posterior brain regions and access in frontal brain regions. (3) He then concludes that mechanisms of access are not constitutive of mechanisms of phenomenology and claims that this conclusion meshes with both the psychology (1) and the neuroscience (2). However, we contend that the separation of phenomenology and access attempted in (1) fails because it relies on too crude a notion of access. Adopting the more precise notion of broadcast (Shanahan & Baars 2005) allows us to preserve the intimate relationship between phenomenology and the global workspace posited by Global Workspace theory (GWT). And we believe the attempt in (2) to isolate "mechanisms of access" from "mechanisms of phenomenology" is also flawed, because it relies on a faulty identification of the global workspace with working memory.

In contemporary elaborations of GWT, the emphasis is on discovering a plausible neurological basis for the hypothesised global workspace – the global neuronal workspace as Dehaene and his colleagues call it (Dehaene et al. 1998). Dehaene and a number of others suggest that phenomenal consciousness of the visual world requires large-scale re-entrant interaction between posterior visual cortex and frontoparietal regions (see also Baars 2002). According to the leading proposals, the neural substrate for the workspace is a network of long-range pathways that interconnects distant cortical sites (Dehaene et al. 2006; Shanahan, in press). A pattern of activation is globally disseminated, or *broadcast*, when it is propagated by this white matter infrastructure, and thereby comes to exercise widespread influence in the brain. Because the great majority of such pathways are bidirectional, the anatomy suggests widespread re-entrant activity triggered by the initial broadcast. The overall pattern of information flow alternates episodes of broadcast with bursts of competition for workspace access, and the typical duration of discrete episodes of broadcast is on the 100 msec scale.

Using the idea that broadcast implements access, we can supply a more fine-grained account of what might go on in the sorts of experiments that, according to Block, demonstrate "phenomenology without access" (Sperling 1960; Landman et al. 2003). It is worth noting that subjects in such experiments believe they are simply reporting the contents of their recent visual consciousness, even after the visual stimulus has disappeared from view. Therefore, subjects' reports do not support a distinction between access and phenomenology.

The theoretical account might go something like this (Shanahan 2005). Information pertaining to the initial visual object gestalt in its entirety is broadcast at the time of presentation for one or more 100 msec broadcast cycles. As a consequence of broadcast, this information can influence widespread local brain processes involved in memory – both short-term and long-term. Subsequently, following some retrieval cue, a detailed memory of a portion of the original scene is recalled from the relevant part of short-term memory and is broadcast, enabling (with the help of many other brain processes) a verbal report. Thanks to this process of retrieval and report, or perhaps because of subsequent activity, or merely due to the time that has elapsed, the fragile contents of the short-term memory are degraded and a good deal of information about the original visual stimulus is lost. So no exhaustive subsequent reports detailing the shapes and orientations of specific objects in the scene are possible. But post-hoc partial reports can be retrieved with very high accuracy, which is of course why the Sperling procedure is called the "partial report procedure." Over multiple trials the entire 12-object Sperling array can be accurately reported, always with a reported phenomenal experience of the retrieved item. Moreover, a general impression of the original visual stimulus is retained in short-term memory. If later evoked, this information can still be broadcast, facilitating a verbal report about the overall arrangement of objects in the scene.

The very possibility of such an account shows that, even if they did supply evidence for phenomenology without "access." in Block's temporally coarse-grained sense of the term, experiments such as those by Sperling and Landman et al. offer no evidence of *phenomenology without broadcast*, to use the more precise and temporally fine-grained technical term found in the contemporary GWT literature. Admittedly, the theoretical concept of broadcast stands in need of further empirical justification. But it is, in this respect, in no worse a position than that of access.

Now, the above explanation makes little sense unless a clear distinction is maintained between (conscious) presence in the global workspace and (unconscious) retention in short-term memory. But Block incorrectly identifies the global workspace of GWT with working memory, leading to further difficulties with his argument. In humans, presence in the global workspace *enables entry to working memory* (Baars & Franklin 2003). But items can then be held in working memory unconsciously. They only become conscious again when they are retrieved from working memory and re-enter the global workspace.

In short, the global workspace and unconscious working memory are quite distinct, and for this reason it does no harm to GWT if it turns out that, as Block suggests, "activation of working memory circuits are not part of the neural basis of phenomenology" (sect. 14, para. 1). What counts for phenomenology, according to contemporary GWT, is broadcast via the long-distance axons alluded to earlier on, and this may or may not necessitate activation of the prefrontal areas associated with working memory.

To summarise, Block's challenge to GWT rests on a faulty view of the concept of the global workspace. GWT is an empirical theory, and as such it is open to both experimental refutation and conceptual critique. So challenges such as this are a welcome opportunity to test the theory, and a chance to revise or discard it if it is found wanting. But to mount a more successful attack on the theory, Block needs to consider more detailed GWT models of broadcast, and clearly to dissociate global workspace activity from unconscious aspects of working memory.

Access for what? Reflective consciousness

doi: 10.1017/S0140525X07003068

Michael Snodgrass and Scott A. Lepisto

Department of Adult Ambulatory Psychiatry, University of Michigan, Ann Arbor, MI 48105.

jmsnodgr@umich.edu **slepisto@umich.edu**

Abstract: Can phenomenality without access occur? We suggest that the crucial issue is not to show phenomenality that *cannot* be accessed, but whether phenomenality sometimes simply *is not* accessed. Considering this question leads to positing a distinct, second form of consciousness: Reflective consciousness. The most important form of access is then from phenomenal (first-order) to reflective (second-order) consciousness.

Block's primary concern involves whether phenomenal consciousness (henceforth, P-consciousness) is independent of what he calls "cognitive accessibility." He argues yes (e.g., that P-consciousness can "overflow" access), and contrasts this view with that of others (e.g., Baars, Dehaene, Dennett, Naccache, etc.) who argue instead

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

that access (for them, a.k.a. being in the "global workspace") is in part constitutive of P-consciousness – and hence that phenomenality without access cannot occur.

These fundamental questions can be usefully engaged from a framework informed by signal detection theory (SDT; e.g., Macmillan & Creelman 1991). To foreshadow, to show that P-consciousness is independent of access does not require demonstrating that there are P-conscious contents that *cannot* be accessed (Block's strategy). Rather, one need only show that sometimes P-conscious contents *are not* accessed, whereas sometimes they are. In this way, even if Block is wrong, one can still show that access is indeed an independent conscious process distinct from P-consciousness proper.

The easiest way to show this is to sketch our theory of consciousness, which derived from attempting to solve a perennial problem in unconscious perception research: How can consciousness be validly indexed? It seems clear that we must rely in some way upon subjects' reports – but what *kind* of report?

First, we could simply inquire about participants' subjective conscious experience; for example, "Can you see (hear, remember, etc.) *X*?" If they say no, they are not conscious of *X*; if yes, they are. Using such methods, ostensibly unconscious effects (including above-chance discrimination) are easily obtained. Despite "subjective threshold" methods' intuitive appeal, however, SDT shows that *all* reports involve not just (1) the perceptual process itself, but also (2) an additional, independent criterion-based decision process wherein the strength of available perceptual evidence is evaluated as to whether it is sufficient to reply in the affirmative. That is, contrary to many researchers' implicit assumptions, there is no such thing as an unmediated "subjective report" – ever. Consequently, the SDT analysis directly implies that if someone says "no," this does *not* show that they have no relevant P-conscious experience, but perhaps merely that such experience, being very weak and hence low-confidence, is below criterion. With this in mind, subjective threshold effects – which are easily obtainable – are plausibly weakly P-conscious, not unconscious.

This "conscious but weak" interpretation is strongly supported by the under-appreciated fact that the same result (i.e., the same estimate of perceptual sensitivity, d') is obtained in both the yes/no form of the detection task *and* in the so-called rating form, wherein participants rate their confidence that a target stimulus (vs. noise only) has appeared. This means that whenever $d' > 0$, participants *do* in fact possess relevant subjective confidence – however weak – which enables above-chance stimulus discrimination.

Despite the above, some argue that subjective threshold stimuli really are phenomenally unconscious (henceforth, P-unconscious) after all. Perhaps the most compelling are various demonstrations of "exclusion failure" (e.g., Merikle et al. 1995), wherein participants respond with a previously presented word above baseline despite instructions not to, apparently suggesting that such stimuli are unconscious. Nonetheless, these effects too can be explained as weakly P-conscious if exclusion is a decision process such that sufficient confidence is required to exclude a candidate response (Snodgrass 2002; see also Haase & Fisk 2001).

Consistent with this analysis, Visser and Merikle (1999) found that exclusion failure disappeared when participants were paid for successful exclusion – a classic criterion manipulation (SDT has long shown criterions can be shifted at will). This finding is crucial, because it shows that subjective threshold stimuli *can* be accessed, even though typically they are not. Here again, this is exactly what one would expect from (initially) below-criterion but nonetheless weakly conscious stimuli.

Perhaps instead, then, we should use the second form of report – namely, by asking participants to objectively discriminate the stimuli rather than deliver opinions about their phenomenal experience. By using SDT methods, we can independently estimate *d'* and the criterion (*c*). In this way, if discrimination responses are required even when confidence is low, criterion artifacts of the sort just mentioned are avoided, and if stimulus conditions are arranged such that $d' = 0$ (i.e., the "objective threshold"), there are strong reasons to conclude that any other obtained effects are truly not P-conscious. Strikingly, under these conditions, clearly qualitatively different phenomena emerge (Snodgrass et al. 2004; Snodgrass & Shevrin 2006) including both facilitation and striking *below*-chance inhibition – which never occurs under subjective conditions.

So are subjective methods simply invalid? Despite the above, we think not, because they *do* index a centrally important phenomenon: whether or not higher-order, controlled (vs. automatic) processing is applied to the relevant stimuli. In other words, subjective methods index our criterion for *access* to higher-order processes including reasoning, planning, and so forth. After all, strategic control processes will not be attempted if we don't think the relevant stimuli exist.

With all this in mind, we (Snodgrass 2002; Snodgrass & Shevrin 2006; Snodgrass et al. 2004) propose that there are two, not one, kinds of consciousness: (1) first-order (phenomenal consciousness), and (2) second-order (reflective consciousness), which takes first-order (and other second-order) conscious states as its objects. In this model, while all reflectively conscious states are also phenomenally conscious, the converse is not the case – that is, there are phenomenal states that are not additionally reflectively conscious. Rather, because reflective consciousness is highly capacity-limited, only the most relevant phenomenally conscious stimuli – even if strong – are selected for additional, optional higher-order processing. At the same time, under usual circumstances essentially any phenomenally conscious content can be selected into reflective consciousness if desired, as the SDT analysis of typically subjectively unconscious stimuli demonstrates.

Putting it all together, stimuli can clearly be P-conscious without being accessed. This is most obvious with subjectively unconscious stimuli, which are typically not accessed, but nonetheless exhibit effects qualitatively identical to clearly P-conscious stimuli whenever further (and optional) reflectively conscious (i.e., strategic) processing is absent. Moreover, they *can* be accessed, given our willingness to do so (e.g., use more liberal criteria). In contrast, clearly (objectively) P-unconscious stimuli are intrinsically inaccessible, and display qualitatively different effects.

Expecting phenomenology

doi: 10.1017/S0140525X0700307X

Maja Spener

Faculty of Philosophy, University of Oxford, Oxford OX1 4JJ, United Kingdom.
maja.spener@philosophy.ox.ac.uk

Abstract: Block's argument against correlationism depends in part on a view about what subjects in certain experiments can be aware of phenomenally. Block's main source of evidence for this view is introspection. I argue that introspection should not be trusted in this respect. This weakens Block's argument and undermines correlationism at the same time.

Block's argument against correlationism depends in part on the assumption that subjects in the central experiments (i.e., Sperling 1960, Landman et al. 2003, and Sligte et al. 2008) are phenomenally aware of all or almost all items in the presented arrays. Block suggests that the main source of evidence for this assumption comes from introspection – that is, from subjects' introspective reports that they have seen all or almost all items during the initial brief presentation.

Someone who adopts the methodology of inference to the best explanation, as Block does, should be ready to defend a proposed range of data as relevant to his or her approach. And when it

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

comes to the above-mentioned introspective evidence, at least, Block is sensitive to such an obligation. He defends his claim that we can take the reports in question at face value against the worry that subjects might confuse potential phenomenology with actual phenomenology.

But there is a more general worry about introspective judgements Block should consider. Introspective judgements might be reliable enough when it comes to off-the-cuff or approximate judgements about phenomenology, as for instance when one introspectively judges that it appears as if something is red. Visual phenomenology arguably does involve colour-related features and, arguably, introspection reveals this fact. For the purposes of theorizing about phenomenology, however, not much follows from this general fact. But introspective judgements are demonstrably unreliable about the subtle details of phenomenology, as the philosophical debate about the nature of perception makes perfectly clear. Participants in that debate disagree fiercely on how to characterize phenomenology on the basis of introspection. Views range from characterizations of phenomenology solely in world-directed terms to ones involving intrinsic properties of experience individuated independently of one's environment. The bone of contention here is not *whether* a given case involves actual phenomenology – on this most participants are agreed – but rather *how to describe it accurately on the basis of introspection*. I have argued elsewhere (Spener, under review) that the best explanation of such disagreements in the philosophical literature involve the idea that one's introspective grasp of phenomenology is thoroughly expectation-driven, and, for this reason, that introspection is thoroughly inadequate on its own to provide reliable knowledge of the finer details of phenomenology.

How does this worry affect Block's claim that we can believe subjects' reports that they saw all the items in the presented arrays? Note, first, that this particular bit of introspective evidence plays a crucial role in Block's argument. It provides the main reason to identify the experimentally elicited informational state with what is present to the subject phenomenally. This identification does not come for free, as demonstrated by the experimenters' explicit attempts to argue for it (see Sligte et al. 2008; Sperling 1960). Moreover, despite giving such arguments, Sperling is careful to say that "[s]hort-term information storage has been *tentatively identified* with the persistence of sensation that generally follows any brief intense stimulation" (Sperling 1960, p. 26, my italics). Without the introspective evidence, the claim that the detailed information about the array of items – that which is experimentally shown to be at the subjects' disposal – is indeed phenomenally present to the subject is on much weaker footing. In turn, this means that Block's argument against correlationism is on much weaker footing.

So, what sort of introspective judgements are we dealing with in Block's argument? Specifically, are they of a detailed sort and hence suspicious, or are they of an approximate sort and hence theoretically uninteresting? Given their role in Block's argument, it seems they have aspects of both. On the one hand, Block takes subjects to report introspective judgements about phenomenology which provide reason to identify the relevant phenomenal state with the short-term informational state demonstrated by the experiments. The latter concerns specific details about all the particular items in the presented array. This suggests that the judgements in some sense reveal detailed phenomenal information. On the other hand, though, subjects cannot exactly report their alleged phenomenal awareness of such plentiful detail, but can only "bring [these experiences] under a general concept" (sect. 9, para. 2). This suggests that the introspective judgements are of an approximate sort.

Recall that the introspective judgements in question are meant in some way to give good evidence for subjects' phenomenal awareness of all the items in the array, a kind of phenomenal awareness that can readily be identified with the informational state demonstrated by the experiments. I think there is trouble for that thought with respect to both the aspects just articulated. Consider the detailed-phenomenology aspect first. As the old debate about the problem of the speckled hen (for instance) makes vivid, the right way to characterize the phenomenology of an experience of a plurality of items *based on introspection* is a matter of great dispute (see, e.g., Chisholm 1942; Jackson & Pinkerton 1973). In this case, too, then there can be different introspective takes on the details of phenomenology. This makes room for my worry that introspection-based characterizations of the specifics of phenomenology are partly expectation-driven. They should not be regarded as straightforward, reliable evidence about phenomenology.

Consider now the approximate-judgement aspect of Block's introspective judgements. There clearly is a large gap between the general nature of the introspective judgement and the phenomenology of particular detail they allegedly reveal. That general introspective take on phenomenology could correctly capture several, non-equivalent ways that things look. It is hard to see how the general judgement could in turn provide evidence to single out one of those ways as captured by it.

By Block's lights, correlationism must be rejected because it places mistaken limits on the kind of data relevant and admissible to the enquiry. I suggest, more radically, that correlationism must be rejected because it limits itself to poor data. Introspective evidence should not be abandoned altogether, of course, but it also should not carry too much weight in an argument. I recommend spreading the weight equally, using introspective data alongside data from other sources, for instance, branches of science and philosophy in which perceptual experience figures prominently. The theory that fits such a motley bag of data best overall, should win our allegiance. I take this suggestion to be very much in the spirit of Block's methodological proposal.

ACKNOWLEDGMENT

Thanks to Scott Sturgeon for helpful comments.

Phenomenal consciousness and cognitive accessibility

doi: 10.1017/S0140525X07003081

Michael Tye

Department of Philosophy, The University of Texas at Austin, Austin, TX 78712.

mtye@mail.utexas.edu

http//www.utexas.edu/cola/depts/philosophy/faculty/tye

Abstract: Block tries to show that the results of the Sperling experiment lend support to the view that phenomenology outstrips cognitive accessibility. I argue that Block fails to make a compelling case for this general claim on the basis of the Sperling data.

Ned Block argues that the results of the Sperling experiment (Sperling 1960) lend support to the view that phenomenology outstrips or overflows cognitive accessibility. I shall argue that Block fails to make a compelling case for this claim on the basis of the Sperling data.

I agree with Block that the subjects in the Sperling experiment see all 12 letters in the array even though they can accurately report the identities of only three or four. The first point I want to make is that there is an important sense in which all twelve letters are cognitively accessible notwithstanding the subjects' inability to identify two-thirds of them; for, with respect to each letter in the array, the subjects' experiences enable them at least to wonder "What is that letter?" Their experiences thus put

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

them in a position, with respect to each letter, to bring the letter under the demonstrative concept *that*. The content of the concept varies with the putative query. Were a subject to wonder "What is that letter?" with respect to, say, the letter S in the top right-hand corner of the array, the cognitive content of the query would be different from the content of the query, expressible in the same sentence, with respect to the letter B in the middle of the bottom row. Of course, subjects do not typically so wonder with respect to each letter in the array. They may not so wonder with respect to any letter. The point is that their experiences enable them to wonder, "What is that?" with respect to each letter, whether or not they actually wonder anything, and the content of this possible cognitive response varies with the letter. So, each letter in the array is cognitively accessible via a cognitive act, the content of which is distinctive to that letter.

This brings me to Block's claim that phenomenology outstrips cognitive accessibility. Let us look more closely at how this is supposed to be the case. One way of understanding Block's claim in a particular instance is that the subjects' phenomenal consciousness of the letter S, say, in the top row, is different from the subjects' phenomenal consciousness of the letter B in the bottom row, even under the condition that the middle tone is heard and subjects are attending to the central row. But this does not show that phenomenology outstrips cognitive accessibility, since the possible act of demonstrative wondering with respect to the former letter has a different content from the possible act of demonstrative wondering with respect to the latter.

What happens when the middle tone sounds? One possible view is that there is no change in phenomenology as the subjects attend to the letters in the central row. The subjects simply focus their attention on the row and report what the letters look like. But patently this does not show that phenomenology outstrips cognitive accessibility. I am happy to allow, of course, that when one focuses on a given letter, one is then put in a position by one's experience to identify the letter as being the letter such-and-such and thus to cognitively classify the letter in a way that one could not before. With the shift in attention, then, one's cognitive possibilities enrichen – one goes from possibly wondering "What is that?" to actually thinking that that is the letter P, for example – even though (I am now supposing) the phenomenology stays the same. But this is not to the point. It shows that cognitive accessibility, content-wise, outstrips phenomenology, not the other way around.

I should add that I am actually disinclined to accept that when one shifts one's attention to a particular row, there is no change in the appearance of the letters. It is true that there is no dramatic or big change, but the situation seems to me similar to the following: Focus on an object – a watch, say – lying on a magazine in the middle of your field of view. There will likely be letters on the magazine cover close to the watch that you won't be able to identify even if you try, given your focus of attention. It is not that the letters are fuzzy or obscure; it is just that they are insufficiently determinate in your experience. Now switch your attention slightly from the watch to the letters. With the new focus, you will be able to identify them, but there will be no marked or sudden change in the phenomenology. The letters will not sharply change in their appearance. Still, the phenomenology does change a little. At a conscious level, there is a detail in the letters that simply was not there before – a detail that now enables you to identify them. Subtle changes of this sort happen all the time as we move our eyes around. But these changes are not dramatic or pre-reflectively obvious.

Block sometimes states his position in a more qualified way. He says that, in his view, the reason why subjects are unable to gain access to all the letters in the Sperling experiment without the tones is that "the 'capacity' of phenomenology, or at least the visual phenomenal memory system, is greater than that of the working memory buffer that governs reporting" (sect. 9, para. 11). The working memory system has a capacity of four items (or less) in humans whereas the capacity of visual phenomenal memory "could be said to be at least 8 to 32 objects – at any rate for stimuli of the sort used in the described [i.e., the Sperling and Landman et al.] experiments" (sect. 9, para. 11).

The point I wish to stress again is that cognitive access can be achieved via the use of the demonstrative concept *that letter*. Seeing the 12 letters, one is put in a position by one's experience to query the identity of each letter and to so do, in each case, via a cognitive act that, in part, subsumes the relevant letter under the concept *letter*. This involves the use of working memory. Furthermore, it permits one to issue the report, "That is a letter."

Presumably, what Block has in mind when he says that the capacity of the working memory system is four items (or less) under a certain restricted range of conditions is the capacity to identify (for example) letters as the specific letters they are. But now his claim about phenomenal consciousness outstripping or overflowing cognitive accessibility, although not without interest, does not have quite the bite it first appeared to have. Certainly, nothing that Block says here shows that there fails to be an important connection between the phenomenology of the experience one undergoes in seeing an object and the object's general cognitive accessibility.

What if phenomenal consciousness admits of degrees?

doi: 10.1017/S0140525X07003093

Robert Van Gulick

Department of Philosophy 541 HL, Syracuse University, Syracuse, NY 13244.
rnvangul@syr.edu

Abstract: If the phenomenality of consciousness admits of degrees and can be partial and indeterminate, then Block's inference to the best explanation may need to be revaluated both in terms of the supposed data on phenomenal overflow and the range of alternatives against which his view is compared.

Block argues, by the method of inference to the best explanation (IBE), for the thesis that cognitive accessibility is not part of the constitutive nature of phenomenal consciousness. He claims his view provides the best explanation of the apparent respect in which phenomenology overflows cognitive access and that it meshes with the neural data, which he takes to show that the machinery underlying cognitive access is distinct from that underlying phenomenal consciousness.

As with any case of IBE, the plausibility of Block's argument will depend on both the data it aims to explain and the range of alternative explanations available. Challenges may be raised in both respects, especially if one allows for the possibility that phenomenal consciousness may admit of degrees rather than being an all or nothing matter. If so, critical cases in which phenomenal consciousness apparently occurs in the absence of some types of cognitive access may be better regarded as cases in which a *lesser* degree of phenomenal consciousness involves a *lesser* degree of cognitive access. We need not suppose, for example, that when we have phenomenal visual experience of a scene we phenomenally represent all of its features. Phenomenal consciousness may be partial and incomplete in ways that correlate with our partial cognitive access.

Does phenomenology in fact overflow cognitive access? In experiments such as those by Sperling (1960) and Landman et al. (2003), subjects often report having experienced more features than they can identify, categorize, or detect changes in. However, such reports do not necessarily show that phenomenal consciousness exceeds cognitive access. Though the subjects in the Sperling experiments cannot cognitively identify many of

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

the specific letters that were briefly presented to them, they do have cognitive access to the fact that they were presented with an array of 12 characters. Their cognitive access is limited, but so too may be their phenomenal consciousness of those letters. If one holds a "movie screen of the mind" model of phenomenal consciousness, it may seem impossible that there could be letters that are phenomenally present as letters without being present as specific letter shapes. But such a model is at best problematic, and if one rejects it, then there seems no reason why the characters of which the subjects are aware could not be indeterminate in ways that exactly match their limited cognitive access to those features.

Moreover, the subjects' reports of what they experienced under such marginal conditions may be more problematic than Block contends. He rejects the idea that they might be hyperillusions – cases in which how it seems that it seems, is not how it really seems. He says he knows of no such illusions. However, the well-known color marker demonstration (Dennett 1991) would seem to be just that. We have the firm belief that our phenomenal experience of the entire visual field is colored; that is how it seems to us. But if we hold fixation to the front and hold a marker whose color we do not know at arms length to the side, we cannot discern its color. If it is gradually moved toward the front, we cannot see its color until it is far toward the center of our field of vision, though we can detect its motion almost as soon as it begins to move – which reflects the fact that the retina lacks cones at the periphery but not motion detectors. Nor is the illusion easily dispelled, as Block says is generally true of cognitive illusions. Despite participating in the marker demonstration, one still experiences the illusion (hyperillusion) that one's whole visual field is phenomenally colored.

The possibility that phenomenal consciousness may admit of degrees also enlarges the space of alternatives against which Block's view must be compared in assessing his IBE argument. He contrasts his position mainly with what he calls metaphysical and epistemological correlationism, but these views, like his own, do not consider the possibility that phenomenal consciousness may be partial, indeterminate, and admitting of degrees. If one allows for that possibility, then there are other hypotheses one might consider about whether cognitive access is sometimes a constitutive element of phenomenal consciousness.

If the neural substrate of phenomenal consciousness involves interactions among multiple brain regions, and the specific regions vary for different experiences – two ideas Block endorses in general – then could it not be the case that sometimes those regions most active in providing cognitive access are among the relevant constitutive elements? The suggestion is not merely hypothetical. As Block acknowledges, phenomenal consciousness involves some sort of awareness. Indeed, it involves the awareness or experience of objects present to the self – that is, objects experienced from the perspective of the self. There need not involve a commitment to any traditional substantial notion of the self. The self and the self-perspective may well be virtual structures. As I look at my desk the phenomenal content of my experience is that a blue mug is present here and now before me. In light of the integrative role played by the machinery of cognitive access, it seems plausible to regard it as constituting part of the substrate of my phenomenal consciousness of that mug as an object in my world, rather than merely a cause of my experience.

Perhaps in marginal and attenuated cases, we may have phenomenal experiences of isolated properties or features, and in such cases perhaps all the constitutive elements of our phenomenal consciousness lie outside the structures that support cognitive access. But those would be cases in which we are phenomenally conscious to a lesser degree than is typical. In more normal cases of full-blown phenomenal consciousness, we experience colors, shapes, and motions as properties of objects present to us as part of our world. And in those cases, it seems plausible to treat the regions that present those objects to us, that give us access to those objects, as parts of the substrate of such phenomenal consciousness.

The challenge of disentangling reportability and phenomenal consciousness in post-comatose states

doi: 10.1017/S0140525X0700310X

Audrey Vanhaudenhuyse[a], Marie-Aurélie Bruno[a], Serge Brédart[b], Alain Plenevaux [c], and Steven Laureys[a]

[a]*Coma Science Group, Cyclotron Research Center, University of Liège, Sart Tilman B30, 4000 Liège, Belgium;* [b]*Department of Cognitive Science, University of Liège, Sart Tilman B32, 4000 Liège, Belgium;* [c]*Department of Cognitive Sciences, Cyclotron Research Center, University of Liège, Sart Tilman B30, 4000 Liège, Belgium.*

avanhaudenhuyse@student.ulg.ac.be **www.comascience.org**
marieaureliebruno@hotmail.com **www.comascience.org**
serge.bredart@ulg.ac.be **http://www.fapse.ulg.ac.be/Lab/Cog/**
alain.plenevaux@ulg.ac.be **http://www.ulg.ac.be/crc/index.html**
steven.laureys@ulg.ac.be **www.comascience.org**

Abstract: Determining whether or not noncommunicative patients are phenomenally conscious is a major clinical and ethical challenge. Clinical assessment is usually limited to the observation of these patients' motor responses. Recent neuroimaging technology and brain computer interfaces help clinicians to assess whether patients are conscious or not, and to avoid diagnostic errors.

Block differentiates phenomenal consciousness and its cognitive access based on empirical data obtained with healthy subjects and neglect syndrome patients. Evidencing consciousness in noncommunicative patients is an important theoretical, medical, and ethical issue. In this commentary, we first emphasize the problem faced when determining the presence of phenomenal consciousness in vegetative and minimally conscious states in absence of any verbal response; second, we show that phenomenal consciousness may occur without verbal or motor reports in patients with a complete locked-in syndrome.

Vegetative state (VS) is characterized by "wakefulness without consciousness," accompanied by reflexive motor activity only, devoid of any voluntary interaction with the environment (The Multi-Society Task Force on PVS 1994). Minimally conscious state (MCS) describes patients who are unable to communicate while demonstrating inconsistent but reproducible behavioral evidence of consciousness. Patients in MCS may show command-following, gestural or verbal yes/no response, intelligible speech, and purposeful behavior (Giacino et al. 2002). Neurological practice illustrates how difficult it is to identify signs of conscious perception in such patients. Misdiagnosis of VS occurs in about one-third of patients (Andrews et al. 1996; Childs et al. 1993). Consciousness is a subjective experience and by definition VS and MCS patients are noncommunicative; we are therefore limited to observing their motor responses and interpreting them in terms of consciousness. Functional neuroimaging procedures provide an opportunity to find the neural correlates of phenomenal consciousness in these patients.

The problem is that there is no validated objective "consciousness meter" that can be used as a proof or disproof of consciousness in such patients. Passive paradigms such as auditory and pain perception showed that brain activations were significantly different between VS and MCS (for a review, see Giacino et al. 2006). But in the absence of a thorough understanding of the neural correlates of consciousness, brain activations observed with passive paradigms are not sufficient to know whether or not these patients are phenomenally conscious. Brain activations observed by using passive paradigms could reflect consciousness but they could also simply reflect nonconscious processing (see

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

studies on subliminal priming or nonconscious processing during sleep and anesthesia). Active paradigms may provide a means for detecting phenomenal consciousness in brain-damaged patients (Boly et al. 2007). Owen et al. (2006) have recently used such paradigms by asking patients to actively perform mental imagery tasks. In one exceptional VS patient studied five months after trauma, activation was observed in the supplementary motor area after the patient was asked to imagine playing tennis; and in premotor, parahippocampal, and posterior parietal cortices when asked to imagine visiting her house. Identical activation was observed in healthy volunteers. The aforementioned patient's neural responses by imagining tasks when asked to do so confirmed that she was phenomenally conscious (Boly et al. 2007; Owen et al. 2006; 2007).

We concur with Block when he claims that conscious states would not magically disappear in a person who is unable to report conscious states. In this context we emphasize the complete locked-in syndrome (LIS). Classical LIS is defined by sustained eye opening, aphonia, quadriplegia, vertical or lateral eye movement or blinking of the upper eyelid to signal yes/no responses, and preserved consciousness (American Congress of Rehabilitation Medicine 1995). Complete LIS consists of total immobility including all eye movements (Bauer et al. 1979). In approximately half of LIS patients the diagnosis initially is missed (Laureys et al. 2005). Julia Tavalaro, a LIS patient whose diagnosis was missed for 6 years, illustrates that phenomenal consciousness can exist for many years in the absence of reportability. Although Tavalaro was actually fully conscious, she was called "the vegetable" – she documents this horrible experience in her book, *Look Up for Yes* (Tavalaro & Tayson 1997). Another testimony comes from an unpublished case we witnessed. This patient was considered as being vegetative for 15 years after a traumatic brain injury, although he was conscious. Now he can talk; and, using a facilitated communication device, he talks of his despair during these years of isolation.

Brain computer interfaces (BCI), also named "thought translation devices," have shown their utility to document consciousness in LIS. BCI is a communication system in which the messages or commands that an individual sends to the external world do not pass through the brain's normal output pathways of peripheral nerves and muscles (Kubler & Neumann 2005). In end-stage amyotrophic lateral sclerosis (ALS), near-complete LIS patients were able to communicate without any verbal or motor report, but were able to do so solely by modulating their EEG (Birbaumer et al. 1999; Hinterberger et al. 2005). BCI can use surface electrodes but is faster when intracortical electrodes are used (Hochberg et al. 2006). Finally, mental manipulation of salivary pH has been used as a form of non-motor mediated communication in one complete LIS ALS patient (Wilhelm et al. 2006). Figure 1 graphically illustrates such "Yes" (imagine lemon)/"No" (imagine milk) communication we obtained from a healthy volunteer.

Bedside evaluation of consciousness in severely brain-damaged patients who cannot verbally or behaviorally report their putative phenomenal consciousness is intrinsically difficult. New functional neuroimaging techniques employing "active" mental imagery paradigms have shown their interest in the assessment of VS and MCS. The rare but horrifying condition of complete LIS illustrates that phenomenal consciousness may remain present for many years in the absence of any overt reportability.

ACKNOWLEDGENTS

Steven Laureys is research associate at the Belgian Fonds National de la Recherche Scientifique (FNRS). Audrey Vanhaudenhuyse is funded by a grant from the French Speaking Community of Belgium (Concerted Research Action 06/11-340).

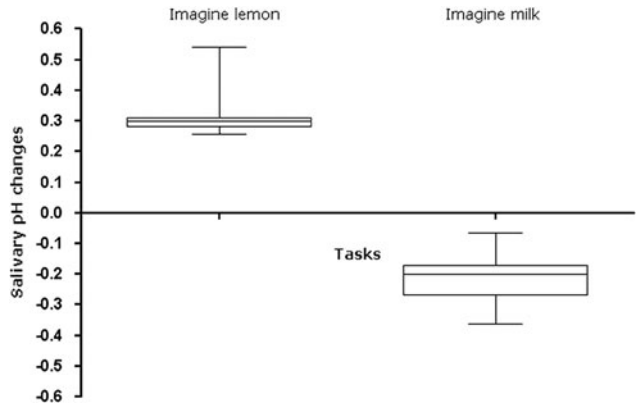


Figure 1 (Vanhaudenhuyse et al.). Nonverbal report of phenomenal consciousness using salivary pH changes. A healthy subject communicated "Yes" (i.e., imagine lemon) or "No" (i.e., imagine milk) while salivary pH changes were monitored (as compared with baseline). Box (mean and standard deviation) and whickers (minima and maxima) obtained during a single untrained session; 2 minutes of imagery for each task.

Author's Response

Overflow, access, and attention

doi: 10.1017/S0140525X07003111

Ned Block

Department of Philosophy, New York University, New York, NY 10003.
ned.block@nyu.edu

Abstract: In this response to 32 commentators, I start by clarifying the overflow argument. I explain why the distinction between generic and specific phenomenology is important and why we are justified in acknowledging specific phenomenology in the overflow experiments. Other issues discussed are the relations among report, cognitive access, and attention; panpsychic disaster; the mesh between psychology and neuroscience; and whether consciousness exists.

R1. Introduction

I have learned a great deal from reading the commentaries and I am gratified that so many respondents are sympathetic to separating phenomenal consciousness from cognitive access to it – a stark contrast to the responses to an earlier BBS paper (Block 1995), in which I argued for similar though slightly stronger views. I don't flatter myself with the supposition that I have convinced anyone – the main factor is that a wide range of accumulating evidence increasingly supports separating phenomenal consciousness and cognitive access. (Evidence beyond what is described in the target article is mentioned in **Gopnik**; **Izard, Quinn, & Most** [**Izard et al.**]; **Koch & Tsuchiya**; **Lamme**; **Landman & Sligte; Malach;** and **Snodgrass & Lepisto**.)

The empirical core of my argument in the target article concerned what I called overflow, that is, that the capacity of the phenomenal system is higher than the capacity of the cognitive access system that underlies reportability of phenomenal states. Many of the respondents (**Burge**;

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

Byrne, Hilbert, & Siegel [**Byrne et al.**]; **Grush**; **Jacob**; **Kouider, de Gardelle, & Dupoux** [**Kouider et al.**]; **Landman & Sligte**; **O'Regan & Myin**; **Naccache & Dehaene**; **Papineau**; **Spener**; **Van Gulick**) have commented on that argument, and my discussion of their responses is in section R2. There are two main issues, one concerning the distinction between generic and specific phenomenology (sect. R2.2) and the other concerning hyperillusions (sect. R2.4). If the overflow argument is correct, the cognitive system underlying the reporting of phenomenology is distinct from the system underlying phenomenology itself, a matter that raises the issue of the role of reports in an empirical investigation of consciousness. I argued that reports are the starting point but that they can be rejected on the basis of the method of "inference to the best explanation." Naccache & Dehaene, **Lau & Persaud**, Papineau, **Prinz**, and **Sergent & Rees** criticize this reasoning. My response – in section R3 – makes use of contributions by **Koch & Tsuchiya**; **Malach**; **Snodgrass & Lepisto**; **and Vanhaudenhuyse, Bruno, Brédart, Plenevaux, & Laureys** [**Vanhaudenhuyse et al.**]. The overflow argument presupposes a view of the cognitive access system that has drawn many comments both for and against (**Balog**; **Clark & Kiverstein**; **Gopnik**; **Harman**; **Kentridge**; **Lau & Persaud**; **Izard et al.**; **Levine**; **Malach**; **Rosenthal**; **Shanahan & Baars**; Snodgrass & Lepisto; **Tye**), and is discussed in section R4. Section R5 concerns the comparison between consciousness and attention in relation to comments by **Lycan**, Koch & Tsuchiya, and Prinz. In the target article I gave another empirical argument in addition to the overflow argument, one based on the claim that a theory that explains the mesh between psychology and neuroscience is to be preferred, other things being equal, to a theory that does not so explain. Commentators Burge, **Grush**, **Lamme**, Landman & Sligte, **Hulme & Whitely**, Prinz, Snodgrass & Lepisto, and Van Gulick discuss this argument regarding the mesh and I reply to them in section R6. Finally, **McDermott** questions whether consciousness even exists, and that issue is discussed in section R7.

R2. The overflow argument

R2.1. Introduction

I appealed to the Sperling (1960), Landman et al. (2003), and Sligte et al. (2008) experiments in the overflow argument. In the Sperling experiment, subjects exposed to a short initial stimulus have the impression that they are aware of up to 12 specific letter-shapes but can report only about 3 or 4 letters. However, if a specific row is cued, subjects can report any 3 or 4 letters in that row, confirming their initial impression. Landman et al. and Sligte et al. get similar results using a paradigm in which subjects give a much more minimal response – with regard to whether a specific cued item has changed orientation – thus minimizing interference between the phenomenal representation and the subjects' own response. (See also the Landman & Sligte commentary.) The overflow argument says that the capacity of phenomenology is much greater than the working memory system underlying report, so they must to some extent be distinct systems.

R2.2. Generic/specific phenomenology

The distinction between generic and specific phenomenology was crucial in the overflow argument, although I did not give it a name or draw enough attention to it (I did speak of experiences "as of . . . specific shapes.") I will explain it here by example. In the Landman et al. experiment, the relevant generic phenomenology would be the phenomenal presentation *that there is* a circle of rectangles. The relevant specific phenomenology would be a phenomenal presentation that specifies for each of the rectangles (or anyway, most of them) whether they are horizontal or vertical. For the Sperling experiment, the relevant generic/specific difference would be that between a phenomenal presentation *that there is* an array of alphanumeric characters and a phenomenal presentation of specific shapes of all or most items in the array. I argued that in the Sperling, Landman et al., and Sligte et al. experiments, there was specific phenomenology involving all or almost all of the items, as well as generic phenomenology. This distinction figures – in those terms – in the commentaries by **Burge**, **Grush**, and **Levine**. **Kouider et al.** express it as the lower/higher distinction, **Papineau** as the scene/item distinction, **Sergent & Rees** as the scene gist/detail distinction, and in the target article as well as in the **Byrne et al.**, **Naccache & Dehaene**, **Jacob**, **Spener**, and **Van Gulick** commentaries, the distinction is deployed without special terminology.

My argument was that before the cue, there is specific phenomenology for all or almost all items (and also generic phenomenology, though that does not figure in the argument). This specific phenomenology is what justifies the claim that the capacity of the phenomenal system is more than 4, whereas the capacity of the access system is 4 or less and thus that the two systems cannot completely coincide. A number of the critiques (most pointedly, **Papineau**'s and **Byrne et al.**'s) challenge the premise that there are more than 4 items of specific phenomenology before the cue. It is important to recognize that the objectors have to agree that before the cue, there are *specific* (not just generic) visual representations of all or almost all of the 8 to 32 items in the Sperling, Landman et al., and Sligte et al. experiments. There have to be such specific representations given that any location can be cued with high accuracy of response. The locus of controversy is *whether those specific representations are phenomenal*.

Here is my evidence for claiming that there is specific phenomenology for all or almost all the items in the overflow experiment:

1. As **Burge** notes, subjects (including myself) in overflow experiments often testify that their responses are based on specific phenomenology that was there all along. (Rogier Landman tells me that the extent to which subjects evince specific phenomenology may be correlated with how well they do in the experiments.) **Shanahan & Baars** say "It is worth noting that subjects in such experiments believe they are simply reporting the contents of their recent visual consciousness, even after the visual stimulus has disappeared from view." In the Sligte et al. experiment, subjects have up to 4 seconds before the cue comes on, so their judgments about their phenomenology are not rushed.

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

2. Subjects are attending to arrays in full view, in good viewing conditions, for half a second in the Landman et al., and some versions of the Sperling experiments, and one second in the Sligte et al. experiment, more than enough time for specific phenomenology. (**Burge** also makes this point.)

3. If there is only generic phenomenology before the cue, and if the cue causes the generic phenomenology to be replaced by specific phenomenology, then there is a shift from generic to specific phenomenology. The fact that subjects report no such phenomenological shift might not be strong evidence against this view, but it is some evidence. The vast literature on this topic (including two Ph.D. theses I have read) contains no mention of such a thing as far as I know. I myself can testify that even looking for such a shift, one does not experience it. The point of the target article's example of the rectangle coming into view as if from a distance was to compare that phenomenal shift with the absence of such a shift in the overflow experiments. If there are some items of specific phenomenology before the cue, and different items of specific phenomenology after the cue, then one might expect that to be noticed as well.

4. There is evidence mentioned in the target article that cortical persistence obtains at all levels of the visual system and therefore at the phenomenal level as well. In particular, there is evidence (mentioned in the target article) that the persistence exists at levels where depth and motion are represented. As **Lamme** and **Landman & Sligte** note, persisting representations obtain at a stage of visual processing past figure-ground segregation and feature binding, properties that "are more associated with conscious processes" (Landman & Sligte's commentary, para. 5). Hence, there is a neural case for phenomenal persistence.

5. In the target article, I mentioned Di Lollo's paradigm using a 5 by 5 grid in which all but one of the squares is filled with a dot. Subjects see a partial grid with 12 of the dots filled in, then, after a delay, another partial grid with a different 12 dots filled in. The subjects' task is to report which square has a missing dot, something they can do easily if they have a visual impression as of the whole matrix of dots. Loftus and Irwin (1998) show that subjects' ability to perform the task correlates nearly perfectly with their phenomenological judgments of whether there appears to be a whole matrix rather than two partial matrices. When writing the target article, I did not know about the variant by Brockmole et al. (2002) in which the appearance of the second partial grid was delayed by as long as 5 seconds and in which subjects were told that a good strategy was to "imagine the dots still being present after they disappeared" (p. 317). The subjects' memory capacity for the 12 dots in the first grid can be computed by the type of errors made. When the delay between the first and second partial grids is 100 msecs, the subjects' retention capacity falls from 12 to 4.1 of the 12 dots in the first partial grid. The striking result was that with delays over 100 msecs, subjects' capacity *increased*, asymptoting at a delay of about 1.5 seconds, at which time their capacity was 10 of 12 dots, and the capacity stayed that high for delays up to 4–5 seconds. Independent estimates of the time to generate a mental image (Kosslyn et al. 2006) are between 1 and 2 seconds, and the authors argue that the subjects were following instructions, generating a visual image of the first array, and integrating that visual image with the percept of the second array. This result constitutes converging evidence for high capacity specific phenomenology: Since the subjects could do the task well, the Loftus and Irwin result suggests they have a visual impression of the whole matrix; and in any case visual imagery is phenomenal. So the representations are phenomenal, and the capacity of 12 dots is substantially more than 4 items. The upshot is that there is a completely different paradigm in which the evidence favors high capacity specific phenomenology.

6. Finally, there is the evidence mentioned by **Kouider et al.**, which I discuss immediately below.

Kouider et al. suggest what they take to be an *alternative* to what I am suggesting but what I take to be a *version* of it. They hypothesize that Sperling-like paradigms result from "partial awareness: subjects have a transient and degraded access to fragments of all the letters in the grid." Kouider et al. take comfort from a common observation that subjects in experiments involving masked stimuli often report seeing bits and pieces of stimuli. What are Kouider et al. saying about specific phenomenology? No specific phenomenology at all is not compatible with their view, since they say subjects are to some degree conscious of and have access to "fragments of all the letters in the grid." And these commentators clearly don't envision full specific phenomenology, so they must envision *partial specific phenomenology*. Their evidence involves cases (Kouider & Dupoux 2004) of brief masked presentations of color words and pseudo-color words (e.g., "green" vs. "gener") in which both have the same effect on subsequent identification of colored stimuli, facilitation in the case of congruent stimuli (e.g., "green" or "gener" followed by green) and inhibition in the case of incongruent stimuli ("green" or "gener" followed by red). Kovider et al. predict, plausibly, that in Sperling experiments that include some letter-like symbols which are not letters, subjects would treat false letters as similar to real letters. More generally, Kouider and Dupoux (2004; 2007) give evidence that genuine semantic priming requires at least partial phenomenology and that totally unconscious stimuli can have only non-semantic effects. (But see Abrams & Grinspan 2007.) To the extent that I disagree with Kouider et. al., that disagreement is on just *how* degraded the specific phenomenology is. One of the experiments in Sligte et al. (2008) used bars that differed in increments of 45 degrees, and subjects still showed high capacities – slightly under 8 for stimuli of 16 bars with the cues presented at 1,000 msecs after the stimulus offset. The specific phenomenology is good enough to make those distinctions between say 135 degrees and 180 degrees pretty well. In the Sperling experiment, the specific phenomenology was good enough for subjects to distinguish among the 26 letters of the alphabet. No doubt if there had been pseudo-letters, subjects would have made more errors, but distinguishing among 26 alternatives is still respectable. It is worth noting that in the experiment by Kouider and Dupoux, stimuli were presented at 29 msecs or 43 msecs and also masked, an intervention aimed at making them harder to see. Landman et al.'s stimuli were normally unmasked and presented for 500 msecs, Sligte et al.'s for 1,000 msecs (and Sperling's stimuli yield the same

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

results with a 500 msec presentation – cf. also **Burge**); so the fragmentariness of the phenomenology in Kouider and Dupoux (2004) could be predicted to be greater than in the overflow experiments. The upshot is that Kouider et al. have presented further evidence for high-capacity specific phenomenology, just what my overflow argument relies on.

Byrne et al. are not impressed with my appeal to subjects' testimony. They raise a plausible objection, that the generic/specific distinction is pretty abstruse, so how can the responses of subjects who don't know the distinction provide support for specific phenomenology? However, when subjects say (cf. **Shanahan & Baars**) that in reporting the letters in the cued row or in telling whether the cued rectangle changed orientation, they are simply reading their answers off of the visual impression that was in existence before the cue, they are evincing specific phenomenology regardless of whether or not they could state the distinction between specific and generic phenomenology.

Readers may be thinking that subjects may have seen the orientations/identities of a few of the items, but that subjects cannot be sure that they saw all or most of them. Recall that *specific representations of all or almost all the items before the cue (though perhaps fragmentary representations) have to be postulated* to explain the fact that subjects can report the items no matter which row is cued. So the options would appear to be either that there was no specific phenomenology before the cue, or that there was specific phenomenology involving all or most of the items, even if fragmentarily. As I just mentioned, subjects' testimony (and lack of surprise in what they can do) suggest the latter.

Papineau hypothesizes that the presence of generic phenomenology is my basic reason for postulating specific phenomenology. I should have drawn more attention to my reasons, as I have now done.

Van Gulick notes that the "movie screen of the mind" view would say that you cannot have generic phenomenology without specific phenomenology, implicitly suggesting that I am relying on the "movie screen of the mind" view, and on the fact of generic phenomenology, to argue for specific phenomenology. (**Grush** uses what would appear to be the movie screen view of pictorial representation, but the other way around from what Van Gulick ascribes to me: Grush says generic phenomenology cannot be an image because if it were an image, clear contents would have to present the details.) I reject the principle – applied by both Grush and Van Gulick – that pictorial representation has to specify the relevant details. I call this principle the "photographic fallacy" (Block 1983). More specifically, the photographic fallacy supposes that pictorial representations have to represent details of anything in view in the manner of a prototypical photograph. To see the fallacy, note that an impressionist painter might represent a hand in broad brush strokes that do not explicitly represent the number of fingers or whether one of them has a ring.

Van Gulick also argues that phenomenology in the overflow cases may be partial, weak, or somewhat indeterminate, but that access may also be limited in the same way, so there is no evidence for a discrepancy. With regard to the issue of whether degree of phenomenology is matched by degree of access, there is experimental evidence to the contrary mentioned by **Sergent & Rees** (e.g., Sergent & Dehaene 2004). Cognitive access appears to be more of a binary phenomenon than Van Gulick supposes.

Spener argues that the fine details of introspective grasp of specific phenomenology are thoroughly expectation-driven. In support, she mentions the disagreements in the philosophical literature about whether or not there is a determinate number of phenomenologically represented speckles on the speckled hen.

I think **Spener** exaggerates the problem. Phenomenological disagreement can be to some extent settled. The speckled hen case, in particular, has been illuminated by work on the different "grain" of vision as opposed to attention. Patrick Cavanagh and colleagues (1999) have shown that the grain of vision is about one-sixtieth of a degree at the fovea. For example, in order for a grating to be distinguishable from a gray field, the individual lines have to subtend more than one-sixtieth of a degree. However, for a subject to *attend* to visible lines and to be able to *move* attention from one to another, the lines must subtend at least five- to ten-sixtieths of a degree. How is this relevant to representations of speckles? Plausibly, the phenomenological disagreement about represented speckles stems from conflating seeing with attending. The speckles in the standard example in which they subtend more than one-sixtieth of a degree, but less than five- to ten-sixtieths of a degree, are visible but not attendable, so that one can see them (and thus phenomenally represent them) but not count them or do anything else that requires moving one's attention from one to another. If one is looking right at the speckles for sufficient time, the phenomenology really does determinately represent many of the individual speckles – specific phenomenology as opposed to generic phenomenology – but the speckles have an "elusive" quality because one cannot attend to them.

Grush argues for an illusion in which generic phenomenology presents an object as affording answers to certain queries, so when the answers to the queries are filled in, the subject does not notice. One might call this putative illusion the *affordance* illusion. (Grush gives it a less mnemonic name.) The affordance illusion is supposed to explain change "blindness" and inattentional "blindness" in terms of sparse generic representations. Of course the same idea, if it worked, could be used to undermine my reason (3) for specific phenomenology. The one argument for this illusion that Grush offers is that if there were an affordance illusion, it would serve to explain change "blindness" in terms of sparse representations. But as I noted in the target article, the Landman et al. and Sligte et al. experiments are themselves evidence against the "sparse representations" view of change "blindness."

Jacob raises the issue of whether what I say in the target article is compatible with what I have said against representationism. I say yes, because the distinction between generic and specific phenomenology requires only that in the cases in question, phenomenology *has* representational content of the appropriate kinds; and that is compatible with the view (which I hold) that there is more to phenomenology than representational content.

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

R2.3. Unconscious representation before the cue

So far, I have been talking about the objection that before the cue, there is only generic phenomenology, and no specific phenomenology. However, there is a more radical view that says that until the cue appears, there is no phenomenological representation of the array at all, either generic or specific. Dehaene and his colleagues (2001; 2006) were arguably committed to this more radical view by the refrigerator-light illusion hypothesis, as emphasized by **Burge**, but **Naccache** & **Dehaene** in their commentary advocate the less radical thesis. The only advocate of the more radical thesis among the commentators is **O'Regan** & **Myin**. My main response to the more radical thesis is that the arguments I have given for specific phenomenology before the cue constitute a case for *some* phenomenology before the cue.

O'Regan & Myin note, correctly, that I am assuming that consciousness is a natural kind and has some sort of neural signature. They claim this hypothesis is speculative and that even if true, it would not solve the "hard problem." One of the points of my target article is that we can discover a lot about the neural basis of consciousness, short of solving the hard problem. On the issue of speculativeness: I am not assuming there is something physical in common to all possible cases of consciousness (e.g., including mammals, birds, octopi, conscious machines, and conscious extra-terrestrials, if there are any). My assumption is that there is a neural signature of consciousness in humans that is shared at least by other mammals with similar sensory systems. This is an assumption that I believe is shared by the field and looks promising so far. O'Regan & Myin note that I claim that phenomenal consciousness has effects on the basis of which we can find evidence about its nature. They interpret my earlier paper in this journal (Block 1995b) as arguing for the opposite view, epiphenomenalism. Although Block (1995b) took epiphenomenalism seriously, the upshot was (and is) that we have reason to reject it.

R2.4. Hyperillusions

Naccache & Dehaene appear to agree with me (and disagree with **Byrne et al.**) that subjects think they have specific phenomenology of all or almost all the items in the overflow experiments; but Naccache & Dehaene think subjects are suffering from what I called a *hyper-illusion.* In ordinary illusions, appearance misleads a subject about reality, but in hyper-illusions, appearance of appearance misleads about appearance. (This may sound glib and meaningless – and perhaps it is – but I think it is fair to describe a putative phenomenon in a way that brings out its peculiarity.) Naccache & Dehaene argue that "we all have the illusion of seeing a world in full color although color-sensitive cones are absent in the periphery of our retina." And **Van Gulick** makes the same claim, citing Dennett's (1991) color marker demonstration: we hold a colored marker to the side of our visual field, bringing it gradually into the fovea, where we supposedly begin to see its color. However, the claim of illusion on the basis of retinal distribution of cones is analogous to (though not quite as obviously wrong as, for reasons I will get to) the idea that there is an illusion involved in seeing the world right-side-up, given that our retina contains upside-down images. What is on the retina is relevant to phenomenal experience *only to the extent that it affects the parts of the brain that determine phenomenology*. Assuming for the sake of the example that activity in V4 is the neural basis of color phenomenology, then the issue of whether the world genuinely appears colored in the periphery would be illuminated by knowing whether V4 has receptive fields in the periphery (that is, whether V4 contains groups of neurons that respond to and represent areas in the periphery), not by knowing what is on the retina.

The two putative illusions (the putative color illusion and the putative upside-down illusion) differ because there is a genuine issue: Does V4 have receptive fields in the periphery? And if it does, what is the purpose of those fields, given low color sensitivity in the retinal periphery? An obvious answer is: Because visual phenomenology depends on integrating information over time. Our visual representations comprising the whole visual field are not built up in an instant! So color representation of something now in the periphery can derive from its representation recently in more central vision. Striking evidence for integration over time in some aspects of perception is provided by evidence that multicolor representations are built up serially – the system processes one color at a time – although locations and shapes are processed in parallel (Huang & Pashler 2007; Huang et al. 2007). There is another response to the question of why cortical color representations would represent color in the periphery: It turns out that hue discrimination at 50 degrees of eccentricity is *as good as in the fovea* (which subtends only 2 degrees – about twice the width of the thumbnail at arms length) if the size of the stimulus is magnified, and there is even some color sensitivity at 80 to 90 degrees (Mullen 1992).

Naccache & Dehaene mention another alleged hyperillusion, the "moving window" experiment by McConkie and colleagues (1975; 1979), in which subjects suppose there is a full page of normal text even though the text outside the small moving window is degraded or changed. The point just made about representations integrated over time applies here too.

R2.5. Phenomenal memory

Jacob notes that my talk of phenomenal memory is misleading. What I meant was that there is neural persistence at all levels of the visual system, including the level that determines phenomenology. So instead of talking about phenomenal memory, I should have just talked about phenomenal persistence. **Byrne et al.** say I don't need high-capacity phenomenal persistence but only high-capacity phenomenology. The role of phenomenal persistence in my target article was in connection with the "looking again" dialectic that I presented. The argument for the high capacity of the phenomenal system can be undermined by the idea that the high capacity derives from the information in the environment or the high capacity of low-level vision, for example, the retina. Tracking the sources of phenomenal persistence allows one to rule out those alternatives. Recall that I argued that phenomenal persistence is at a level that involves binocularity (Engel 1970) and motion (Treisman 1975), is not

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

disturbed by a light mask that disturbs positive afterimages (Sligte et al. 2008), and is disturbed by a pattern mask that does not disturb retinal afterimages (Sligte et al. 2008). All these data provide reason for thinking that the phenomenal persistence is due to a persistence in the underlying basis of phenomenology that is not being driven by an earlier persistence. For these reasons, iconic memory is, as **Landman & Sligte** note, a window into phenomenal consciousness.

R3. Reports

The overflow argument says that the cognitive system underlying reporting differs from the phenomenal system. This raises the issue of the extent to which reports are pipelines to phenomenology.

R3.1. Are reports privileged?

Naccache & Dehaene pose a dilemma for me: "If one wants to define phenomenal consciousness differently from conscious reportability, then one should resist the temptation to make use of subjects' reports" as evidence for phenomenal consciousness. However, no one would think there is such a dilemma for, say, acidity or heat. An acid is a proton donor and heat is molecular kinetic energy. These are good scientific definitions, but no one thinks that these definitions preclude any kind of evidence. Naccache & Dehaene see inconsistency looming: How can I rely on reports in the Sperling experiment while at the same time claiming that subjects' reports that they don't see something can be wrong? **Prinz** maintains that reports are "authoritative."

Vanhaudenhuyse et al. exhibit the ingenuity of researchers in finding ways that subjects who have no control of motor and verbal systems – and hence no capacity to make a motor or verbal report – can nonetheless report conscious states by, for example, thinking of milk rather than lemon. This point does not, however, challenge Dehaene & Naccache, since it does not show any daylight between consciousness and *possibility* of reporting it. My view is that reports are the *starting point* for building a theory of phenomenal consciousness, but can be rejected if the best explanatory theory requires it. The arguments for this view include the methodological points about inference to the best explanation, the sketch of an actual explanatory account on which reports can be wrong, and the points made forcefully by **Snodgrass & Lepisto** (and in Block 2005) involving signal detection theory.

It is obvious that reports fail to be authoritative in that we can have conclusive evidence against the truth of introspective reports. As **Koch & Tsuchiya** note, in Anton's syndrome, subjects are blind but think and report that they see. More generally, anosognosics deny their perceptual and motor disabilities, making all sorts of false reports about their own experience. Introspective reports do have a certain priority: we have no choice but to start with reports in investigating consciousness. I am in complete agreement with Koch & Tsuchiya when they say, "in the absence of compelling, empirical evidence to the contrary... if the subject denies any phenomenal experience, this should be accepted." One very notable form of empirical evidence that can conflict with report is evidence about subjects' decision processes evaluated according to signal detection theory. As **Snodgrass & Lepisto** note, "contrary to many researchers' implicit assumptions, there is no such thing as an unmediated 'subjective report' – ever."

In practice, **Naccache & Dehaene** use the methodology I advocate, not the one they themselves advocate. They concede that reports can be partial and biased and argue that "reports of a rich phenomenality cannot be taken at face value," citing the moving-window paradigm and the claim that we have the illusion of seeing a world in full color. Further, both Naccache & Dehaene and **Sergent & Rees** suppose that reports have to be measured using high-tech machinery, leading to a concern about what they could mean by *report.* Is a report just (1) *any* evidence of consciousness? Or is it (2) evidence of consciousness from a subject's performing one voluntary action rather than another? Or is it (3) evidence of global broadcasting? If (1), these commentators' focus on report as the gold standard for evidence of consciousness is trivial, because it just means that evidence is the gold standard of evidence. If (3), their claim that the evidence for consciousness is always evidence for global broadcasting is trivial, amounting to the claim that evidence for global broadcasting is evidence for global broadcasting. If (2), the claim that reports are the gold standard is substantive but probably false. As **Malach** has shown in a previous study (see Hasson et al. 2004), you can get evidence about consciousness from people just watching a movie in a scanner and not making any voluntary response. In another such case mentioned by **Sergent & Rees,** Lumer and Rees (1999) put subjects in a binocular rivalry experiment without requiring any voluntary response and found alternation between neural activations, providing evidence of the neural bases of different perceptual experiences without a voluntary response. An advantage of sticking to normal uses of words is that we don't have to *guess* what people mean.

R3.2. Panpsychic disaster

A number of commentators argue that once you give up the special authority of reports, you will have no way of avoiding attributing consciousness to lampposts. **Papineau** notes that I regard some states as uncontroversially unconscious and wonders, "what makes a state 'uncontroversially unconscious' if it is not that subjects tell us so." He argues that once we allow that a state can be conscious even though normal subjects systematically deny it, there may be no uncontroversially unconscious states. **Prinz** says: "Block must either concede that reports are authoritative or deny that we can rule out the possibility of conscious states in V1, the LGN, and the retinae." **Lau & Persaud** say that the methodology I rely on to support recurrent loops as necessary assumes that subjects are to be believed when they report that they don't see something. But given that I allow for phenomenology without access, Lau & Persaud wonder, how can I rule out that these subjects have phenomenology despite reporting that they don't? **Sergent & Rees** and **Naccache & Dehaene** would no doubt agree. But these critiques ignore the obvious failure of authoritativeness of reports just mentioned and the power of inference to the best

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

explanation to sort the good from the bad attributions of consciousness.

Perhaps what **Naccache & Dehaene** and **Sergent & Rees** are supposing is that if biomarkers of consciousness conflict with actual verbal report, verbal report wins. But there is no evidence that this *always wins*. If items of evidence conflict, the right scientific procedure is to find an explanation of the conflict, doing further experiments if necessary to see which item of evidence is misleading.

R4. Access

Thus far, I have talked about the overflow argument (sect. R2) and the methodology of reports (sect. R3). I now turn to issues concerning the cognitive accessibility system.

R4.1. Awareness, access, and agency

Levine notes that the idea of a phenomenal consciousness that has nothing to do with any kind of access-relation to the subject doesn't really seem like any kind of consciousness at all. A major difference between my position in the target article and Block (1995b) is that in the target article I acknowledge this point and accommodate it within a framework that avoids any constitutive connection between that access-relation (which I called *Awareness* in the target article and Levine calls *awareness-access*) and *cognitive* access, which I identify in the target article with broadcasting in the global workspace. **Jacob** wonders why I have changed my mind, noting that the kind of awareness I now say is a necessary part of conscious experience is not full-fledged self-awareness of the kind a rational thinking creature might sometimes have but that a mouse will presumably not have. In 1995, the only option I saw for explaining awareness-access in non-cognitive terms was as a kind of phenomenal property I called "me-ishness." But now I see that awareness-access can be adequately understood in terms of "same-order" and deflationary theories, so there is no need for cognitive or other "higher-order" accounts.

Levine makes a plausible case that the resistance many feel to the idea that the machinery of phenomenal consciousness is separable from the machinery underlying report (i.e., broadcasting in the global workspace) stems from conflating broadcasting-access with awareness-access; and he mentions two interesting suggestions for ways in which the relation between awareness-access and the self might be further elucidated. One of them is that there is no reason why subjectivity cannot involve a fragmented subject. Yes, but the subject cannot be *so* fragmented that the experience is not for-the-subject. If G.K. has the face experience on the left that he denies, what keeps him from acknowledging it is his inability to attend to the left side of space. But that is compatible with the face experience on the left being part of the *same visual field* as whatever he is seeing on the right. (I use this term, "visual field", as is standard in philosophy to mean the entirety of visual space, including left and right.) That is, the relations in the visual field would be normal and have normal awareness-access, but the subject would have poor broadcast-access to them. So we can make sense of the idea that experience of the face on the left is an experience *for him* despite some fragmentation of the self.

Levine's hypothesis illuminates **Rosenthal**'s critique, in which Rosenthal asks: If phenomenology necessarily involves awareness as I say it does, and given that awareness normally results in cognitive access, why would cognitive access fail in the G.K. case? Rosenthal is assimilating or at least supposing too tight a connection between awareness-access and broadcasting access. G.K. has awareness-access on the left if he has phenomenology on the left, but it is his lack of broadcasting-access that explains why he cannot report what is on the left. Rosenthal claims that higher-order thought theories of consciousness (known as HOT theories) explain why ordinary phenomenology always involves some awareness of it, whereas the global-workspace theories I favor as theories of access do not. But the global workspace theories are theories of *cognitive* access, not theories of *awareness*-access. Again, **Levine**'s hypothesis is confirmed. The competition to higher-order theories as theories of awareness-access is offered by the same-order theory and the deflationary theory, not by the global workspace theory.

Rosenthal suggests that infant pain might have "weak" phenomenology and that weak phenomenology may be what I am supposing G.K.'s face experience has as well. I am not sure what "weak" phenomenology is supposed to be, but I guess it stands for the phenomenology you get when there is no global broadcasting and no higher-order thought. There is no reason to suppose that infant pain or G.K.'s face experience is any less *intense* than adult pain (cf. **Gopnik** and **Izard et al.**). Gopnik adds another layer of evidence to what I presented in the target article about infant phenomenology and **Malach** adds yet another layer relevant both to infant and adult phenomenology. Endogenous attention matures late, making it even less likely that infants are capable of higher-order thought. Inhibitory neurotransmitters, a major contributor to unconscious states, mature late, making it less likely that infants' perceptual states are unconscious. Younger infants are slower to habituate, another source of unconscious perceptual states (although this may be because younger infants are slower to encode stimuli rather than anything to do with the machinery of habituation). Evidence against *both* higher-order thought *and* unconscious states puts Rosenthal in a vise, since these states exhaust his options. Izard et al. also give suggestive evidence that emotion and action systems reveal phenomenal states that may not involve either global broadcasting or higher-order thought. The upshot is to increase the empirical squeeze on higher-order thought theories of consciousness, thereby increasing the attractiveness of same-order theories.

Rosenthal says standard same-order views are just as cognitive as HOT. Rather than argue about the texts he cites, let me just say that Caston's influential paper on Aristotle's same-order theory (Caston 2002) emphasizes some of the advantages of the same-order view over higher order theories that I mentioned in the target article, namely, that there is no need to postulate that whenever there is a token conscious pain, there is also a token thought about it; and that there is no puzzle about why my thought about a teacup or your pain does not make the teacup or the pain conscious, whereas a higher-order thought about my own pain does. (See also **Malach**'s Figure 2, part B, which depicts a way of thinking about the same-order account.)

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

Astonishingly, **Rosenthal** claims in regard to infant pain that, "because nonconscious pains have largely the same causal connections to behavior and to a strong desire for pain to cease cease, they are just about as bad." It is well known that pain in infants has bad effects, but I have never heard of any experimental evidence to the effect that those bad effects do not depend on whether the pain is conscious or not, nor does Rosenthal cite any. Newborns who are circumcised without anesthesia or analgesia are more stressed by vaccination even 6 months afterward (Taddio et al. 1995; Qiu 2006). But this evidence does not approach the issue of whether the infant pain is conscious. I imagine that Rosenthal is taking the evidence that higher-order thought is unlikely in infants, and the evidence that pain in infants has bad effects, to come to the conclusion that unconscious pain in infants has bad effects. But this response suggests that Rosenthal is retreating from a substantive empirical claim to a trivial linguistic claim. Rosenthal can, if he likes, simply use the word "*unconscious*" to mean something on the order of *unaccompanied by higher-order thought.* But pains that are unconscious in that trivial sense cannot be supposed to be pains such that there is nothing it is like to have them. The danger of promoting such a triviality into a substantive thesis can be seen in Peter Carruthers' infamous claim (Carruthers 1989; 1992) that, given that pains in dogs, cats, sheep, cattle, pigs, and chickens are not available to be thought about, these pains are not felt and hence are not appropriate objects of sympathy or concern and are of no moral significance. (Carruthers [1999] backpedals, not on the ground that animal pains are appropriate objects of concern or sympathy, but on the ground that the frustration of animal desires are of moral significance.) Drug testers and slaughterhouses take notice! Given the evidence presented in the target article and in **Gopnik** and in **Izard et al.**, I really don't see how advocates of this unattractive view can avoid applying it to human infants.

Snodgrass & Lepisto give a very plausible argument concerning Jacoby-style "exclusion" experiments (Debner & Jacoby 1994) – that they involve phenomenally conscious stimuli or parts of stimuli which the subject has little confidence in having seen, and hence no higher-order belief in having seen it. (There was also an argument to this effect in Block 2001.) For example, the subject is instructed to complete the stem "rea__" with an ending that is not a word that might have just been flashed briefly. If the word is "reason" and the subject saw all or part of it, but has low confidence, the subject is more likely to complete "rea__" with "son" than if no word at all was flashed. As noted in Snodgrass (2002), subjects who are penalized for errors do better at excluding, suggesting on a signal detection analysis that they really did see the word (or parts of it, as suggested in **Kouider et al.**) but had low confidence and no higher-order state that would lead to a report. This provides another piece of empirical evidence against the higher-order point of view. Perhaps the advocates of the higher-order point of view will take the degree of confidence to be an index of degree of consciousness, but the signal detection models show that consciously seeing the stimulus and confidence that one has seen it can vary independently, each being influenced by a variety of different factors. (For example, changing the catch rate can influence confidence without influencing perception. See Supèr et al. [2001a] and Block [2005].)

Clark & Kiverstein argue for a view opposed to **Rosenthal** and to **Levine** and me, that for a representation to be phenomenal, it must be "available for use in the planning and selection of deliberate, stored-knowledge–exploiting, and goal-reflecting and goal-responsive, actions." Clark & Kiverstein makes a cogent case, but it doesn't involve any empirical evidence. Rather, it is an a priori conceptual appeal. "To count as a conscious experience, an informational state must" be available for "rational choices and considered actions." One and the same condition of the skin, intrinsically described, will *count* as sunburn if it is caused by the sun, but not so if it is caused by a fire. Similarly, according to Clark & Kiversteins's view, one and the same neural-informational state, intrinsically described, can count as a conscious experience if it is properly connected to rational choices and considered actions, and not so if it isn't. But our ordinary concept of consciousness is too vague to support this kind of claim. Further, one reasonable precisification of our ordinary concept of consciousness will make it the concept of a natural kind. The key feature of that natural kind-property is the way it feels to have it. My own view is that there is an ordinary concept of *phenomenal* consciousness. **Gopnik** gives evidence that infant consciousness may "not be accessible for goal-directed planning nor be the subject of the endogenous attention that accompanies such planning." I don't give this as *evidence* against Clark & Kiverstein's view since conceptual claims are not subject to evidence. Instead, if Gopnik is right, Clark & Kiverstein's conceptual claim would dictate that our concept of consciousness does not apply to infants. But the absurdity of this step strongly indicates that our concept of consciousness – or at least one of our concepts of consciousness – is the concept of a phenomenal feel that has no conceptually necessary relation to agency of the sort Clark & Kiverstein describe.

Malach presents exciting evidence that intense experience such as watching an engrossing movie (a spaghetti western?) activates an "extrinsic" sensory system in the same ways in different people; but it activates an "intrinsic" system based in frontal areas (Hasson et al. 2004) less strongly and not in the same ways in different people. Malach's intrinsic system overlaps considerably with a "default" system that is active when the subject is "doing nothing" and is inactive when the subject is performing intense goal-directed tasks (Raichle et al. 2001). In one study, Malach and colleagues (Goldberg et al. 2006) used the same stimuli (pictures and audio clips) in an introspective task and a difficult rapid categorization task – but with different instructions. In the introspective task, subjects were asked to categorize their emotional reaction as positive, negative, or neutral. In the categorization task, the stimuli were presented at triple the rate and the subjects were asked to categorize the stimuli, for example as animals or not. Subjects also rated their own self-awareness. Their self-awareness was high in the introspective task and low in the rapid categorization task. Part of the intrinsic system – see Malach's Figure 1 – was activated both in the introspective task and other "self"-relevant tasks, including judging the application of emotionally neutral words to oneself. This area is suppressed in the rapid categorization task. The upshot is

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

that the intrinsic system is less active when the subject is doing a demanding perceptual task, or watching Clint Eastwood, and that a part of it that is especially tuned to self-relevant tasks is especially suppressed in difficult perceptual tasks. This study presents a challenge to views such as **Lycan**'s, **Harman**'s, **Rosenthal**'s, and **Van Gulick**'s that emphasize the importance of the self in phenomenal experience because it provides evidence that the self is lost in intense perceptual activity.

R4.2. Fractionation of access

As I have already mentioned in commenting on **Clark & Kiverstein** and **Rosenthal**, the evidence presented in **Gopnik** and **Izard et al.** is highly relevant to the issues of higher-order thought and agency that connect with consciousness. However, I want to make a few more remarks about developmental evidence.

1. The main methodological point made by **Gopnik** and **Izard et al.** is well taken. Adult capacities fractionate in infants. By studying that fractionation and the process by which the distinct capacities and abilities are integrated, we can learn a lot about adult consciousness and cognition.

2. Much of the behavior that seems at first glance to be a reflection of cognitive access in infants may bypass global broadcasting altogether. One example from adults mentioned by **Izard et al.** serves to illustrate the point: the example concerns dorsal visual system guidance of action as exemplified in Goodale and Milner's famous patient D.F. who can post a letter in a variable angled slot about as well as normal persons and avoid obstacles, but who does not have the globally broadcast visual contents that are created by the ventral visual system. D.F. knows about the angle of the slot only by noticing her own actions (Milner & Goodale 2008). The data presented by **Gopnik** and Izard et al. might make one wonder about the extent to which infants even have global broadcasting. But young infants – at least by 10 months of age – do show working memory systems that function in much the way those of adults do, albeit with a slightly lower capacity (Feigenson & Carey 2003; 2005; Feigenson et al. 2002).

3. I agree with **Gopnik**'s cluster of points having to do with the preponderance of exogenous (stimulated from outside) over endogenous (internally directed) attention in infants. However, I am not sure that this fact suggests that infants have a lantern of consciousness instead of a spotlight. Why can't exogenous attention focus the spotlight?

4. Although I agree in general with **Gopnik**'s point that infant cognition is not geared towards goal directed planning in the early months, infants are certainly capable of goal-directed action by age 5 months, and there is evidence of some capacity for goal-directed action (and perception of it) at age 3 months (Sommerville et al. 2005).

R4.3. Kinds of access

Representationism is the view that the phenomenology of an experience is the experience's representational (intentional) content. **Harman**, **Jacob**, **Lycan,** and **Tye** are advocates of this view; I am an opponent. However, I am willing to allow that every phenomenological state has representational content, even that the phenomenology consists – in part – in its having that representational content. I think that phenomenology goes beyond representational content, but this opposition played no role in the target article and will play no role in this reply. With this bone of contention cleared away, Harman's argument can be streamlined.

Harman argues that clearing up some unclarities renders the claim that phenomenology does not require cognitive access non-scientifically true or non-scientifically false. In my weight-lifting this morning, I experienced "flow." Among my visual experiences was seeing the barbell. Is the cognitive access in question with regard to the experience of seeing the barbell or with regard to the barbell itself? If the cognitive access in question is with regard to the experience of seeing the barbell, then we know from nonscientific observations of flow (Csikszentmihalyi 1975) that cognitive access to the experience itself destroys the experience, or at least its flow aspect. However, cognitive access to the barbell itself is guaranteed by the experience being an experience *of* the barbell. In neither case, Harman concludes, is there any need for science.

Recall that my target article agress with the idea of access as global broadcasting accepted by **Shanahan & Baars** and **Naccache & Dehaene**. What is globally broadcast is a representation with both representational and phenomenal content. What makes the access cognitive is that the consuming systems in terms of which global broadcasting is defined include the cognitive machinery of reasoning, planning, problem solving, categorization, memory, and the like. **Harman** assumes that what makes access cognitive is just that the accessed state is about something. But a frog might have a perceptual state that is about a fly without engaging any cognitive machinery. Let's go back to the first step of Harman's argument, that access has to be to the experience or to the barbell. In the framework I was using, access is neither to the experience nor the barbell but to the content of the state, since that is what is broadcast. So Harman's dilemma leaves out the crucial case. It may seem as if the issue between me and Harman is verbal – what is meant by *cognitive* and *access.* However, I chose the global broadcasting framework because it seems the best empirical framework for thinking about access and accessibility. Within that framework, the issue of phenomenology without access becomes an empirical issue. So, contrary to Harman, the issue is doubly scientific. Scientific considerations go into choosing the terms of the debate, and then, once they are chosen, scientific considerations determine the answer.

Balog, like **Harman**, is concerned with the question "Access to what?" She mentions, in effect, three options: global broadcasting of either (1) the nonconceptual contents of perception; or (2) the conceptual contents of perception; or (3) the introspective conceptual contents that are concerned with the phenomenal character of the experience itself. Option (3) involves phenomenal concepts, that is, concepts whose tokening itself involves phenomenology. (See papers in Alter & Walter [2006].) Balog argues that my interpretation of what subjects in the overflow experiments say "crucially relies on the subjects' introspective report of the phenomenality of their *entire visual experience*, including those aspects of the experience whose content is not access conscious" (in the

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

sense of option 2; Balog's emphasis), and this motivates her to suggest a sense of *access* (as in option 3) in which there is no gap between phenomenality and access.

Balog says my view of access involves the second option described earlier – that I define it in terms of conceptual contents. There is some justification for that in my 1995 paper (Block 1995b, see especially Note 11), which was written before I had adopted the global broadcasting model of access. My model of access in 1995 was dominated by inference and other reasoning, and it is often taken as part of the definition of nonconceptual content that it can play no role in inference. However, once I had adopted the global broadcasting model (Block 2002) the issue of whether the broadcast contents were conceptual or nonconceptual no longer looms large. I don't see why broadcast contents cannot be a mix.

Balog is responding to a problem that comes up repeatedly in the commentaries (especially **Naccache & Dehaene**'s and **Sergent & Rees**'s): How can I use reports to demonstrate inaccessible (including non-reportable) phenomenology? My reply is that in some cases, one infers unreportable or unreported phenomenal states via theoretical considerations that are themselves based on reports. In the overflow experiments, the situation is this: Generic contents are reportable, and one of those generic contents is that the subject has had specific phenomenology. That generic content need not deploy any phenomenal concepts in the sense of a concept whose tokening itself involves phenomenology. So there is no need to appeal to phenomenal concepts.

Tye notes that in one sense of the term *accessible*, phenomenology does not overflow accessibility. The sense he has in mind is that for each letter that the subject sees, the subject can demonstrate it with, for example, "that letter." He thinks that this point deprives my claim that phenomenology overflows accessibility from its "bite." Recall that I distinguished between wide and narrow accessibility. In my terms, what Tye has done is to provide another wide sense of accessibility – demonstration as opposed to identification, as in Sperling, or comparison, as in Landman et al. and Sligte et al. Demonstration, identification, and comparison are all different and have different cognitive demands. Demonstration is in a way more demanding (see **Koch & Tsuchiya**) because, whereas identification can take place without attention or with limited attention, demonstration requires attention. As I noted in the discussion of **Spener**, the grain of vision is finer than the grain of attention – you can see more than you can attend to. But regardless of whether Tye's version of wide accessibility is wide enough for his purposes, my point is about something different. The overflow point in the target article was this: The Sperling, Landman et al., and Sligte et al. experiments show that the phenomenal system is at least in part distinct from the global workspace because the capacity of the phenomenal system is greater than that of the global workspace. This point can be defended without being concerned with different senses of "accessibility," so I do not agree that providing another wide sense of accessibility affects the bite.

R4.4. Phenomenology and working memory

Lau & Persaud suppose that the apparent difference in capacity between phenomenology and working memory is just a consequence of deterioration of information, as in the "Telephone" or "Chinese whisper game" in which information is whispered from one person to another, becoming ever more garbled. However, what decreases in the Chinese whisper game is *quality* or *amount* of information, not *capacity* of an informational system. Lau & Persaud note that there are cases in which forced-choice responses overestimate phenomenology (blindsight) or yield inconsistent measures of phenomenology. They conclude that forced-choice reports are not ideal for measuring phenomenology. My replies to this are as follows:

1. Forced choice measures in the Landman et al. and Sligte et al. experiments are measures of working memory, not phenomenology. Subjects' ability to make the comparisons depends on the representations being present in the global workspace.

2. The Sperling version of the partial-report superiority effect used free recall, not forced choice as in Landman et al. and Sligte et al., and got similar results. It is the use of different methodologies with the same results that avoids difficulties of the sort that **Lau & Persaud** are mentioning.

Shanahan & Baars say their aim is to defend GWT (Global Workspace theory) from my challenge. More specifically, they say I use an outdated and imprecise concept of access and perpetuate a common misunderstanding of GWT that conflates the workspace with working memory. My main disagreement with them is simple: they say what is broadcast in the global workspace is all there is to consciousness, whereas I argue that it is an empirical issue whether there is more to consciousness and that the tentative answer is, Yes, there is. Shanahan & Baars' remarks do not seem to speak to this main issue. On the "misunderstanding" of working memory: the issue here is purely verbal. I use the term *working memory* to mean what Baars and Franklin call *active working memory* (Baars & Franklin 2003). One of the first things that strikes a reader of the "working memory" literature is that the term *working memory* is used differently by different theorists. Cowan (in press) notes "Working memory has been conceived and defined in three different, slightly discrepant ways." It would take too long to explain all three ways, but what I can say briefly is that a pretty standard model (Cowan 2005) includes three ingredients: (i) an activated part of long-term memory (often called "short-term memory"), (ii) attentional processes that operate to broadcast perceptual and short-term memory contents, and (iii) what is broadcast in the global workspace. I tend to use *short-term memory* for the activated part of long-term memory, and *working-memory* for what is broadcast in the global workspace; but judging from their commentary, Shanahan & Baars prefer to use both *short-term memory* and *working memory* for what I call *short-term memory*. The bottleneck of working memory on any definition I have seen, including that of Shanahan & Baars, is the capacity of the global workspace, so if that bottleneck is what is of interest, my use of *working memory* is apt. Whether or not I am attacking Global Workspace theory depends on whether Shanahan & Baars' controversial ideas about consciousness are part of it. Baars gets credit for many of the original ideas about the global workspace, but now that the ideas have been widely adopted, he no longer gets to dictate the terminology. My overflow argument

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

could be put entirely in terms of global workspace terminology, excluding all talk of memory. Here it is: The capacity of phenomenology is greater than the capacity of the global workspace, so phenomenology must contain machinery not part of the machinery of the global workspace.

R4.5. Report and color

Kentridge describes an intriguing case: Cerebral achromatopsics can recognize borders without being able to distinguish the colors on either side of the borders. Of course, guitar strings that are less than a JND (just-noticeable-difference) apart in pitch create interference "beats" that allow one to tell that they are different. And color patches that are less than a JND apart can be discriminated by normal subjects if they share a border. As I understand it, Kentridge et al. (2004) rule out such explanations, so phenomenal presentation of colors that the subject cannot report is a live option in the case that Kentridge describes. This case is as much a challenge for my views as an opportunity, since there are issues regarding how one can tell whether Kentridge's patient M.S. can see colors that he cannot report, and whether he has some kind of unconscious appreciation of color differences without phenomenal presentations of the colors themselves.

R5. Attention

R5.1. Phenomenology and attention

Lycan asks me to explain how there could be a form of awareness that is intrinsic to consciousness but does not necessarily involve attention or cognitive access. He notes (and I agree) that any form of awareness (worthy of the name) is representational and that there can be representation without attention. However, he does not see how cases of representation without attention could fail to be *sub*-personal. According to Lycan, representing by the whole person has to involve attending, or at least he does not see how it could not involve attending. The claim that representing by the whole person has to involve attention is in part empirical and I believe the evidence goes against it.

One kind of evidence derives from unconscious perception, more specifically perception that is unconscious because the person is not paying attention. To pick one of many examples: In visuospatial extinction, as mentioned in the target article, subjects often claim not to be able to see something on the left if there is a stimulus that competes for attention on the right. But as first noted by Volpe et al. (1979) (and nailed down by Verfaellie et al. 1995), the subject's claim not to see the object on the left is often combined with the ability to make comparisons between the thing on the left and the thing on the right. The thing on the left is certainly *seen* and, it is usually supposed, unconsciously so. Unconscious seeing is still seeing, and seeing by the person, not by a part of the person. This is seeing and therefore representing without, as far as we can tell, attention (whether or not the seeing on the left is unconscious). Therefore, the principle that Lycan can see no alternative to no personal-level representation without attention, is empirically suspect.

In the discussion of **Spener**, I mentioned another type of evidence against the claim that there is no personal level representation without attention, namely, that one can see more than one can attend to. Seeing the speckles on the speckled hen is personal and phenomenal, but there is no attention to the individual speckles; the individual speckles are conscious but unattended.

Lycan (1996; 2004) follows Locke (1689/1975) and Armstrong (1968) in holding that what makes a state conscious is that it is the object of an "inner sense," which Lycan seems to think is to be explained in terms of attention. The point I just made is not directly relevant to that view, but rather to supporting it in terms of personal-level representation requiring attention. However, as noted by **Koch & Tsuchiya** (and Koch & Tsuchiya 2007), *conscious* gist perception of peripheral non-targets persists for certain kinds of stimuli even when attention is maximally drained off by a demanding central task. So there is direct evidence against the attentional version of the inner sense thesis as well.

In cases of inattentional blindness, subjects say they do not see the stimulus – and this contrasts with the Sperling and the Landman et al. cases in which subjects clearly do see the stimulus. Relying on this contrast, **Prinz** and **Grush** are convinced (against **Naccache & Dehaene**) that phenomenology outstrips *actual* broadcasting in the global workspace, but, like **Snodgrass & Lepisto**, they think I go too far in supposing that perhaps phenomenology outstrips even potential broadcasting, as in the G.K. case. Prinz, like Lycan, argues that the flaw in my position is to allow phenomenology without attention. Prinz's account is based on the idea that attention is necessary for phenomenology and for accessibility in the sense of potential for broadcast in the global workspace (Prinz 2000; 2005). On both issues, I think the evidence is piling up against Prinz's view. First, on the claim that attention is necessary for accessibility, that is, potential broadcast: Dehaene and his colleagues (Dehaene 2006; Kouider et al. 2007) have convincingly shown that in conditions in which attention is maximally drained away by other tasks, representations that are as unattended as can be ensured by such conditions are nonetheless very strongly activated. They call this category "preconscious" as opposed to "unconscious" to indicate the strength of the activations, their strong influence, and their potential to be globally broadcast. Tse et al. (2005) have obtained similar results. These representations are accessible in that with a shift of attention, they will be accessed. Regardless of whether or not attention is necessary for actual access, it is not necessary for potential access, contrary to Prinz.

On the more important issue of the relation between attention and phenomenology, as **Koch & Tsuchiya** note, the evidence points toward the conclusion that neither is necessary for the other (Koch & Tsuchiya 2007). On attention without phenomenology, subjects' attention is drawn by nudes of the opposite sex (for heterosexuals) in conditions in which the nudes are invisible according to experimental standards (Jiang et al. 2006). (Whether the invisible pictures attract attention to one side of the visual field or the other is judged by performance on a subsequent task to which distribution of attention is crucial.) Kentridge et al. (1999) have shown the effects of attention in blindsight. On phenomenology

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

without attention, again in a "dual task" paradigm, subjects are able to see and report certain kinds of scene gists (e.g., the presence of a face in the periphery) despite maximal siphoning off of attention by a very demanding task. Work by Alvarez and Oliva (in press) suggests that without attention (or much attention) conscious visual representations represent much more abstract properties of stimuli, for example, center of mass rather than shape. As Koch & Tsuchiya (2007) point out, it is difficult to make absolutely sure that there is no attention devoted to a certain stimulus; but given this limitation, this evidence points away from both of **Prinz**'s claims. And the same point about the speckled hen case I made in discussing **Spener** and **Lycan** applies also to **Prinz**.

R6. Mesh between psychology and neuroscience

Burge argues that the conclusion of my target article can be derived from the psychological evidence alone without the argument from mesh with neuroscience. Here is a version of his argument. First, there is specific (not just generic) phenomenology prior to the cue in the Sperling-type experiments. Second, conscious events occur; they happen at a time. So we would have to have special reason to think that something that happens *after* a conscious event – or a disposition for something to happen after the conscious event – is necessary to the identity of the conscious event. If I see something red now, why should we think that the conditional occurrence of my later saying "I saw something red" if asked is partly constitutive of my sensation of red now? So the best explanation of the psychological evidence is that the machinery of the processes of access do not constitutively overlap with the machinery of consciousness itself.

Snodgrass & Lepisto say that to show that the machinery of consciousness and access are distinct, we do not need to show that there is phenomenality that cannot be accessed, but only to show that there is phenomenality that *is not in fact* accessed. And of course that is shown by the Sperling, Landman et al. and Sligte et al. experiments – *if*, as I have argued, those experiments reveal specific phenomenology without specific access. However, Snodgrass & Lepisto are neglecting the position that **Burge** refutes, the view that the machinery of access can be constitutively related to the machinery of phenomenology even via a connection of potential. Philosophers in the functionalist (and behaviorist) tradition have been impressed with the analogy between consciousness and dispositions such as solubility. (For example: **Grush** and **Prinz** advocate potential access as a condition of phenomenology.) Burge's argument gives a powerful reason for dissatisfaction with this analogy. The mesh argument gives an experimentally based way of reaching Burge's conclusion about overlap of machinery. However, I did not try to directly rule out potential access as a constitutively necessary condition, and so Burge's argument is a welcome addition.

Hulme & Whiteley note that inference to the best explanation requires a comparison (a point also noted by **Van Gulick**). If there is only one horse in the race, it wins. The comparison I had in mind was between the model on which recurrent activation in the back of the head without global broadcasting is sufficient for phenomenology and a model on which global broadcasting is required. My point was that the former yields a mechanism of overflow, whereas the latter does not. Hulme & Whiteley suggest another option: that recurrent activation is actually best thought of as a kind of binding. According to their hypothesis, activation of the relevant areas over threshold, but without binding, would produce amodal experience; and they give one example involving patient G.Y. that could be interpreted that way. I would not want to base much on G.Y.'s descriptions since he has given so many somewhat different and perhaps incompatible descriptions of his experience. Also, I don't follow Hulme & Whiteley's comment on the disruption of moving phosphenes by zapping V1 with trans-cranial magnetic stimulation. The subjects say they have stationary phosphenes, which does not sound amodal. However, Hulme & Whiteley's model would be fine for my purposes since it shares a feature not shared by the competitor I was trying to refute, namely, the view that a representation cannot be phenomenal unless it is globally broadcast.

Lamme says I have not integrated fully unconscious (neither accessible nor phenomenal) processing into the mesh argument and have not adequately considered arguments intrinsic to neuroscience. He and I agree almost completely about what is conscious and what is not and the reasoning why, but not on the abstract methodological description. I say it is inference to the best explanation. He says the key is criteria that are intrinsic to neuroscience. Lamme says the issue is whether Dehaene's I_2 is more similar to I_1 or to I_3, using criteria intrinsic to neuroscience. He makes a convincing case that I_2 is indeed more similar to I_3 than to I_1, but I think he misdescribes his own reasoning. Similarity in neuroscience and everywhere is *always in a respect* – according to a certain metric. There is no *abstract* issue of whether I_2 is more like I_1 than I_3. This point was first rammed down the throats of philosophers by Nelson Goodman (1976) who pointed out that any two objects chosen at random are similar in an infinity of respects and dissimilar in another infinity of respects. For example, you and I are similar in being more than one meter from the Eiffel Tower, more than 1.1 meter from the Eiffel Tower, and so on. The same point can be made *within neuroscience. One can come up with neural similarity metrics that give opposite rankings*. Representations in I_2 are more similar to those in I_1 than in I_3, in that the category I_3 involves frontal activations, whereas the representations in I_1 and I_2 do not. This respect of similarity is just as *intrinsic to neuroscience* as the one Lamme uses. Figure-ground organization, binding, and perceptual organization are important dimensions of similarity not because they are neural but because they are associated with *known conscious processing*, as Lamme's former and current students, **Landman & Sligte** note; and feature extraction is *excluded* because we have reason to think it happens unconsciously in the feed-forward sweep. Further, Lamme says that our opponents can give their own mesh argument; but the opposition's mesh argument cannot explain the overflow phenomenon, which is only my mesh argument is better. These points all fit into the category of classic inference to the best explanation reasoning of the sort that I am recommending.

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

R7. Does consciousness even exist?

McDermott says that the ultimate theory of how the brain works "will of course not refer to anything like phenomenology, but only to neural structures," concluding that as science marches on, notions of phenomenal consciousness will give way to neurally specified cognitive access. I have two criticisms. First, why *replacement* rather than *reduction*? The distinction I am appealing to is described in every introductory Philosophy of Science text (e.g., Rosenberg 2005). To illustrate: The concept of "phlogiston" has been *replaced* by the concept of oxygen. By contrast, we still have the concepts of heat and temperature. Heat has been *reduced* to molecular kinetic energy: heat exists and is molecular kinetic energy. Reductionist physicalists (a category that includes people as diverse as me and the Churchlands) hold that phenomenal consciousness can be reduced in neuroscientific terms. McDermott speaks of the buzz saw that is cutting through the science of consciousness. But the buzz saw of the revolution in chemistry in the eighteenth century did not show that there was no such thing as heat, temperature, pressure, or entropy, but rather, that they could be understood in molecular terms, that is, reduced rather than replaced. Of course, there are some cases to which the reduction/replacement distinction does not neatly apply. One much discussed example is the gene (Darden & Tabery 2007) for which there is no straightforward answer to the question of whether there are genes and they are snippets of DNA, or whether genes have been shown to not exist. Perhaps the most charitable interpretation of McDermott's remarks on life and subjectivity is that he predicts that the case of consciousness will end up resembling the case of the gene.

My second criticism of **McDermott** is: Why suppose that the reduction or replacement of the future will be in terms of access as opposed to lower-level neuroscience; for example, in terms of recurrent activation of neural connections? Computer scientists tend to assume – without argument – that anything a neuroscientist might discover about what consciousness is will be basically computational. They often assume it will be implementable in a silicon computer. The underlying disagreement here is between physicalist and functionalist reduction (or replacement). The difference is a form of a dispute about the mind/body problem that has been around in one form or another for ages and is discussed in detail in my two most recent books (Block 2007; 2008).

References

[The letters "a" and "r" before author's initials stand for target article and response references, respectively.]

Abrams, R. L. & Grinspan, J. (2007) Unconscious semantic priming in the absence of partial awareness. *Consciousness and Cognition* 16:942–53. [rNB]

Adelson, E. H. (1978) Iconic storage: The role of rods. *Science* 201:544–46. [aNB]

Aimola Davies, A. M. (2004) *Disorders of spatial orientation and awareness: Cognitive and behavioral rehabilitation: From neurobiology to clinical practice.* Guilford Press. [aNB]

Alkire, M. T. & Miller, J. (2005) General anesthesia and the neural correlates of consciousness. *Progress in Brain Research* 150:229–44. [aNB]

Alter, T. & Walter, S. (2006) *Phenomenal concepts and phenomenal knowledge: New essays on consciousness and physicalism*. Oxford University Press. [rNB]

Alvarez, G. A. & Oliva, A. (in press) The representation of ensemble visual features outside the focus of attention. *Psychological Science*. [rNB]

American Congress of Rehabilitation Medicine (1995) Recommendations for use of uniform nomenclature pertinent to patients with severe alterations of consciousness. *Archives of Physical Medicine and Rehabilitation* 76:205–209. [AV]

Anderson, A. K. & Phelps, E. A. (2001) Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature* 411:305–09. [CEI]

Andrade, J. (2001) The contribution of working memory to conscious experience. In: *Working memory in perspective*, ed. J. Andrade. Psychology Press. [TB]

Andrews, K., Murphy, L., Munday, R. & Littlewood, C. (1996) Misdiagnosis of the vegetative state: Retrospective study in a rehabilitation unit. *British Medical Journal* 313(7048):13–16. [AV]

Aristotle (1955) On dreams. In: *Aristotle. Parva naturalia*, ed. W. D. Ross, pp. 175–223. Clarendon Press. [aNB]

Armstrong, D. M. (1968) *A materialist theory of the mind*, Humanities Press. [rNB]

(1977) The causal theory of the mind. *Neue Heft für Philosophie* 11:82–95. [aNB]

(1981) What is consciousness? In: *The nature of mind and other essays*, by D. M. Armstrong, pp. 55–67. Cornell University Press. [WGL]

Arnell, K. M., Killman, K. V. & Fijavz, D. (2007) Blinded by emotion: Target misses follow attention capture by arousing distractors in RSVP. *Emotion* 7:465–77. [CEI]

Avidan, G., Hasson, U., Hendler, T., Zohary, E. & Malach, R. (2002) Analysis of the neuronal selectivity underlying low fMRI signals. *Current Biology* 12(12):964–72. [RM]

Baars, B. J. (1988) *A cognitive theory of consciousness.* Cambridge University Press. [KB, aNB]

(1997) *In the theater of consciousness: The workspace of the mind.* Oxford University Press. [KB, aNB, RWK, DM]

(2002) The conscious access hypothesis: Origins and recent evidence. *Trends in Cognitive Sciences* 6(1):47–52. [MSh]

(2005) Global Workspace theory of consciousness: Toward a cognitive neuroscience of human experience. *Progress in Brain Research* 150:45–53. [VAFL]

Baars, B. & Franklin, S. (2003) How conscious experience and working memory interact. *Trends in Cognitive Sciences* 17(4):166–72. [rNB, MSh]

Baars, B. J., Ramsoy, T. Z. & Laureys, S. (2003) Brain, conscious experience and the observing self. *Trends in Neurosciences* 26(12):671–75. [RM]

Bartels, A. & Zeki, S. (2004) Functional brain mapping during free viewing of natural scenes. *Human Brain Mapping* 21. [aNB]

Basso, A., Speinnler, H., Vallar, G. & Zanobio, M. E. (1982) Left hemisphere damage and selective impairment of auditory verbal short term memory: A case study. *Neuropsychologia* 20:263–74. [aNB]

Bauer, G., Gerstenbrand, F. & Rumpl, E. (1979) Varieties of the locked-in syndrome. *Journal of Neurology* 221(2):77–91. [AV]

Bechara, A., Damasio, H. & Damasio, A. (2000) Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex* 10:295–307. [CEI]

Bermúdez, J. & Macpherson, F. (1998) Nonconceptual content and the nature of perceptual experience. *Electronic Journal of Analytical Philosophy* 6 (1998). Archived at: http://ejap.louisiana.edu/archives.html [AC]

Birbaumer, N., Ghanayim, N., Hinterberger, T., Iversen, I., Kotchoubey, B., Kubler, A., Perelmouter, J., Taub, E. & Flor, H. (1999) A spelling device for the paralysed. *Nature* 398(6725):297–98. [AV]

Block, N. (1978) Troubles with functionalism. *Minnesota Studies in the Philosophy of Science* 9:261–325. [aNB]

(1980) What is functionalism? In: *Readings in the philosophy of psychology*, ed. N. Block, pp.171–84. Harvard University Press. [aNB]

(1983) The photographic fallacy in the debate about mental imagery. *Nous* 17:651–62. [rNB]

(1990) Consciousness and accessibility. *Behavioral and Brain Sciences* 13(4):596–98. [KB, aNB]

(1992) Begging the question against phenomenal consciousness. *Behavioral and Brain Sciences* 15(2):205–206. [aNB]

(1995a) How many concepts of consciousness? *Behavioral and Brain Sciences* 18(2):272–84. [aNB]

(1995b) On a confusion about a function of consciousness. *Behavioral and Brain Sciences* 18(2):227–47. [KB, arNB, LN, JKO, JP, WGL]

(1997) Biology versus computation in the study of consciousness. *Behavioral and Brain Sciences* 20(1):159–66. [aNB]

(2001) Paradox and cross purposes in recent work on consciousness. *Cognition* 79(1–2):197–220. [arNB, LN]

(2002) Concepts of consciousness. In: *Philosophy of mind: Classical and contemporary readings*, ed. D. Chalmers, pp. 206–18. Oxford University Press. [KB, rNB]

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

(2003) Mental paint. In: *Reflections and replies: Essays on the philosophy of Tyler Burge*, ed. M. Hahn & B. Ramberg. MIT Press. [PJ]
(2005) Two neural correlates of consciousness. *Trends in Cognitive Sciences* 9(2):46–52. [arNB, VAFL, RL, LN]
(2007) Remarks on chauvinism and the mind-body problem. In: *Consciousness, function, and representation*, vol. 1, ed. N. Block, pp. 7–12. MIT Press. [rNB]
(2007) Wittengstein and qualia. *Philosophical Prespectives* 21(1):73–115. [PJ]
(in press) Functional reduction. In: *Supervenience in mind*, ed. T. Horgan, M. H. Sabates & D. Sosa, MIT Press. [rNB]
Block, N. & Dworkin, G. (1974) IQ, heritability, and inequality I. *Philosophy and Public Affairs* 3(4):331–409. [aNB]
Boghossian, P. A. (2006) *Fear of knowledge: Against relativism and constructivism.* Oxford University Press. [aNB]
Bohm, D. & Hiley, B. J. (1993) *The undivided universe: An ontological interpretation of quantum theory.* Routledge. [DM]
Boly, M., Coleman, M. R., Davis, M. H., Hampshire, A., Bor, D., Moonen, G., Maquet, P. A., Pickard, J. D., Laureys, S. & Owen, A. M. (2007) When thoughts become action: An fMRI paradigm to study volitional brain activity in non-communicative brain injured patients. *Neuroimage* 36(3):979–92. [AV]
Breitmeyer, B. G., Kropfl, W. & Julesz, B. (1982) The existence of and role of retinotogic and spatiotopic forms of visual persistence. *Acta Psychologica* 52:175–96. [RL]
Breitmeyer, B. G., Ro, T. & Singhal, N. S. (2004) Unconscious color priming occurs at stimulus- not percept-dependent levels of processing. *Psychological Science* 15(3):198–202. [RL]
Brentano, F. (1874/1924) *Psychologie vom Empirischen Standpunkt.* Felix Meiner. [aNB, DMR]
Britten, K. H., Shadlen, M. N., Newsome, W. T. & Movshon, A. (1992) The analysis of visual motion: A comparison of neuronal and psychophysical performance. *Journal of Neuroscience* 12:4745–65. [aNB]
Brockmole, J. R., Wang, R. F. & Irwin, D. E. (2002) Temporal integration between visual images and visual percepts. *Journal of Experimental Psychology: Human Perception and Performance* 28(2):315–34. [rNB]
Burge, T. (2006) Reflections on two kinds of consciousness. In: *Philosophical essays, vol. II: Foundations of mind*, ed. T. Burge, pp. 392–419. Oxford University Press. [aNB]
Bushnell, I. W. R., Sai, F. & Mullin, J. T. (1989) Neonatal recognition of the mother's face. *British Journal of Developmental Psychology* 7:3–15. [CEI]
Byrne, A. (2004) What phenomenal consciousness is like. In: *Higher order theories of consciousness*, ed. R. Gennaro. John Benjamins. [aNB]
Carrasco, M. (2007) Visual attention alters appearance: Psychophysical studies of subjective experience. In: *Oxford companion to consciousness*, ed. P. Wilken, T. Bayne & A. Cleeremans. Oxford University Press. [aNB]
Carrasco, M., Ling, S. & Read, S. (2004) Attention alters appearance. *Nature Neuroscience* 7(3):308–313. [aNB, RWK]
Carruthers, P. (1989) Brute experience. *Journal of Philosophy* 86:258–69. [rNB]
(1992) *The animals issue: Moral theory in practice.* Cambridge University Press. [rNB]
(1999) Sympathy and subjectivity. *Australasian Journal of Philosophy* 77:465–82. [rNB]
(2000) *Phenomenal consciousness: A naturalistic theory.* Cambridge University Press. [aNB]
Caston, V. (2002) Aristotle on consciousness. *Mind* 111(444):751–815. [arNB]
Cattell, J. M. (1885) The inertia of the eye and brain. *Brain* 8(3):295–312. [aNB]
Cavanagh, P. He, S. & Intriligator, J. (1999) Attentional resolution: The grain and locus of visual awareness. In: *Neuronal basis and psycological aspects of consciousness*, ed. C. Taddai-Ferretti, C. Musio. World Scientific. [rNB]
Chalmers, D. (1996) *The conscious mind: In search of a fundamental theory*. Oxford University Press. [aNB, DM]
(1997) Availability: The cognitive basis of experience. *Behavioral and Brain Sciences* 20(1):148–49. [aNB].
(1998) On the search for the neural correlate of consciousness. In: *Toward a science of consciousness II: The second Tucson discussions and debates*, ed. S. Hameroff, A. Kaszniak & A. Scott, pp. 219–30. MIT Press. [aNB]
(2000) What is a neural correlate of consciousness? In: *Neural correlates of consciousness: Empirical and conceptual issues*, ed. T. Metzinger, pp. 17–40. MIT Press. [aNB]
Childs, N. L., Mercer, W. N. & Childs, H. W. (1993) Accuracy of diagnosis of persistent vegetative state. *Neurology* 43(8):1465–67. [AV]
Chisholm, R. (1942) The problem of the speckled hen. *Mind* 51:368–73. [MSp]
Chun, M. M. & Potter, M. C. (1995) A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance* 21:109–27. [CEI]
Churchland, P. S. (1986) *Neurophilosophy: Toward a unified science of the mind-brain*. MIT Press. [aNB]
(1994) Can neurobioloby teach us anything about consciousness? *American Philosophical Association Proceedings* 67(4):23–40. [aNB]
(2002) *Brain-wise: Studies in neurophilosophy.* MIT Press. [aNB]
(2005) A neurophilosophical slant on consciousness research. *Progress in Brain Research* 149:285–92. [aNB]
Claxton, L. J., Keen, R. & McCarty, M. E. (2003) Evidence of motor planning in infant reaching behavior. *Psychological Science* 14:354–56. [CEI]
Cohen, J. (2002) The grand illusion illusion. *Journal of Consciousness Studies* 9(5–6):141–57. [aNB]
Coltheart, M. (1980) Iconic memory and visible persistence. *Perception and Psychophysics* 27(3):183–228. [AB, aNB]
Cowan, N. (1995) *Attention and memory: An integrated framework.* Oxford University Press. [RL]
(2001) The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences* 24:87–185. [aNB]
(2005) Working-memory capacity limits in a theoretical context. In: *Human learning and memory: Advances in theory and applications*, ed. C. Izawa & N. Ohta, pp. 155–75. Erlbaum. [rNB]
(in press) What are the differences between long-term, short-term, and working memory? In: *The essence of memory*, ed. W. Sossin, J.-C. Jacaille, V. F. Castellucci & S. Belleville. Elsevier/Academic Press. [rNB]
Cowan, N., Morey, C. & Chen, Z. (2007) The legend of the magical number seven. In: *Tall tales about the brain: Things we think we know about the mind, but ain't so*, ed. S. Della Sala, pp. 45–59. Oxford University Press. [aNB]
Cowey, A. & Walsh, V. (2000) Magnetically induced phosphenes in sighted, blind and blindsighted subjects. *NeuroReport* 11:3269–73. [aNB]
Crick, F. & Koch, C. (1995) Are we aware of neural activity in primary visual cortex? *Nature* 375:121–23. [aNB, CK]
(1998a) Consciousness and neuroscience. *Cerebral Cortex* 8:97–107. [CK]
(1998b) Why is there a hierarchy of visual cortical and thalamic areas: The no-strong loops hypothesis. *Nature* 391:245–49. [CK]
(2005) What is the function of the claustrum? *Philosophical Transactions of the Royal Society of London B* 360:1271–319. [CK]
Csikszentmihalyi, M. (1975) *Beyond boredom and anxiety*. Jossey-Bass. [rNB]
(1990) *Flow: The psychology of optimal experience*. Harper and Row. [GH]
Curtis, C. & D'Esposito, M. (2003) Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences* 7(9):415–23. [aNB]
Damasio, A. (1999) *The feeling of what happens: Body and emotion in the making of consciousness.* Harcourt Brace. [aNB]
Darden, L. & Tabery, J. (2007) Molecular biology. In: *The Stanford encyclopedia of philosophy (Fall 2007 edition)*, ed. E. N. Zalta. Online publication. Available at: http://plato.stanford.edu/archives/fall2007/entries/molecular-biology. [rNB]
de Fockert, J. W., Rees, G., Frith, C. D. & Lavie, N. (2001) The role of working memory in visual selective attention. *Science* 291:1803–06. [aNB]
Debner, J. A. & Jacoby, L. L. (1994) Unconscious perception: Attention, awareness, and control. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20:304–17. [rNB, HL]
Dehaene, S. & Changeux, J. P. (2004) Neural mechanisms for access to consciousness. In: *The cognitive neurosciences III*, ed. M. Gazzaniga, pp. 1145–58. MIT Press. [aNB]
Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J. & Sergant, C. (2006) Conscious, preconscious, and subliminal processing: A testable taxonomoy. *Trends in Cognitive Sciences* 10(5):204–11. [arNB, AC, SK, RL, VAFL, LN, CS, MSh]
Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J. B., Le Bihan, D. & Cohen, L. (2004) Letter binding and invariant recognition of masked words: Behavioral and neuroimaging evidence. *Psychological Science* 15(5):307–13. [LN]
Dehaene, S., Kerszberg, M. & Changeux, J.-P. (1998) A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences USA* 95:14529–34. [aNB, MSh]
Dehaene, S. & Naccache, L. (2001) Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition* 79(1–2):1–37. [aNB, LN, JP, DMR]
Dehaene, S., Naccache, L., Le Clec, H. G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., van de Moortele, P. F. & Le Bihan, D. (1998) Imaging unconscious semantic priming. *Nature* 395(6702):597–600. [CS]
Dehaene, S., Sergent, C. & Changeux, J. P. (2003) A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proceedings of the National Academy of Sciences USA* 100:8520–25. [CS]
Demkiw, P. & Michaels, C. (1976) Motion information in iconic memory. *Acta Psychologica* 40:257–64. [aNB]
Dennett, D. C. (1988) Quining Qualia. In: *Consciousness in contemporary science*, ed. A. Marcel & E. Bisiach, pp. 381–414. Oxford University Press. [aNB]
(1991) *Consciousness explained.* Little, Brown. [arNB, RVG, PJ, CS]
(1993) The message is: There is no medium. *Philosophy and Phenomenological Research* 53(4):889–931. [aNB]
(2001) Are we explaining consciousness yet? *Cognition* 79:221–37. [aNB]

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

Di Lollo, V. (1980) Temporal integration in visual memory. *Journal of Experimental Psychology: General* 109:75–97. [aNB]

Di Lollo, V., Enns, J. T. & Rensink, R. A. (2000) Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General* 129:481–507. [CEI]

Dienes, Z. & Scott, R. (2005) Measuring unconscious knowledge: Distinguishing structural knowledge and judgment knowledge. *Psychological Research* 69(5–6):338–51. [CS]

Dretske, F. (1993) Conscious experience. *Mind* 102(406):263–83. Also published in: Dretske, F. (2000) *Perception, knowledge and belief*. Cambridge University Press. [PJ]

(1994) Differences that make no difference. *Philosophical Topics* 22(1–2):41–57. Also published in: Dretske, F (2000) *Perception, knowledge and belief*. Cambridge University Press. [PJ]

(2004) Change blindness. *Philosophical Studies* 120(1–3):1–18. [aNB]

Driver, J. & Vuilleumier, P. (2001) Perceptual awareness and its loss in unilateral neglect and extinction. *Cognition* 79(1–2):39–88. [aNB, CK]

Edelman, G. M. (2004) *Wider than the sky: The phenomenal gift of consciousness.* Yale University Press. [aNB]

Eimas, P. D. & Quinn, P. C. (1994) Studies on the formation of perceptually based basic-level categories in young infants. *Child Development* 65:903–17. [CEI]

Ellsworth, C. P., Muir, D. W. & Hains, S. M. J. (1993) Social competence and person-object differentiation: An analysis of the still-face effect. *Developmental Psychology* 29:63–73. [CEI]

Engel, G. R. (1970) An investigation of visual responses to brief stereoscopic stimuli. *Quarterly Journal of Experimental Psychology* 22:148–60. [aNB]

Evans, G. (1982) *The Varieties of reference*, ed. J. McDowell. Oxford University Press. [AC]

Fearing, F. (1970) *Reflex action.* MIT Press. [CK]

Feigenson, L. & Carey, S. (2003) Tracking individuals via object files: Evidence from infants' manual search. *Developmental Science* 6:568–84. [arNB]

(2005) On the limits of infants' quantification of small object arrays. *Cognition* 97(3):295–313. [rNB]

Feigenson, L., Carey, S. & Hauser, M. (2002) The representations underlying infants' choice of more: Object files versus analog magnitudes. *Psychological Science* 13:150–56. [arNB]

Fodor, J. A. (1983) *The modularity of mind.* MIT Press. [aNB]

Fuster, J. M. (1973) Unit activity in prefrontal cortex during delayed-response performance: Neuronal correlates of transient memory. *Journal of Neurophysiology* 36:61–78. [aNB]

Gaillard, R., Del Cul, A., Naccache, L., Vinckier, F., Cohen, L. & Dehaene, S. (2006) Nonconscious semantic processing of emotional words modulates conscious access. *Proceedings of the National Academy of Sciences USA* 103(19):7524–29. [CS]

Gazzaniga, M. S., Ivry, R. & Mangun, G. (2002) *Cognitive neuroscience: The biology of the mind.* W. W. Norton. [aNB]

Gazzaniga, M. S., LeDoux, J. E. & Wilson, D. H. (1977) Language, praxis, and the right hemisphere: Clues to some mechanisms of consciousness. *Neurology* 27(12):1144–47. [LN]

Gennaro, R. (1996) *Consciousness and self-consciousness*. John Benjamins. [WGL]

Giacino, J. T., Ashwal, S., Childs, N., Cranford, R., Jennett, B., Katz, D. I., Kelly, J. P., Rosenberg, J. H., Whyte, J., Zafonte, R. D. & Zasler, N. D. (2002) The minimally conscious state: Definition and diagnostic criteria. *Neurology* 58(3):349–53. [AV]

Giacino, J. T., Hirsch, J., Schiff, N. & Laureys, S. (2006) Functional neuroimaging applications for assessment and rehabilitation planning in patients with disorders of consciousness. *Archives of Physical Medicine and Rehabilitation* 87(12, Suppl. 2):S67–76. [AV]

Gill, N. F. & Dallenbach, K. M. (1926) A preliminary study of the range of attention. *The American Journal of Psychology* 37(2):247–56. [aNB]

Goldberg, I. I., Harel, M. & Malach, R. (2006) When the brain loses its self: Prefrontal inactivation during sensorimotor processing. *Neuron* 50(2):329–39. [rNB, RM, JP]

Goldman-Rakic, P. (1987) Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In: *Handbook of physiology: The nervous system, higher functions of the brain*, ed. F. Plum, pp. 373–417. American Physiological Society. [aNB]

Golland, Y., Bentin, S., Gelbard, H., Benjamini, Y., Heller, R., Nir, Y. U., Hasson & Malach, R. (2007) Extrinsic and intrinsic systems in the posterior cortex of the human brain revealed during natural sensory stimulation. *Cerebral Cortex* 17(4):766–77. [RM]

Goodale, M. A., Milner, A. D., Jakobson, L. S. & Carey, D. P. (1991) A neurological dissociation between perceiving objects and grasping them. *Nature* 349:154–56. [HL]

Goodman, N., (1976) *Languages of art: An approach to a theory of symbols*, 2nd edition. Hacket Publishing Company. [HL]

Greenwald, A. G., Abrams, R. L., Naccache, L. & Dehaene, S. (2003) Long-term semantic memory versus contextual memory in unconscious number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29(2):235–47. [CS]

Grill-Spector, K., Sayres, R. & Ress, D. (2006) High-resolution imaging reveals highly selective nonface clusters in the fusiform face area. *Nature Neuroscience* 9:1177–85. [aNB]

(2007) Corrigendum. *Nature Neuroscience* 10:133. [aNB]

Haase, S. & Fisk, G. (2001) Confidence in word detection predicts word identification: Implications for an unconscious perception paradigm. *American Journal of Psychology* 114:439–68. [MSn]

Hanley, J. R., Young, H. W. & Pearson, N. A. (1991) Impairment of visuo-spatial sketch pad. *Quarterly Journal of Experimental Psychology A* 43:101–25. [aNB]

Harman, G. (1965) The inference to the best explanation. *The Philosophical Review* 74(1):88–95. [aNB]

(1990) The intrinsic quality of experience. *Philosophical Perspectives* 4:31–52. [GH]

Harrison, R. V., Harel, N., Panesar, J. & Mount, R. J. (2002) Blood capillary distribution correlates with hemodynamic-based functional imaging in cerebral cortex. *Cerebral Cortex* 12(3):225–33. [CK]

Hasselmo, M. E. & Stern, C. E. (2006) Mechanisms underlying working memory for novel information. *Trends in cognitive sciences* 10(11): 487–93. [aNB]

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. & Malach, R. (2004) Intersubject synchronization of cortical activity during natural vision. *Science* 303(5664):1634–40. [arNB, RM]

Hauser, M., Carey, S. & Hauser, L. (2000) Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proceedings of the Royal Society of London: Biological Sciences* 267:829–33. [aNB]

Haynes, J. D., Driver, J. & Rees, G. (2005) Visibility reflects dynamic changes of effective connectivity between V1 and fusiform cortex. *Neuron* 46(5): 811–21. [VAFL, LN]

Haynes, J. D. & Rees, G. (2006) Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience* 7(7):523–34. [aNB, LN]

Heath, R. G., Carpenter, M. B., Mettler, F. A. & Kline, N. S. (1949) Visual apparatus: Visual fields and acuity, color vision, autokinesis. In: *Selective partial ablation of the frontal cortex: A correlative study of its effects on human psychotic subjects; problems of the human brain*, ed. F. A. Mettler, pp. 489–91. P. B. Hoeber. [aNB]

Heeger, D. J., Boynton, G. M., Demb, J. B., Seideman, E. & Newsome, W. T. (1999) Motion opponency in visual cortex. *Journal of Neuroscience* 19:7162–74. [aNB]

Heywood, C. A. & Kentridge, R. W. (2003) Achromatopsia, color vision and cortex. *Neurological Clinics of North America* 21:483–500. [RWK]

Heywood, C. A., Kentridge, R. W. & Cowey, A. (1998) Form and motion from colour in cerebral achromatopsia. *Experimental Brain Research* 123:145–53. [RWK]

Hinterberger, T., Birbaumer, N. & Flor, H. (2005) Assessment of cognitive function and communication ability in a completely locked-in patient. *Neurology* 64:1307. [AV]

Hochberg, L. R., Serruya, M. D., Friehs, G. M., Mukand, J. A., Saleh, M., Caplan, A. H., Branner, A., Chen, D., Penn, R. D. & Donoghue, J. P. (2006) Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature* 442(7099):164–71. [AV]

Hofer, H., Singer, B. & Williams, D. R. (2005) Different sensations from cones with the same photopigment. *Journal of Vision* 5:444–54. [RWK]

Hopkin, M. (2006) 'Vegetative' patient shows signs of conscious thought. *Nature* 443(7108):132–33. [aNB]

Huang, L. & Pashler, H. (2007) A Boolean map theory of visual attention. *Psychological Review* 114(3):599–631. [rNB]

Huang, L., Treisman, A. & Pashler, H. (2007) Characterizing the limits of human visual awareness. *Science* 317:823–25. [rNB]

Huk, A. C., Ress, D. & Heeger, D. J. (2001) Neuronal basis of the motion aftereffect reconsidered. *Neuron* 32:161–72. [aNB]

Huntley-Fenner, G., Carey, S. & Solimando, A. (2002) Objects are individuals but stuff doesn't count: Perceived rigidity and cohesiveness influence infants' representations of small numbers of discreet entities. *Cognition* 85:203–21. [aNB]

Hupe, J. M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P. & Bullier, J. (1998) Cortical feedback improves discrimination between figure and background by V1, V2, and V3 neurons. *Nature* 394(6695):784–87. [RL]

Hurley, S. L. (1997) Non-conceptual self-consciousness and agency: Perspective and access. *Communication and Cognition* 30(3/4):207–48. (Part 1 of Special Issue: *Approaching Consciousness*). [AC]

Iwasaki, S. (1993) Spatial attention and two modes of visual consciousness. *Cognition* 49:211–33. [CK]

Izard, C. E. (in press) Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspectives on Psychological Science*. [CEI]

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

Izard, C. E., Fantauzzo, C. A., Castle, J. M., Haynes, O. M., Rayias, M. F. & Putnam, P. H. (1995) The ontogeny and significance of infants' facial expressions in the first 9 months of life. *Developmental Psychology* 31:997–1013. [CEI]

Jackson, F. & Pinkerton, R. J. (1973) On an argument against sensory items. *Mind* 82:269–72. [MSp]

James, T. W., Culham, J., Humphre, G. K., Milner, A. D. & Goodale, M. A. (2003) Ventral occipital lesions impair object recognition but not object-directed grasping: An fMRI study. *Brain* 126:2463–75. [CEI]

James, W. (1890) *Principles of psychology.* Henry Holt. [aNB]

Jiang, Y., Costello, P., Fang, F., Huang, M. & He, S. (2006) A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences USA* 103:17048–52. [rNB, CEI]

Jolij, J. & Lamme, V. A. F. (2005) Repression of unconscious information by conscious processing: Evidence from affective blindsight induced by transcranial magnetic stimulation. *Proceedings of the National Academy of Sciences USA* 102(30):10747–51. [RL]

Kamitani, Y. & Tong, F. (2005) Decoding the visual and subjective contents of the human brain. *Nature Neuroscience* 8:679–85. [aNB]

Kammer, T. (1999) Phosphenes and transient scotomas induced by magnetic stimulation of the occipital lobe: Their topographic relationship. *Neuropsychologia* 37:191–98. [aNB]

Kanwisher, N. (2001) Neural events and perceptual awareness. *Cognition* 79:89–113. [aNB]

(2006) What's in a face? *Science* 311(5761):617–18. [aNB]

(2007) Does the fusiform face area contain subregions highly selective for nonfaces? *Nature Neuroscience* 10(1):3–4. [aNB]

Kastner, S., De Weerd, P., Desimone, R. & Ungerleider, L. G. (1998) Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science* 282(5386):108–11. [LN]

Kentridge, R. W., Heywood, C. A. & Cowey, A. (2004) Chromatic edges, surfaces and constancies in cerebral achromatopsia. *Neuropsychologia* 42:821–30. [RWK]

Kentridge, R. W., Heywood, C. A. & Weiskrantz, L. (1999) Attention without awareness in blindsight. *Proceedings of the Royal Society of London: Biological Sciences* 266(1430):1805–11. [rNB]

Kentridge, R. W., Heywood, C. A. & Cowey, A. (2004) Chromatic edges, surfaces, and constancies in cerebral achromatopsia. *Neuropsychologia* 42(6):821–30. [rNB]

Kobes, B. (1995) Access and what it is like. *Behavioral and Brain Sciences* 18(2):260. [aNB]

Koch, C. (2004) *The quest for consciousness: A neurobiological approach.* Roberts. [aNB, CK, DM]

Koch, C. & Tsuchiya, N. (2007) Attention and consciousness: Two distinct brain processes. *Trends in Cognitive Sciences* 11:16–22. [rNB, CK, SK]

Kosslyn, S. M., Thompson, W. L. & Ganis, G. (2006) *The case for mental imagery.* Oxford University Press. [rNB]

Kouider, S., Dehaene, S., Jobert, A. & Le Bihan, D. (2007) Cerebral bases of subliminal and supraliminal priming during reading. *Cerebral Cortex* 17(9):2019–29. [arNB, SK, JP]

Kouider, S. & Dupoux, E. (2004) Partial awareness creates the "illusion" of subliminal semantic priming. *Psychological Science* 15(2):75–81. [rNB, SK]

(2007) How "semantic" is response priming restricted to practiced items? A reply to Abrams & Grinspan (2007). *Consciousness and Cognition* 16:954–56. [rNB]

Kourtzi, Z. & Kanwisher, N. (2000) Activation in human MT/MST by static images with implied motion. *Journal of Cognitive Neuroscience* 12:48–55. [aNB]

Kriegel, U. (2005) Naturalizing subjective character. *Philosophy and Phenomenological Research* 71:23–57. [aNB, WGL, DMR]

Kriegel, U. & Williford, K. (2006) *Self-representational approaches to consciousness.* MIT Press. [aNB]

Kriegeskorte, N., Sorger, B., Naumer, M., Scwarzbach, J., van den Boogert, E., Hussy, W. & Goebel, R. (2003) Human cortical object recognition from a visual motion flowfield. *Journal of Neuroscience* 23(4):1451. [aNB]

Kubler, A. & Neumann, N. (2005) Brain-computer interfaces – the key for the conscious brain locked into a paralyzed body. *Progress in Brain Research* 150:513–25. [AV]

Lamme, V. A. F. (2003) Why visual attention and awareness are different. *Trends in Cognitive Sciences* 7(1):12–18. [aNB, CK, RL, VAFL]

(2004) Separate neural definitions of visual consciousness and visual attention; a case for phenomenal awareness. *Neural Networks* 17:861–72. [VAFL]

(2006) Towards a true neural stance on consciousness. *Trends in Cognitive Sciences* 10(11):494–501. [aNB, RL]

Lamme, V. A. F. & Roelfsema, P. R. (2000) The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neuroscience* 23(11):571–79. [RL, VAFL, CS]

Lamme, V. A. F., Zipser, K. & Spekreijse, H. (1998) Figure-ground activity in primary visual cortex is suppressed by anesthesia. *Proceedings of the National Academy of Sciences USA* 95(6):3263–68. [RL, VAFL]

(2002) Masking interrupts figure-ground signals in V1. *Journal of Cognitive Neuroscience* 14(7):1044–53. [RL]

Landman, R., Spekreijse, H. & Lamme, V. A. F. (2003) Large capacity storage of integrated objects before change blindness. *Vision Research* 43(2):149–64. [AB, KB, arNB, TB, AC, RG, CK, RL, VAFL, LN, JKO, DMR, CS, MSh, MSp, RVG]

(2004) The role of figure-ground segregation in change blindness. *Psychonomic Bulletin and Review* 11(2):254–61. [RL]

Lau, H. C. & Passingham, R. E. (2006) Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proceedings of the National Academy of Sciences USA* 103(49):18763–68. [HL, CS]

(2007) Unconscious activation of the cognitive control system in the human prefrontal cortex. *Journal of Neuroscience* 27:5805–11. [VAFL]

Laureys, S. (2005) The neural correlate of (un)awareness: Lessons from the vegetative state. *Trends in Cognitive Sciences* 9(2):556–59. [aNB]

Laureys, S., Pellas, F., Van Eeckhout, P., Ghorbel, S., Schnakers, C., Perrin, F., Berré, J., Faymonville, M. E., Pantke, K. H., Damas, F., Lamy, M., Moonen, G. & Goldman, S. (2005) The locked-in syndrome: What is it like to be conscious but paralyzed and voiceless? *Progress in Brain Research* |V150: 495–511. [LN, AV]

Levine, J. (1983) Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly* |V64:354–61. [aNB]

(2001) *Purple haze: The puzzle of consciousness.* Oxford University Press. [aNB]

(2006) Conscious awareness and (self)representation. In: *Self-representational approaches to consciousness*, ed. U. Kriegel & K, pp. 173–97. Williford. MIT Press. [aNB]

Levy, I., Hasson, U. & Malach, R. (2004) One picture is worth at least a million neurons. *Current Biology* 14(11):996–1001. [RM]

Lieberman, M. D., Eisenberger, N. I., Crockett, M. J., Tom, S. M., Pfeifer, J. H. & Way, B. M. (2007) Putting feelings into words: Affect labeling disrupts amygdala activity in response to affective stimuli. *Psychological Science* 18: 421–28. [CEI]

Llinás, R. R. (2001) *I of the vortex.* MIT Press. [aNB]

Llinás, R. R., Ribary, U., Contreras, D. & Pedroarena, C. (1998) The neuronal basis for consciousness. *Philosophical Transactions of the Royal Society B* 353:1841–49. [aNB]

Locke, J. (1689/1975) An essay concerning human understanding. Clarenda Press. [rNB]

Loftus, G. & Irwin, D. (1998) On the relations among different measures of visible and informational persistence. *Cognitive Psychology* 35:135–99. [AB, arNB]

Logothetis, N. K. (2003) The underpinnings of the BOLD functional magnetic resonance imaging signal. *Journal of Neuroscience* 23(10):3963–71. [CK]

Logothetis, N. K. & Wandell, B. A. (2004) Interpreting the BOLD signal. *Annual Review of Physiology* 66:735–69. [CK]

Long, G. (1980) Iconic memory: A review and critique of the study of short term visual storage. *Psychological Bulletin* 88:785–820. [aNB]

(1985) The variety of visual persistence: Comments on Yeomas and Irwin. *Perception and Psychophysics* 38:381–85. [aNB]

Luck, S. J. & Vogel, E. K. (1997) The capacity of visual working memory for features and conjunctions. *Nature* 390:279–81. [aNB]

Lumer, E. D., Friston, K. J. & Rees, G. (1998) Neural correlates of perceptual rivalry in the human brain. *Science* 280:1930–34. [CS]

Lumer, E. D. & Rees, G. (1999) Covariation of activity in visual and prefrontal cortex associated with subjective visual perception. *Proceedings of the National Academy of Sciences USA* 96(4):1669–73. [rNB, CS]

Lycan, W. (1996) *Consciousness and experience.* MIT Press. [arNB]

(2004) The superiority of HOP to HOT. In: *Higher-order theories of consciousness*, ed. R. Gennaro, pp. 93–114. John Benjamins. [rNB]

Mack, A. & Rock, I. (1998) *Inattentional blindness.* MIT Press. [CEI]

Mackay, D. J. C. (2004) *Information theory, inference, and learning algorithms.* Cambridge University Press. [OJH]

Macmillan, N. A. & Creelman, C. D. (1991) *Detection theory: A user's guide.* Cambridge University Press. [CS, MSn]

Marcel, A. J. (1993) Slippage in the unity of consciousness. In: *Experimental and theoretical studies of consciousness: The Ciba Foundation symposium*, ed. G. R. Bock & J. Marsh, p. 174. Wiley. [HL]

Mathiesen, C., Caesar, K., Akgoren, N. & Lauritzen, M. (1998) Modification of activity-dependent increases of cerebral blood flow by excitatory synaptic activity and spikes in rat cerebellar cortex. *Journal of Physiology* 512(Pt 2):555–66. [CK]

McCarty, M. E., Clifton, R. K. & Collard, R. R. (1999) Problem solving in infancy: The emergence of an action plan. *Developmental Psychology* 35:1091–101. [CEI]

McConkie, G. W. & Rayner, K. (1975) The span of the effective stimulus during a fixation in reading. *Perception and Psychophysics* 17:578–86. [arNB, SK]

McConkie, G. W. & Zola, D. (1979) Is visual information integrated across successive fixations in reading? *Perception and Psychophysics* 25:221–24. [arNB]

McDermott, D. (2001) *Mind and mechanism.* MIT Press. [DM]

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

McElree, B. (2006) Accessing recent events. In: *The psychology of learning and motivation, vol.* 46, ed. B. H. Ross, pp 155–201. Academic Press. [aNB]
McGinn, C. (1991) *The problem of consciousness.* Oxford University Press. [aNB]
McLaughlin, B. P. (1992) The rise and fall of British emergentism. In: *Emergence or reduction?: Prospects for nonreductive physicalism*, ed. A. Beckermann, H. Flohr & J. Kim, pp. 49–93. De Gruyter. [aNB]
McRae, K., Butler, B. E. & Popiel, S. J. (1987) Spatiotopic and retinotopic components of iconic memory. *Psychological Research* 49(4):221–27. [RL]
Meadows, J. C. (1974) Disturbed perception of colours associated with localized cerebral lesions. *Brain* 97:615–32. [RWK]
Merikle, P., Joordens, S. & Stolz, J. (1995) Measuring the relative magnitude of unconscious influences. *Consciousness and Cognition* 4:422–39. [MSn]
Merker, B. (2007) Consciousness without a cerebral cortex: A challenge for neuroscience and medicine. *Behavioral and Brain Sciences* 30:63–81. [aNB, CEI]
Metzinger, T. (2003) *Being no one.* MIT Press. [aNB]
Miller, G. (1956) The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63:81–97. [aNB]
Milner, A. D. & Goodale, M. A. (1995) *The visual brain in action.* Oxford University Press. [CEI]
(2008) Two visual systems re-viewed. *Neuropsychologia* 46:774–85. [rNB]
Minsky, M. (1986) *The society of mind.* Simon & Schuster. [aNB]
Montague, D. P. F. & Walker-Andrews, A. S. (2001) Peekaboo: A new look at infants' perception of emotion expressions. *Developmental Psychology* 37:826–38. [CEI]
Most, S. B., Chun, M. M., Widders, D. M. & Zald, D. H. (2005a) Attentional rubbernecking: Cognitive control and personality in emotion-induced blindness. *Psychonomic Bulletin and Review* 12:654–61. [CEI]
Most, S. B., Scholl, B. J., Clifford, E. R. & Simons, D. J. (2005b) What you see is what you set: Sustained inattentional blindness and the capture of awareness. *Psychological Review* 112:217–42. [CEI]
Most, S. B., Simons, D. J., Scholl, B. J., Jimenez, R., Clifford, E. & Chabris, C. F. (2001) How not to be seen: The contribution of similarity and selective ignoring to sustained inattentional blindness. *Psychological Science* 12:9–17. [CEI]
Muir, D. & Lee, K. (2003) The still-face effect: Methodological issues and new applications. *Infancy* 4:483–91. [CEI]
Mullen, K. T. (1992) Colour vision as a post-receptoral specialization of the central visual field. *Vision Research* 31(1):119–30. [rNB]
Naccache, L. (2006a) *Le nouvel inconscient. Freud, Christophe Colomb des neurosciences.* Odile Jacob. [LN]
(2006b) Psychology. Is she conscious? *Science* 313(5792):1395–96. [aNB, LN, CS]
Nachev, P. & Husain, M. (2007) Comment on "Detecting awareness in the vegetative state." *Science* 315(5816):1221; author reply: 1221. [CS]
Nagel, T. (1974) What is it like to be a bat? *The Philosophical Review* 83(4):435–50. [aNB]
Nakayama, K. (1990) The iconic bottleneck and the tenuous link between early visual processing and perception. In: *Vision: Coding and efficiency*, ed. C. Blakemore, pp. 411–22. Cambridge University Press. [aNB]
Nakayama, K., He, Z. J. & Shimojo, S. (1995) Visual surface representation: A critical link between lower-level and higher-level vision. In: *Visual cognition: An invitation to cognitive science*, vol. 2, ed. S. M. Kosslyn & D. Osherson, pp. 1–70. MIT Press. [aNB]
Nature, Editorial (2006) Flickers of consciousness. *Nature* 443(7108):121–22. [aNB]
Noë, A. (2002) Is the visual world a grand illusion? *Journal of Consciousness Studies* 9(5–6):1–12. [aNB]
(2004) *Action in perception.* MIT Press. [aNB]
(2005) What does change blindness teach us about consciousness? *Trends in Cognitive Sciences* 9(5):218. [aNB]
Noë, A. & Thompson, E. (2004) Are there neural correlates of consciousness? *Journal Of Consciousness Studies* 11(1):3–28. [aNB]
O'Craven, K. & Kanwisher, N. (2000) Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience* 12:1013–23. [aNB]
O'Regan, J. K. & Noë, A. (2001) A sensorimotor approach to vision and visual consciousness. *Behavioral and Brain Sciences* 24:883–975. [aNB]
O'Regan, K. (1992) Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology* 46(3):461–88. [aNB]
Oakes, L. M. & Ribar, R. J. (2005) A comparison of infants' categorization in paired and successive presentation familiarization tasks. *Infancy* 7:85–98. [CEI]
Owen, A. M., Coleman, M. R., Boly, M., Davis, M. H., Laureys, S. & Pickard, J. D. (2006) Detecting awareness in the vegetative state. *Science* 313(5792):1402. [aNB, LN, CS, AV]
Owen, A. M., Coleman, M. R., Boly, M., Davis, M. H., Laureys, S., Jolles, D. & Pickard, J. (2007) Response to comments on "Detecting awareness in the vegetative state." *Science* 315(5816):1221c. [AV]
Papineau, D. (1998) Functionalism. In: *Routledge encyclopedia of philosophy*, ed. E. Craig. Routledge. (Web-based encyclopedia.) Retrieved April 22, 2007 from: http://www.rep.routledge.com/article/V015SECT4 [aNB]
(2002) *Thinking about consciousness.* Oxford University Press. [aNB, KB, DP]
Pascual-Leone, A. & Walsh, V. (2001) Fast backprojections from the motion to the primary visual area necessary for visual awareness. *Science* 292(5516):510–12. [aNB, OJH, HL, RL]
Pearson, D. G. (2001) Imagery and the visuo-spatial sketchpad. In: *Working memory in perspective*, ed. J. Andrade. *Psychology Press.* [TB]
Peirce, C. S. (1903) *Collected papers of Charles Sanders Peirce.* Harvard University Press. [aNB]
Persaud, N. & McLeod, P. (2007) Wagering demonstrates subconscious processing in a binary exclusion task. *Consciousness and Cognition.* DOI:10.1016/.concog.2007.05.003. [HL]
Persaud, N., McLeod, P. & Cowey, A. (2007) Post-decision wagering objectively measures awareness. *Nature Neuroscience* 10:257–61. [HL, CS]
Phillips, W. A. (1974) On the distinction between sensory storage and short-term visual memory. *Perception and Psychophysics* 16:283–90. [aNB]
Piattelli-Palmarini, M. (1994) *Inevitable illusions: How mistakes of reason rule our minds.* Wiley. [aNB]
Pins, D. & Ffytche, D. H. (2003) The neural correlates of conscious vision. *Cerebral Cortex* 13:461–74. [aNB]
Pollen, D. A. (1999) On the neural correlates of visual perception. *Cerebral Cortex* 9(1):4–19. [aNB, OJH]
(2003) Explicit neural representations, recursive neural networks and conscious visual perception. *Cerebral Cortex* 13:807–14. [aNB]
(in press) The fundamental requirements for primary visual perception. *Cerebral Cortex*. Advance access published on December 17, 2007; doi:10.1093/Cercor/bhm226. [aNB]
Potter, M. C. (1993) Very short-term conceptual memory. *Memory and Cognition* 21(2):151–61. [aNB]
Prinz, J. J. (2000) A neurofunctional theory of visual consciousness. *Consciousness and Cognition* 9(2):243–59. [arNB]
(2005) A neurofunctional theory of consciousness. In: *Cognition and the brain: Philosophy and the neuroscience movement*, ed. A. Brook & K. Akins, pp. 381–96. Cambridge University Press. [rNB, JP]
(2007) Mental pointing: Phenomenal knowledge without concepts. *Journal of Consciousness Studies* 14(9–10):184–211. [JP]
(forthcoming) *The conscious brain*. Oxford University Press. [JP]
Putnam, H. (1975) The meaning of "meaning." *Minnesota Studies in the Philosophy of Science* 7:131–93. [aNB]
(1981) *Reason, truth and history.* Cambridge University Press. [aNB]
Pylylshyn, Z. (2003) *Seeing and visualizing.* MIT Press. [aNB]
Qiu, J. (2006) Does it hurt? *Nature* 444(9):143–45. [rNB]
Quine, W. V. (1969) Natural kinds. In: *Ontological reality and other essays*, ed. W. V. Quine, pp.114–39. Columbia University Press. [aNB]
Quinn, P. C. (2006) On the emergence of perceptual organization and categorization in young infants: Roles for perceptual process and knowledge access. In: *Child psychology: A handbook of contemporary issues*, 2nd edition, ed. L. Balter & C. Tamis-LeMonda, pp. 109–31. Psychology Press/Taylor & Francis. [CEI]
Quinn, P. C. & Eimas, P. D. (1996) Perceptual organization and categorization in young infants. In: *Advances in infancy research*, vol. *10*, ed. C. Rovee-Collier & L. P. Lipsitt, pp. 1–36. Ablex. [CEI]
(1997) A reexamination of the perceptual-to-conceptual shift in mental representations. *Review of General Psychology* 1:271–87. [CEI]
Quinn, P. C., Eimas, P. D. & Rosenkrantz, S. L. (1993) Evidence for representations of perceptually similar natural categories by 3- and 4-month-old infants. *Perception* 22:463–75. [CEI]
Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A. & Shulman, G. L. (2001) A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America* 98(2):676–82. [rNB, RM]
Raymond, J. E., Shapiro, K. L. & Arnell, K. M. (1992) Temporary suppression of visual processing during RSVP: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance* 18:849–60. [CEI]
Rayner, K. & Bertera, J. H. (1979) Reading without a fovea. *Science* 206(4417):468–69. [LN]
Reddy, L., Wilken, P. & Koch, C. (2004) Face-gender discrimination is possible in the near absence of attention. *Journal of Vision* 4:106–17. [CK]
Rees, G., Kreiman, G. & Koch, C. (2002a) Neural correlates of consciousness in humans. *Nature Reviews Neuroscience* 3(4):261–70. [aNB, CS]
Rees, G., Wojciulik, E., Clarke, K., Husain, M., Frith, C. & Driver, J. (2000) Unconscious activation of visual cortex in the damaged right hemisphere of a parietal patient with extinction. *Brain* 123(Pt 8):1624–33. [aNB, CK, CS]
(2002b) Neural correlates of conscious and unconscious vision in parietal extinction. *Neurocase* 8:387–93. [aNB, CK]

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

Reiss, J. E. & Hoffman, J. E. (2006) Object substitution masking interferes with semantic processing: Evidence from event-related potentials. *Psychological Science* 17:1015–20. [CEI]

Rensink, R. A., O'Regan, J. K. & Clark, J. J. (1997) To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science* 8:368–73. [aNB]

Rosenberg, A. (2005) *Philosophy of science: A contemporary introduction*. Routledge. [rNB]

Rosenthal, D. (2005) *Consciousness and mind*. Oxford University Press. [aNB, DMR]

Ross, W. D., (1955) Aristotle: On dreams. In: *Aristotle. Parva naturalia*, ed. W. D. Ross, pp. 175–223. Clarendon Press. [aNB]

Ross, W. D., ed. (1961) *Aristotle's De anima*. Clarendon Press. [aNB]

Rousselet, G., Fabre-Thorpe, M. & Thorpe, S. (2002) Parallel processing in high-level visual scene categorization. *Nature Neuroscience* 5:629–30. [aNB]

Rousselet, G. A., Thorpe, S. J. & Fabre-Thorpe, M. (2004a) How parallel is visual processing in the ventral pathway? *Trends in Cognitive Science* 8:363–70. [TB]

(2004b) Processing of one, two, or four natural scenes in humans: The limits of parallelism. *Vision Research* 44: 877–94. [TB]

Ruff, C. C., Kristjansson, A. & Driver, J. (in press) Readout from iconic memory involves similar neural processes as selective spatial attention. *Psychological Science*. [CS]

Sackeim, H. A., Nordlie, J. W. & Gur, R. C. (1979) A model of hysterical and hypnotic blindness: Cognition, motivation, and awareness. *Journal of Abnormal Psychology* 88(5):474–89. [CK]

Scholte, H. S., Witteveen, S. C., Spekreijse, H. & Lamme, V. A. F. (2006) The influence of inattention on the neural correlates of scene segmentation. *Brain Research* 1076:106–15. [VAFL]

Searle, J. (1992) *The rediscovery of the mind*. MIT Press. [aNB]

(2005) Consciousness: What we still don't know. *The New York Review of Books* 52(1)36–39. [aNB]

Sergent, C., Baillet, S. & Dehaene, S. (2005) Timing of the brain events underlying access to consciousness during the attentional blink. *Nature Neuroscience* 8(10):1391–400. [VAFL, CS, LN]

Sergent, C. & Dehaene, S. (2004) Is consciousness a gradual phenomenon? Evidence for an all-or-none bifurcation during the attentional blink. *Psychological Science* 15:720–28. [arNB, CS]

Shanahan, M. P. (2005) Global access, embodiment, and the conscious subject. *Journal of Consciousness Studies* 12(12):46–66. [MSh]

(in press) A spiking neuron model of cortical broadcast and competition. *Consciousness and Cognition*. [MSh]

Shanahan, M. P. & Baars, B. J. (2005) Applying global workspace theory to the frame problem. *Cognition* 98(2):157–76. [MSh]

Shoemaker, S. (1981) Some varieties of functionalism. *Philosophical Topics* 12:93–119. [aNB]

Siegel, S. (2006a) The contents of perception. In: *The Stanford encyclopedia of philosophy*, ed. E. N. Zalta, pp. Online publication. Available at: http://plato.stanford.edu/archives/spr2006/entries/perception-contents/ [aNB]

(2006b) Which properties are represented in perception? In: *Perceptual experience*, ed. T. Gendler & J. Hawthorne, pp. 481–503. Oxford University Press. [aNB]

Silvanto, J., Cowey, A., Lavie, N. & Walsh, V. (2005a) Striate cortex (V1) activity gates awareness of motion. *Nature Neuroscience* 8(2):143–44. [aNB]

Silvanto, J., Lavie, N. & Walsh, V. (2005b) Double dissociation of V1 and V5/MT activity in visual awareness. *Cerebral Cortex* 15:1736–41. [aNB, VAFL]

Simons, D. J. (1997) Change blindness. *Trends in Cognitive Sciences* 1:261–67. [aNB]

Simons, D. J. & Ambinder, M. (2005) Change blindness: Theory and consequences. *Current Directions in Psychological Science* 14:44–48. [aNB]

Simons, D. J. & Chabris, C. F. (1999) Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception* 28:1059–74. [CEI]

Simons, D. J. & Rensink, R. (2005a) Change blindness: Past, present and future. *Trends in Cognitive Sciences* 9(1):16–20. [aNB]

(2005b) Change blindness, representations and consciousness: Reply to Noë. *Trends in Cognitive Sciences* 9(5):219. [aNB]

Singer, W. (1995) Development and plasticity of cortical processing architectures. *Science* 270:758–64. [VAFL]

Slater, A. & Lewis, M., eds. (2007) *Introduction to infant development*, 2nd edition. Oxford University Press. [CEI]

Slater, R., Cantarella, A., Gallella, S., Worley, A., Boyd, S., Meek, J. & Firzgerald, M. (2006) Cortical pain responses in human infants. *The Journal of Neuroscience* 26(14):3662–66. [aNB]

Sligte, I. G., Scholte, H. S. & Lamme, V. A. F. (2008) Are there multiple visual short-term memory stores? *Plos One* 3(2):1–9. [AB, arNB, TB, RL, JKO, MSp]

Smith, D. W. (1986) The structure of (self-)consciousness. *Topoi* 5:149–56. [aNB]

Snodgrass, M. (2002) Disambiguating conscious and unconscious inferences: Do exclusion paradigms demonstrate unconscious perception? *American Journal of Psychology* 115:545–80. [rNB, MSn]

Snodgrass, M., Bernat, E. & Shevrin, H. (2004) Unconscious perception: A model-based approach to method and evidence. *Perception and Psychophysics* 66:846–67. [MSn]

Snodgrass, M. & Shevrin, H. (2006) Unconscious inhibition and facilitation at the objective detection threshold. *Cognition* 101(1):43–79. [aNB, MSn]

Solomon, J. (2004) The effects of spatial cues on visual sensitivity. *Vision Research* 44:1209–16. [RWK]

Sommerville, J. A. & Woodward, A. L. (2005) Pulling out the intentional structure of action: The relation between action processing and action production in infancy. *Cognition* 95:1–30. [CEI]

Sommerville, J. A., Woodward, A. L. & Needham, A. (2005) Action experience alters 3-month-old infants' perception of others' actions. *Cognition* 96:B1–B11. [rNB]

Sosa, E. (2002) Privileged access. In: *Consciousness: New philosophical perspectives*, ed. Q. Smith & A. Jokic, pp. 273–94. Oxford University Press. [aNB]

Spener, M. (under review) Phenomenal adequacy and introspective evidence. [MSp]

Sperber, D. (2001) In defense of massive modularity. In: *Language, brain and cognitive development: Essays in honor of Jacques Mehler*, ed. E. Dupoux, pp. 47–57. MIT Press. [aNB]

Sperling, G. (1960) The information available in brief visual presentations. *Psychological Monographs: General and Applied* 74(11, Whole No. 498):1–29. [Whole issue.] [KB, arNB, TB, AC, RG, GH, CK, SK, DM, LN, JKO, DP, DMR, CS, MSh, MSp, MT, RVG]

Srinivasan, R., Russell, D. P., Edelman, G. M. & Tononi, G. (1999) Increased synchronization of neuromagnetic responses during conscious perception. *Journal of Neuroscience* 19(13):5435–48. [CS]

Stern, D. N. (1974) *Mother and infant at play: The dyadic interaction involving facial, vocal, and gaze behavior*. Wiley. [CEI]

Stoerig, P. (2001) The neuroanatomy of phenomenal vision: A psychological perspective. In: *Cajal and consciousness: Scientific approaches to consciousness on the centennial of Ramon y Cajal's Textura*, ed. P. C. Marijuan, pp. 176–94. Annals of the New York Academy of Sciences. [aNB]

Strawson, G. (2003) What is the relation between an experience, the subject of the experience, and the content of the experience? *Philosophical Issues* 13(1):279–315. [aNB]

Supèr, H., Spekreijse, H. & Lamme, V. A. F. (2001a) A neural correlate of working memory in the monkey primary visual cortex. *Science* 293:120–24. [arNB]

(2001b) Two distinct modes of sensory processing observed in monkey primary visual cortex (V1). *Nature Neuroscience* 4(3):304–10. [aNB, VAFL, RL]

Taddio, A., Goldbach, M., Ipp, M., Stevens, B. & Koren, G. (1995) Effect of neonatal circumcision on pain responses during vaccination in boys. *The Lancet* 345(8945):291–92. [rNB]

Tavalaro, J. & Tayson, R. (1997) *Look up for yes*. Kodansha America. [AV]

The Multi-Society Task Force on PVS (1994) Medical aspects of the persistent vegetative state (1). *New England Journal of Medicine* 330(21):1499–508. [AV]

Theoret, H., Kobayashi, M., Ganis, G., Di Capua, P. & Pascual-Leone, A. (2002) Repetitive transcranial magnetic stimulation of human area MT/V5 disrupts perception and storage of the motion aftereffect. *Neuropsychologia* 40(13):2280–87. [aNB]

Thomas, N. (1999) Are theories of imagery theories of imagination? *Cognitive Science* 23(2):207–45. [aNB]

Thompson, K. G. & Schall, J. D. (1999) The detection of visual signals by macaque frontal eye field during masking. *Nature Neuroscience* 2:283–88. [VAFL]

Tong, F., Nakayama, K., Vaughan, J. T. & Kanwisher, N. (1998) Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron* 21(4):753–59. [aNB]

Tononi, G. & Edelman, G. M. (1998) Consciousness and complexity. *Science* 282:1846–51. [aNB]

Treisman, A., Russell, R. & Green, J. (1975) Brief visual storage of shape and movement. In: *Attention and performance*, ed. S. Dornic & P. Rabbitt, pp. 699–721. Academic Press. [arNB]

Treue, S. & Martînez Trujillo, J. C. (1999) Feature-based attention influences motion processing gain in macaque visual cortex. *Nature* 399:575–79. [RWK]

Tronick, E. Z. (1989) Emotions and emotional communication in infants. *American Psychologist* 44:112–19. [CEI]

Tsao, D. Y., Tootell, R. B. H. & Livingstone, M. S. (2006) A cortical region consisting entirely of face-selective cells. *Science* 311(5761):670–74. [aNB]

Tse, P. U., Martinez-Conde, S., Schlegel, A. A. & Macknik, S. L. (2005) Visibility, visual awareness and visual masking of simple unattended targets are confined to areas in the occipital cortex beyond human V1/V2. *Proceedings of the National Academy of Sciences USA* 102:17178–83. [arNB]

Tsuchiya, N. & Koch, C. (2008) The relationship between consciousness and attention. In: *The neurology of consciousness*, ed. S. Laureys & G. Tononi. Elsevier. [CK]

Uebel, T. (2006) Vienna Circle. *The Stanford Encyclopedia of Philosophy Fall 2006 Edition*, ed. E. N. Zalta. Online publication. Available at: http://plato.stanford.edu/archives/fall2006/entries/vienna-circle [aNB]

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press

VanRullen, R. V. & Thorpe, S. J. (2001) Is it a bird? Is it a plane? Ultra-rapid visual categorization of natural and artificial objects. *Perception* 30:655–68. [TB]
Verfaellie, M., Milberg, W. P., McGlinchey-Berroth, R., Grande, L. & D'Esposito, M. (1995) Comparison of cross-field matching and forced-choice identification in hemispatial neglect. *Neuropsychology* 9:427–34. [rNB]
Visser, T. & Merikle, P. (1999) Conscious and unconscious processes: The effects of motivation. *Consciousness and Cognition* 8:94–113. [MSn]
Vogel, E. K., Woodman, G. F. & Luck, S. J. (2001) Storage of features, conjunctions and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance* 27:92–114. [aNB]
Volpe, B. T., LeDoux, J. E. & Gazzaniga, M. S. (1979) Information processing of visual stimuli in an extinguished field. *Nature* 282:722–24. [rNB]
Vorberg, D., Mattler, U., Heinecke, A., Schmidt, T. & Schwarzbach, J. (2003) Different time courses for visual perception and action priming. *Proceedings of the National Academy of Sciences USA* 100(10):6275–80. [CS]
Walker-Andrews, A. S. (1998) Emotions and social development: Infants' recognition of emotions in others. *Pediatrics* 102:1268–71. [CEI]
Weiskrantz, L. (1986) *Blindsight: A case study and implications.* Oxford University Press. [HL]
(1997) *Consciousness lost and found: A neuropsychological exploration*. Oxford University Press. [aNB, CS]
Weiskrantz, L., Cowey, A. & Hodinott-Hill, I. (2002) Prime-sight in a blindsight subject. *Nature Neuroscience* 5:101–102. [aNB]
Wilhelm, B., Jordan, M. & Birbaumer, N. (2006) Communication in locked-in syndrome: Effects of imagery on salivary pH. *Neurology* 67(3):534–35. [AV]
Wilken, P. (2001) *Capacity limits for the detection and identification of change: Implications for models of visual short-term memory.* University of Melbourne Press. [aNB]
Williams, M. A., Dang, S. & Kanwisher, N. G. (2007) Only some spatial patterns of fMRI response are read out in task performance. *Nature Neuroscience* 10(6):685–86. [LN]
Wolfe, J. (1999) Inattentional amnesia. In: *Fleeting memories*, ed. V. Coltheart, pp. 71–94. MIT Press. [aNB]
Woodman, G. F. & Luck, S. J. (2003) Dissociations among attention, perception, and awareness during obect-substitution masking. *Psychological Science* 14:605–11. [CEI]
Xu, Y. (2002) Encoding colour and shape from different parts of an object in visual short-term memory. *Perception and Psychophysics* 64:1260–80. [aNB]
Xu, Y. & Chun, M. M. (2006) Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature* 440:91–95. [aNB]
Yang, W. (1999) *Lifetime of human visual sensory memory: Properties and neural substrate.* New York University Press. [aNB]
Zeki, S. (2001) Localization and globalization in conscious vision. *Annual Review of Neuroscience* 24:57–86. [aNB]
Zeki, S. & Bartels, A. (1999) Toward a theory of visual consciousness. *Consciousness and Cognition* 8(2):225–59. [aNB, OJH]
Zeki, S. & ffytche, D. H. (1998) The Ridoch syndrome: Insights into the neurobiology of conscious vision. *Brain* 121(Pt 1):25–45. [aNB, OJH]
Zihl, J., von Cramon, D. & Mai, N. (1983) Selective disturbance of movement vision after bilateral brain damage. *Brain* 106:313–40. [aNB]

https://doi.org/10.1017/S0140525X07003056 Published online by Cambridge University Press