Philosophia (2008) 36:509–529
DOI 10.1007/s11406-008-9126-8

The Scientific Untraceability of Phenomenal Consciousness

Hilla Jacobson-Horowitz

Received: 14 December 2007 / Accepted: 4 March 2008 /
Published online: 27 March 2008

Abstract It is a common conviction among philosophers who hold that phenomenal properties, qualia, are distinct from any cognitive, intentional, or functional properties, that it is possible to trace the neural correlates of these properties. The main purpose of this paper is to present a challenge to this view, and to show that if "non-cognitive" phenomenal properties exist at all, they lie beyond the reach of neuroscience. In the final section it will be suggested that they also lie beyond the reach of psychology, so that they may be said to lie beyond the reach of science.

Keywords Consciousness · Phenomenal consciousness · Qualia · Neural correlate · NCC · The neural correlate of consciousness · Phenomenism · Non-cognitivism

Introduction

It is a pretty common conviction among philosophers who endorse a non-cognitivist (or "phenomenist") view of phenomenal consciousness—ones who hold that phenomenal properties, qualia, are "distinct from any cognitive, intentional, or functional propert[ies]" (Block 1995, p. 231)—that it is possible to trace the neural correlates of these properties.[1,2] Attempts to find the neural correlates of phenomenal properties—to find the minimal set of neural properties whose instantiations give rise to phenomenal consciousness - have become widespread (see, e.g., Crick and Koch 1990, 1995; Bogen 1998; Edelman 1989; Newman 1997; Llinas et al. 1994; Logothetis and Schall 1989), and some of those philosophers interpret such attempts

[1]On the non-cognitivist or phenomenist conception, see also Block 1996, 2003.

[2]The expression "the neural correlate of consciousness" (often abbreviated as "NCC") might sound as if it presupposes dualism, for correlation is only possible between distinct entities, but most of the philosophers and researchers writing on the issue of NCC are physicalists. For this reason, perhaps the expression "the neural realizer of consciousness" would be more appropriate. Nevertheless, I follow the literature in using the "correlate" terminology.

H. Jacobson-Horowitz (✉)
Department of Philosophy, Ben-Gurion University of the Negev, P.O. Box 653,
Beersheba 84105, Israel
e-mail: hillaj@bgu.ac.il

as just this: attempts to trace the neural correlates of *non-cognitive* phenomenal consciousness (see, e.g., Block 1994a, 1995, 1998, 2005 and unpublished; Chalmers 1996, 1998, 2000).[3] This conviction is shared by some of the scientists who take part in this endeavor themselves (though they are usually not as explicit on this issue as the philosophers).[4] The main purpose of this paper is to present a challenge to the view that non-cognitive phenomenal properties are neurologically traceable. Some philosophers, as well as scientists, believe that we can only find the neural correlates of cognitive attitudes that are associated with non-cognitive phenomenally conscious properties (if there are such properties at all),[5] but to the best of my knowledge no systematic and comprehensive case against the possibility of tracing the neural correlate of non-cognitive phenomenal consciousness has been made, and this is what I will try to do in what follows.[6]

The issue may be presented by considering the main lines of the debate regarding the existence of non-cognitive phenomenal properties between Dennett and some of his critics.[7] Dennett has discussed some pathological, abnormal or otherwise special cases (both real and fictional), in which, so he has argued, it is undecidable whether or not a certain phenomenal event occurs. From this (epistemic) undecidability he derives the conclusion that there is no fact of the matter regarding the question of the occurrence of such events. Dennett's arguments have been criticized by, among others, Block (1993) and Flanagan (1992), who argue that although it is true that in each one of the cases Dennett discusses, taken on its own, the evidence does not favor either of the alternatives, we should be able to decide between them by using knowledge about phenomenal/neural correlations gained by studying "paradigmatic" or "unproblematic" cases of phenomenal consciousness. I shall attempt to show that such a strategy cannot work, due to an essential obstacle which stands in the way of gaining the required knowledge in the paradigmatic cases, an obstacle which concerns the fact that it is impossible to identify phenomenally conscious features in isolation from some form of cognitive consciousness. My argument, then, aims at establishing that if non-cognitive qualia exist at all, they lie beyond the reach of neuroscience. In the final section I shall suggest that they also lie beyond the reach of psychology, so that we may say that they lie beyond the reach of science.

[3] Chalmers (2004) now endorses a version of representationalism of phenomenal consciousness, but this is a non-reductive form of representationalism which in essence should be classified within the category of non-cognitive views of phenomenal consciousness.

[4] Crick and Koch (1990) claim that one of the functions of consciousness is *to present the result* of various underlying computations, and thus seem to presuppose the non-cognitivist (or the Cartesian) conception of consciousness as a show presented to a spectator in a theater. Commenting on this claim of Crick and Koch, Dennett asks (1991, p. 255): "... but to whom [is consciousness supposed to present this result]? The Queen?.."

[5] Among the scientists who maintain that only the neural correlate of a cognitive form of consciousness (what they call "access to consciousness") can be found are Dehaene and Changeux (2004).

[6] The very notion of the neural correlate of consciousness is not free of difficulties (see, e.g., Noe and Thompson 2004), but I will not pursue this topic here.

[7] I am concerned in this paper with the positions of Dennett and his critics only insofar as they may serve to highlight the issue of the neural traceability of phenomenal consciousness, and so I will ignore various aspects and subtleties of their works which are irrelevant to this issue.

I believe that these claims that I advance may be taken to contribute to the (ontological) case against non-cognitive phenomenal properties.[8] But here I will confine myself to epistemological and methodological points, and will not pursue what may be viewed as their ontological implication. For the sake of argument, I shall assume that phenomenal properties—the ways things look or sound, the way pain feels, etc.—are indeed non-cognitive.[9] (Hereafter, unless context indicates otherwise, the terms "phenomenal consciousness", "phenomenal properties", "qualia", and so forth, should be understood in a theory-laden sense, so as to imply that their referents are non-cognitive.) The claim that phenomenal properties are non-cognitive does not entail that phenomenal properties are epiphenomenal, or that they do not participate in cognitive processes (I shall later say a few words on this point); all the more so, it does not entail that they do not correlate with cognitive properties. Indeed, I shall not assume any of these.

Dennett's Argument for the Indeterminacy of the Occurrence of "Brief Flickers of Consciousness"

Dennett's argument is based on a series of examples, and I shall concentrate on one of them—that of the phenomenon of metacontrast (see Fehrer and Raab 1962). Subjects in the metacontrast experiment are exposed to a brief (about 30 ms) visual presentation of a disc, which is immediately followed by a presentation of a ring that fits closely outside the space where the disc was displayed. The subjects report to have seen only the ring. (If they are exposed for the same short duration only to the first stimulus, they do report seeing the disc.) Nevertheless, if they are asked to guess whether there were two stimuli, the subjects do much better than chance.

Dennett argues that it is possible to account for this phenomenon by either one of two theories. According to the "Stalinesque" theory, the second stimulus somehow prevents conscious experience of the first stimulus. That is, the brain process initiated by the arrival of information about the disc, information that would otherwise lead to a conscious experience of the disc, is interrupted by the arrival of information about the ring, so that this expected experience does not occur. According to the "Orwellian" theory, on the other hand, processing the first stimulus does result in a conscious event, yet the arrival of the ring stimulus initiates an immediate obliteration of this conscious experience from memory.

It seems that we are facing two distinct theories, which differ as to whether the subject had a brief flicker of consciousness (i.e., awareness of the disc) which she does not remember, or didn't have at all. But Dennett argues that the accounts given by the two theories are not really distinct. The difference between the theories "is a difference that makes no difference" (Dennett 1991, p. 125). Both theories are

[8] Though I share their rejection of the notion of (non-cognitive) phenomenal consciousness, I disagree with current cognitivists about which cognitive characteristics should figure in an account of phenomenal consciousness. The resources they appeal to are too meager. See Jacobson-Horowitz 2003.

[9] Obvious examples of phenomenally conscious states are perceptions, feelings and emotions. Some philosophers hold that propositional attitudes also have phenomenal aspects (see, e.g. Block 1994a, Flanagan 1992; Jackendoff 1987; see also note 27 below). I shall focus here on perceptions.

consistent with all data—the data we have at present and the data we may have in the future—and thus differ about nothing which is in principle discernible. To begin with, both account for the subject's reports of what she sees: "One theory says they are innocently mistaken, while the other says they are accurate reports of experienced mistakes." (pp. 124–25) The subject's verbal behavior then, is neutral with respect to the two theories. (The same is true, of course, with respect to the subject's non-verbal behavior.) A more disputable claim of Dennett (p. 125) is that the rival theories agree about what happens in the subject's brain. Suppose that neuroscientists discover the exact "time of arrival" of any content anywhere in the subject's nervous system, and that they can follow its track in the brain. Still, Dennett and Kinsbourne argue, since there is no disagreement between the two theories as to where and when the discriminations of the content "disc" and of the content "ring" occur—the (alleged) disagreement concerns the question whether the content "disc" occurs "inside or outside the charmed circle of consciousness" (Dennett and Kinsbourne 1992, p. 193), such findings cannot decide between the two theories. The claim that neuroscience cannot decide between the rival theories will be the focus of the following two sections.

It might be natural to think that what matters most for deciding between the rival theories—what has the decisive authority—is truly "inside information", that is, information gained from the first-person point of view. But Dennett claims that "... the inability to distinguish these two possibilities does not just apply to the outside observers who might be supposed to lack some private data to which the subject had 'privileged access'." (p. 123) Even you, the subject, "could not discover anything in the experience from your own first-person perspective that would favor one theory over the other..." (*ibid*.) Both theories are consistent with everything the subject may think or remember: she cannot "tell the difference between misbegotten experiences and immediately misremembered experiences." (p. 125)[10]

Thus, Dennett takes the Orwellian theory and the Stalinesque theories to agree on all possible evidence: on all possible neurological, behavioral and introspective data. Now is it possible that this epistemic tie between the two theories is compatible with there being a real difference between them—with there being "brute facts of the matter" unreachable neither by an external observer nor by the subject herself? Dennett rejects such a possibility as far as the "arena of subjectivity" is concerned. He holds that there could be no such unreachable *subjective* facts, and thus that there is no difference between the rival theories in question. Dennett acknowledges this reasoning to be verificationist or operationalist, and argues that in the arena of subjectivity such reasoning makes manifest good sense (p. 132). One reason he suggests for this contention has to do with his denial of the reality/appearance distinction with respect to subjectivity, when "appearance" seems to mean

[10] Dennett (1991, p. 123) also claims that "the experience would 'feel the same' on either account", and Block (1993, p. 189) denies this. Clearly, Dennett cannot take it for granted that the two cases feel the same, in the phenomenal sense of "feel" (exactly as Block cannot take it for granted that the cases do *not* feel the same), on pain of begging the question. But in his argument to the effect that there is no fact of the matter regarding the occurrence of "brief flickers of consciousness" Dennett uses "feel the same" as more or less synonymous with "is indistinguishable by the subject". In this sense, his claim that "the experience would 'feel the same' on either account" should not be controversial.

"appearance to the subject".[11,12] Dennett's conclusion, then, is that there is no fact of the matter regarding the occurrence of "brief flickers of consciousness", so brief as to leave no traces on the subject's memory, nor influence her cognitive life in any other way. Thus, the notion of a non-cognitive phenomenally conscious event, that notion which allows for the possibility of conscious events which leave no cognitive traces, is undermined.[13]

Many commentators have pointed our alleged flaws in Dennett's argument. Here I will present one trend of criticism, the prominent advocates of which are Block (1993) and Flanagan (1992). The discussion of this trend will serve as a basis for the presentation of a general argument for the neurological untraceability of phenomenal consciousness. The trend of criticism in question does not call into question Dennett's verificationist reasoning, but rather objects to Dennett's intermediary conclusion concerning the unavailability of evidence that can decide between the two alternative theories (a contention which expresses the prevalence of the "initial conditions" that are supposed to allow for the verificationist move). More specifically, this attack is directed at the claim that neuroscience will never be able to reach a warranted conclusion as to which of the two hypotheses that account for phenomena like metacontrast is correct. Thus, Block and Flanagan argue that it is in principle possible to neurologically distinguish between the rival hypotheses in the "problematic" cases Dennett discusses, from using knowledge gained by studying "normal" (Block's term) or "paradigmatic" (Flanagan's term) cases involving phenomenal consciousness. These unproblematic cases involve no perceptual or memory tricks, and there is no controversy over what contents the subject is conscious of in each of them. Both Block and Flanagan suggest that by studying such cases we may come to know that paradigmatic conscious events have some salient set of neural properties (they refer, as an example, to Crick and Koch's conjecture that correlates phenomenal consciousness with 40–75 Hz oscillation in the cerebral cortex (Crick and Koch 1990)), and examine whether these properties are involved in the representation of the relevant stimuli (like the disc stimulus in metacontrast) in the problematic cases. If, and only if, these properties are involved in this representation, we will be entitled to conclude that the subject was conscious of the stimulus.

[11] Rejecting this kind of verificationism, Dennett (1991, p. 461) adds, leads to "nonsense" like epiphenomenalism, zombies, indistinguishable inverted spectra, and others.

[12] Dennett's verificationist reasoning certainly raises questions and deserves a wider discussion (for example, the role that is played in this reasoning by the claim that no *third-person* evidence is available should be elucidated). But examining the validity of this reasoning is not important for the epistemological and methodological issues with which this paper is concerned.

[13] According to Dennett, when the question of whether contents are conscious concerns cases like metacontrast, in which—due to the short period of time in which the content is accessible to the subject's cognitive system—it cannot leave cognitive traces, the question presupposes more fine-grained temporal boundaries than the phenomena admit (see Dennett and Kinsbourne 1992, pp. 235–36). This emphasis on the temporal element should not induce us to overlook the general lesson: the principal reason for Dennett's taking the question to be unanswerable concerns the fact that the stimulus does not influence the subject's cognitive life. (We shall later encounter cases which exemplify similar inaccessibility of perceptual contents to subjects' cognitive systems but in which no temporal constraints are in effect.) It has been argued by some commentators that Dennett's rejection of the notion of such (non-cognitive) brief flickers of consciousness in fact positions him as Stalinesque. Again, assessing this claim is not essential for the epistemological issues with which this paper is concerned.

The Neurological Untraceability of Phenomenal Consciousness

For the sake of argument, I will grant that we may have good reasons to know in many cases about the existence of phenomenally conscious states. I will accept the natural view that specific kinds of behavior on the part of the subject, notably verbal behavior, may provide good reasons for attributing specific conscious contents to her in particular cases. I do not mean to suggest that these reasons are conclusive. (For one thing, I do not take introspective judgments to be incorrigible.) I mention this point in order to make it clear that my objection is not that the knowledge Block and Flanagan hold to be attainable is not certain. They would surely not claim that it is. Rather, my claim is that there can be no such "knowledge" which satisfies even minimal standards of scientific knowledge. Block (1994b) and Flanagan (1992, Chapter 4) accuse Dennett of undermining a "straw concept": he assumes that qualia—phenomenal properties—are supposed to be incorrigibly apprehensible, atomic, intrinsic, unanalyzable, essentially private and unreachable by science. But the argument to be presented below assumes none of these features. In particular, the claim that phenomenal consciousness is in principle unreachable by science will not be the argument's premise, but rather its conclusion.

Let us examine, then, the possibility of gaining information about the neural correlates of phenomenally conscious states from paradigmatic cases, cases in which the occurrence of such states is, practically, unquestionable (which is not to say, of course, that it is beyond any doubt). We have good reasons to assume that they occur, and no reasons that suggest that they do not occur. (In the context of the debate between Dennett and his critics, this is actually an examination of the first step of Block's and Flanagan's suggestion to resolve Dennett's special cases by importing neurological knowledge from the paradigmatic cases, the second step being the employment of this information in the special cases.) What characterizes these normal or paradigmatic cases? What is it that makes their involving phenomenal consciousness unquestionable and distinguishes them in this respect from, e.g., Dennett's problematic cases? In some of these paradigmatic cases, we learn from the subject's behavior (which may be either linguistic or non-linguistic) that she believes that she had the experience in question. In these cases, then, in addition to her having the relevant experience, the subject also exhibits "reflective consciousness". A mental state is conscious in this sense if it is accompanied—non-inferentially and non-observationally—by a higher-order thought to the effect that one is in this state (Block 1995 p. 232, 1993 p. 182, and 1994a p. 213). In other paradigmatic cases, what would be taken to show that the subject is phenomenally conscious of the content in question is the fact that her behavior can be best explained on the assumption that this content is used by her in particular ways: as a premise in practical or theoretical reasoning, or in (participating in) rationally controlling her (verbal or non-verbal) behavior. In these cases, then, in addition to her being phenomenally conscious of the content, the subject also exhibits something similar to what Block calls "access consciousness". A state is conscious in this sense if in virtue of one's having it, a representation of its content (1) is freely available for one as a premise in reasoning; (2) is poised for rational control of one's action, and (3) is poised for rational control of one's speech (Block 1993 p. 182, 1994a p. 213, 1995 p. 232). (Davies and Humphreys (1993, p. 11) explicate Block's notion of a content's being *freely* available for a subject's reasoning as the subject's ability to use this content

simply in virtue of being in a state with this content.) There is strong affinity between this notion of access and the notion of access as global broadcasting—as broadcasting contents, that is, to systems such as those engaging with reasoning, planning, problem solving, categorization, and memory (see Baars 1988)—and in his recent writings on the topic Block (2005 and *unpublished*) adopts the latter notion of access. For our purpose, the distinction between these notions is inessential.[14]

Thus, what is common to all these paradigmatic cases is that the phenomenally conscious state is accompanied by cognitive attitudes or processes of some typical sorts. It is the presence of such cognitive attitudes or processes that forms the evidence for the occurrence of phenomenal consciousness and makes it unquestionable. (Let us refer to these patterns of cognitive attitudes and processes as forms of *cognitive consciousness*.)

We should also note that if a case does not involve cognitive consciousness of the relevant forms, its involving phenomenal consciousness would not be unquestionable. Block, an advocate of non-cognitive phenomenal consciousness, presents an array of examples which allegedly involve "normal" phenomenal consciousness, but no—or deficient—cognitive consciousness (see Block 1994a, 1995). An examination of each of these cases would show that if indeed cognitive consciousness is missing or significantly deficient, there is no sound reason to suppose that phenomenal consciousness is present. Consider the case of a person who is engaged in intense thought when suddenly at midnight she realizes that there is now and has been for some time a deafening pounding noise going on. According to Block, the person was phenomenally conscious of the noise all along, but only at midnight did she become access conscious of it (Block 1994a p. 215, 1995 p. 234). (Block believes that such a case refutes also the higher-order thought theory of phenomenal consciousness, since, lacking access to the phenomenal state before midnight, the subject could not form a belief about it.) Yet what reasons support the claim that the person was phenomenally conscious of the noise all along? The fact that the noise was going on surely does not justify this conclusion, since why should we prefer the hypothesis that the person was phenomenally conscious of the noise and had no access to the relevant (phenomenally conscious) content over the hypothesis that there was no such phenomenally conscious

[14] Block's notion of access consciousness is a dispositional notion. A subject who is access conscious of a certain content in his sense need not actually use the content in the ways described above; it suffices that the content be freely *available* for reasoning and *poised* for rational control of action or speech. But I speak of actual uses of contents in those ways, since, of course, if we are interested in identifying phenomenally conscious states, we need to identify actualizations of such dispositions. In spite of this difference, and in spite of the (related) fact that Block's purpose in introducing this notion is to characterize a non-phenomenal type of consciousness whereas my purpose is to suggest an epistemic linkage between phenomenal consciousness and cognitive attitudes and processes of certain sorts, I will use Block's term—"access consciousness" to refer to these cognitive processes and the attitudes they involve. It should also be emphasized that my arguments do not depend on whether the cognitive patterns I speak of exactly match Block's characterizations. In the next section I will further clarify the evidential linkage between phenomenal consciousness and the various kinds of cognitive consciousness. I will only add here that a content's being *freely* available for the cognitive usages in questions in the sense explained above is an important evidential factor. Block correctly points out, with respect to blind sight subjects—who need to be prompted by the experimenter's request in order to guess which stimuli are present in their "blind" visual field—that the contents representing these stimuli are not freely available to them. This illustrates that behavioral data can tell us whether certain contents are *freely* available for cognitive usages or not. (I will discuss the phenomenon of blind sight below.)

content to be accessed in the first place? That the subject realizes that the noise was going on before midnight indeed supports the claim that the information about the noise was received and processed in her brain, but why does it make the hypothesis that this process involved phenomenal consciousness more probable than the hypothesis that it was wholly unconscious?[15,16]

This point is general. What characterizes such cases is that the subject's overt behavior supports the supposition that her cognitive functioning (in the noise example—the pre-midnight functioning) differs, quite significantly, from the cognitive functioning of other subjects in similar perceptual situations (i.e., situations in which other subjects stand in similar external relations to similar objects). The similarity in the perceptual situations might seem to suggest a *prima facie* reason in favor of the assumption of phenomenal similarity; and the fact that the perceptually gained information is received and has some (though minimal) effect on the subject's cognitive system might seem to strengthen this assumption. Yet the cognitive difference testifies that there is enough dissimilarity between the subject's brain state in such special cases and the brain states of subjects in normal cases to undermine the simple analogical inference to phenomenal similarity. The (limited) employment of the perceptual information on the part of the subject does not make the analogy more compelling: there is just no reason to prefer the assumption that the information processed—which, *ex hypothesi*, is processed in manners quite different from the usual ones—is phenomenally conscious, over the assumption that it is not. The assumption that the special cases in question indeed instantiate the relevant phenomenal properties rests, then, on shaky foundations, and this reflects the fact that the absence of evidence for the subject's instantiating the relevant forms of cognitive consciousness (i.e., for the subject's instantiating more or less typical cognitive processing of the relevant contents) is also an absence of evidence for her being phenomenally conscious of the contents in question. (This claim would not be true, of course, were there independent neurological evidence for the subject's being phenomenally conscious of the relevant contents, but the availability of such evidence is what's at stake.)

The claims regarding the evidential linkage between phenomenal consciousness and forms of cognitive consciousness that has so far been presented can be summed up as follows: when we have (good) evidence for the subject's instantiating one of the forms of cognitive consciousness, we have good evidence for the subject's being phenomenally conscious of the relevant contents (those contents which figure in the relevant cognitive attitudes). When we have (good) evidence for the subject's

[15] In his later work Block (1998) expresses awareness to the fact that various hypotheses may account for the phenomenon in question.

[16] Another example that Block (1995, pp. 239–41) takes to involve phenomenal but no (or very deficient) cognitive consciousness is the example of epileptic patients performing some activity (walking home or playing the piano) during seizures. Block accuses Penefield (1975), Searle (1992) and van Gulick (1989) of taking these patients to lack phenomenal consciousness, while actually the evidence shows only that they lack access consciousness (or that their access consciousness is very deficient). There is no reason to doubt, he says, that these subjects are phenomenally conscious: "they do show every sign of *normal sensation*" (Block 1995, p. 239) Again, according to the reasoning I am suggesting, the controversy over this phenomenon is not accidental: in the absence of access consciousness we cannot be in a position to know that the patients are phenomenally conscious.

lacking (wholly or partly) one of those forms of cognitive consciousness, we do not have reasons which justify the assumption that the subject is phenomenally conscious of the relevant contents. (Neither do we have such reasons in cases in which we simply lack evidence for the subject's instantiating any of these forms of cognitive consciousness.) One point that is left vague in this summary of the phenomenal/cognitive evidential linkage concerns cases in which we have (good) evidence for the subject's instantiating one form of cognitive consciousness (e.g., access consciousness) *and* (good) evidence for her lacking another form (e.g., reflective consciousness), with respect to the same content. Such cases will be discussed shortly, and a more comprehensive analysis of the evidential linkage will be suggested in the next section. What has so far been said suffices, I believe, to serve as a basis for the general line of reasoning to which I now turn.

Due to the evidential linkage between phenomenal consciousness and cognitive consciousness, I shall now argue, it is impossible to find the neural correlates of phenomenal consciousness in paradigmatic cases. As we have just seen, paradigmatic phenomenally conscious states, *qua* states whose being phenomenal is unquestionable, are accompanied by cognitive attitudes of certain kinds. Thus, if we trace some neural feature characteristic of such cases, we would not be able to tell (on the basis of examining these cases alone) whether it correlates with the phenomenal state itself, with the accompanying cognitive attitudes, or with both. We face, that is, an indeterminacy of phenomenal and cognitive consciousness from a neurological point of view. Since no external behavior can be taken as evidence for the subject's being in a phenomenally conscious state without also being taken as evidence for her having some such cognitive attitudes, this indeterminacy is doomed to arise in any paradigmatic case, and hence to undermine the possibility of gaining the desired neurological knowledge.

We can thus construct a general argument against the neurological traceability of phenomenal consciousness in the form of a dilemma: either a phenomenally conscious state is accompanied by cognitive consciousness, or it is not accompanied by it. If it is not accompanied by cognitive consciousness, then we have no justification for believing that the case in question indeed involves phenomenal consciousness; if, that is, there is no behavioral indication that the subject has a belief to the effect that he is in the phenomenal state in question, nor is there any indication that he uses the content of this state as a premise in practical or theoretical reasoning etc., then we have no way of knowing about the existence of this phenomenally conscious state. *A fortiori*, it would be impossible to gain any knowledge of its neural correlate in such a case. If, on the other hand, a phenomenal state is accompanied by cognitive consciousness, we would not be able to correlate neural features with the phenomenal state rather than with the accompanying cognitive attitude, or with both.

Block (unpublished (an extended version of Block 2005), p. 16) denies that this linkage between phenomenal and cognitive consciousness implies that we cannot find the neural correlate of phenomenal consciousness: "observed electrons can provide evidence about electrons that cannot in principle be observed, for example electrons that are too distant in space and time (e.g. outside our light cone) to be observed. Why should we suppose matters are any different for consciousness?" But it isn't clear how the analogy is supposed to work—how instances of "observed" (that is, introspected or otherwise accessed) instances of phenomenal consciousness can provide evidence

about unobserved (or unobservable) instances of phenomenal consciousness. Certainly, the idea cannot be that we can make triangulations from what we know about observed phenomenal consciousness to arrive at predictions about unobserved instances of phenomenal consciousness. Can we employ analogical inference of the form: since phenomenal consciousness accompanies some features in the observed cases, it probably accompanies these features also in the unobserved cases? We cannot, because the candidate features are either neural ones or behavioral ones. The former are irrelevant, since we are now looking for a way to find them; the latter are irrelevant since the behavior in question is mediated by cognitive attitudes, which (by definition) are absent from the cases of unobserved phenomenal consciousness. (As to similarity in perceptual situations, we already saw that it is ineffective.)[17]

Posing the dilemma in question is not enough for undermining the traceability of the neural correlate of consciousness. This dilemma involves cases of two kinds: cases of phenomenally conscious states which are accompanied by cognitive consciousness, and cases of phenomenally conscious states which are not accompanied by cognitive consciousness. It might be thought that the difficulty of indeterminacy that is involved in the second horn of the dilemma can be resolved by appealing to cases of a third kind, cases of cognitive consciousness which do not involve phenomenal consciousness. We can use these cases of "isolated" cognitive consciousness, it might be suggested, for finding the neural correlates of the relevant cognitive attitudes, "subtract" them from the neural correlates of the "sum" of phenomenal consciousness and cognitive consciousness involved in the paradigmatic cases, and thus identify the neural correlate of the phenomenally conscious property. Such a strategy, however, cannot work either, even if we avail ourselves of the relevant notion of subtraction.[18] For in the allegedly non-phenomenal cases we must identify the same cognitive attitudes as in the allegedly phenomenal cases, since otherwise we would not know what to subtract in the allegedly phenomenal cases. But here the same evidential linkage between phenomenal and cognitive consciousness strikes again: if we do identify the same cognitive attitudes in both kinds of cases, we would have no reason to think that the allegedly non-phenomenal cases are indeed non-phenomenal (or that the allegedly phenomenal cases are indeed phenomenal).

It is thus no wonder that Block, who argues for a strict distinction between cognitive consciousness and phenomenal consciousness, cannot (as he admits) point out an actual case exemplifying the former but not the latter with respect to the same content. Even if such cases obtain, there is no way for us to know about their existence. Block is thus forced to appeal to the imaginary phenomena of "super

[17] Block also raises in the same context several objections to the idea that the only possible evidence for phenomenal consciousness is introspective evidence or introspective reports. But I by no means assume that this is the only possible evidence for phenomenal consciousness. (On this issue, see also the discussion of SDI in the next section.)

[18] Such a subtraction—referred to in cognitive neuroscience as "cognitive subtraction"—is not free of difficulties. As pointed out by Poeppel (1996), it is threatened by various findings that suggest that neural processing stages may be sensitive to adding or deleting other processing stage. (The discussion of metacontrast below has some affinity to this issue.) Whether or not there are conditions under which cognitive subtraction can nevertheless be justified is a matter of dispute. Poeppel himself answer this question positively; for a negative reply to this question see, e.g., Sartori and Umilta 2000.

blindsight" and "super-duper-blindsight", whose examination may help to further clarify the present point. Blindsight patients are supposed to have "blind" areas in their visual fields, as a result of damage to the frontal cortex. If they are exposed to visual stimuli (say, simple shapes) in their blind visual field, they report to see nothing. The interesting thing about them is that if they are asked to guess what is being "shown" and the experimenter supplies a small set of alternatives, they are able to "guess" quite reliably (much better than chance) which of the alternative stimuli was shown to them.[19] Super blindsight is the fictional phenomenon of subjects who are similar to blindsight subjects but are also capable of prompting themselves at will to "guess" what is in their "blind" field. Super-duper-blindsight is the fictional phenomenon of subjects who differ from normal perceivers only in that they deny seeing the objects in their visual field albeit they know about them everything that a seeing subject knows (i.e., they differ from normal perceivers in having what may be called "a negative reflective belief".)

Block takes these fictional phenomena to exemplify the conceptual possibility of cognitive consciousness (specifically, access consciousness) with no phenomenal consciousness. Since we are now examining the evidential linkage between phenomenal consciousness and cognitive attitudes that typically accompany it, and cases of super-duper-blindsight are closer than the other mentioned cases to exemplifying full typical cognitive consciousness, they are more relevant to our present concern. But even they cannot serve to refute the principled evidential linkage in question, since they do *not* exemplify full typical cognitive consciousness. Even if we can accept the conceptual possibility of such cases, we should deny that we can ever be in a position to know about their occurrence. Given that a person behaves exactly like a normal seeing person in all respects except those which pertain to introspective reports—raises his hand to catch a ball that was thrown to him without a warning, declares that he prefers the color of "this" carpet for his living-room instead of "that one" etc., it seems that *we simply cannot determine* that she is *not* phenomenally conscious of the visual stimulus she is facing. This seems to hold even in the face of what seems to be the strongest counter-evidence, namely that the subject herself insists that she is "blind". This might seem odd: shouldn't we take a subject's word for her own perceptual experiences? But the present debate—the debate concerning cases like super-duper-blindsight—may be said to be situated in the realm of oddities. It is indeed strange and counter-intuitive to claim that a subject who denies seeing is actually seeing; but it is equally strange to claim that a subject who never bumps into furniture, who can play football, and can give us his detailed opinion of an art exhibition, does not see. It is important to emphasize that I by no means argue that we are justified in holding that this subject is phenomenally conscious of the stimuli he is encountering; rather, I argue that we are precluded from justifiably judging in favor of either hypothesis. The same holds (though maybe less strikingly) in cases of somewhat deficient access consciousness (super blindsight cases), and Block himself even raises the possibility that real blindsight subjects (subjects with very limited access consciousness with respect to the relevant stimuli) may actually be phenomenally aware of objects in their alleged blind field. But even if we do not accept

[19] Blindsight was discovered by Lawrence Weiskrants (1986).

these last claims, I hope that this discussion makes it clear that we cannot justifiably judge of cases which exhibit normal cognitive functioning (and in particular normal access consciousness) that they do not involve phenomenal consciousness.

The impossibility of gaining the required neurological knowledge from cases of the three kinds I have discussed can be further clarified by considering the following example. Suppose that neuroscientists attempt to decide between the rival hypotheses (Orwellian and Stalinesque) in the case of metacontrast. What would be more natural for them than to examine, first, the case in which a subject is exposed only to the first stimulus, the disc? After gaining the neurological information about this (paradigmatic) case, they can check whether its typical neural features appear also in the metacontrast case. It might seem that according to whether or not these features are found in the metacontrast case, the scientists would be able to establish whether or not the subject in the metacontrast experiment is phenomenally conscious of the disc.

This tactic, however, cannot work, for surely, not all of the typical neural features of the paradigmatic case (disc only) are present in the metacontrast case. Had they been present, the subject would have had in the metacontrast case all the cognitive attitudes he has in the paradigmatic case, and in particular, he would have believed himself to be phenomenally conscious of the disc; yet he lacks this belief. The typical neurological information gained from the paradigmatic case concerns not only the phenomenally conscious apprehension of the disc, but also a specific cognitive attitude, namely the belief about this apprehension.[20] Thus, on the basis of the information gained from the paradigmatic case alone, we would not be able to isolate the neural correlate of the phenomenal apprehension from the neural correlate of the belief, and, trivially, we wouldn't know the presence of which neural feature is indicative of the phenomenal apprehension.

Can we go the other way around, that is, gain information about the neural correlate of the cognitive attitude that is involved in our paradigmatic case (in this case the reflective belief), and then "subtract" it from the entire typical neurological information? In order to gain this information we would have to appeal to cases in which this belief is isolated—that is, it is not accompanied by a phenomenal state of the same kind. But again, we will never be in a position to judge that such cases occur, since these are cases in which the belief that should be isolated—i.e., the belief that the phenomenal state occurs—is the very same belief whose presence is the best indication for the presence of such phenomenal states. I do not assume that introspective beliefs are incorrigible. The point is that unless we have an independent way of identifying the phenomenal state that an introspective belief is about, then, even if such a belief is false, we cannot be in a position to know that. To assume that we do have such an independent way is to put the cart before the horses: the very purpose of the present move was to point out a way of identifying phenomenal consciousness.

Now consider the following suggestion. Suppose we have a well-corroborated hypothesis ("H_1") that phenomenal consciousness typical of, say, visual experiences

[20] Can the occurrence of this belief be prevented? Plausibly it can (see the discussion of SDI in the next section). But what cannot be prevented is the occurrence of the phenomenal apprehension of the disc with no accompanying cognitive attitudes whatsoever.

("P"), is correlated with a certain neural feature ("N"). (We may suppose that H_1 is Crick's and Koch's hypothesis.) H_1 seems to be well-corroborated since in many paradigmatic cases of visual phenomenal consciousness the neural feature was found to be present, and in no such case was it found to be missing. But it must be admitted that the neural feature was found to be correlated not only with the presence of phenomenal consciousness but also with some typical cognitive attitudes ("C"), for the corroborating cases must be cases in which we have good reasons to assume that P occurs. But this means that the hypothesis we are entitled to at this stage is not H_1, but rather the hypothesis that N correlates with P or that it correlates with C or that it correlates with both. (Let's refer to this latter hypothesis as "H_2"). I argued that it is impossible to isolate the relevant phenomenal states or the relevant cognitive attitudes, and if so, it seems that we cannot discover which of the three disjunctions of H_2 is the right one. But suppose, further, that N was then found to be present in cases that do not involve C. It might be claimed that since N correlates with P or with C or with both, and the new evidence shows that it does not correlate with C, it follows that it correlates with P alone. In short, couldn't the cases of N and non-C count as evidence for H_1?

They could not. For there is no reason to prefer the assumption that cases of N and non-C support the hypothesis of N–P correlation (H_1) over the assumption that these cases are simply refuting counter-examples to the hypothesis that N correlates with P or with C or with both (H_2), and *a fortiori* refuting counter-examples to H_1. After all, in cases of N and non-C we cannot know that P occurs. The only possible reason in favor of the first assumption is *independent* evidence for the occurrence of P in cases of N and non-C (independent, that is, of the occurrence of N), but as was shown above, no such evidence is available. We saw that even in cases in which the perceptual circumstances are the same as those that obtain in the paradigmatic cases and the subject uses the relevant perceptual content in various cognitive operations (though in a limited way, for the cases in question are, *ex hypothesi*, cases of non-C), we have no reason to assume that the subject is phenomenally conscious of this content, rather than that the content is processed unconsciously. The presence of N in both kinds of cases (paradigmatic and non-paradigmatic) can be explained by its realizing an unconscious stage in the processing of this content (a processing that only in cases of one kind results in a phenomenal apprehension of this content).

This might be the place to mention that the non-cognitivist, who claims that phenomenal properties cannot be identified with cognitive or functional roles, is clearly not committed to the claim that phenomenal properties do not play a role in cognition, and specifically, to the claim that they play no role in the production of those cognitive attitudes that serve as evidence for their own instantiations. But this latter claim has no bearing on the argument for the neurological indeterminacy of the phenomenal and the cognitive. For, as we have just seen, given that we identify a neural feature (N) which is present also in cases in which the cognitive processing does not result in the occurrence of the cognitive attitudes which serve as evidence for the occurrence of a phenomenally conscious state, we would never be in a position to ascertain that this neural feature is indeed the correlate of a phenomenal property. The question of whether this neural feature is the occupant of the functional role responsible for this processing (or merely correlates with the occupant of this role) simply makes no difference for this claim.

Further Clarifications of the Role of Cognitive Evidence

Since the main line of reasoning of this paper revolves around the role of cognitive evidence in indicating phenomenal consciousness, further clarifying this role will make this line of reasoning itself clearer, and, I believe, more persuasive. What is the evidential linkage that the argument for the phenomenal/cognitive indeterminacy requires (and for which I have argued)? This argument requires that having good evidence for a subject's cognitively employing a content in one of certain ways is *a necessary condition* for being justified in taking him to be phenomenally conscious of this content. But, furthermore, the argument requires that in any situation in which we are justified in judging that the subject employs the relevant content in one of the relevant ways (that is, that she has a cognitive attitude of the sort necessary for inferring that a phenomenally conscious state occurs), we would not be justified in preferring the assumption that the relevant phenomenal state does not occur over its negation. (We may express this idea by saying that the required cognitive attitudes must not be *strongly defeasible* as evidence for phenomenal consciousness.) The argument requires this, since if there were situations in which we are justified both in accepting the occurrence of the cognitive attitude and in denying the occurrence of the phenomenally conscious state, it would be possible to isolate the neural correlate of the cognitive attitude, and consequently to isolate the neural correlate of the phenomenally conscious state (by "subtraction").

Thus, the argument requires that the identification of the occurrence of one of the relevant cognitive attitudes constitute a reason of a rather strong kind in favor of the assumption that the subject is phenomenally conscious of the content in question. Still, the argument allows that there be situations in which, in spite of the occurrence of one such cognitive attitude, the obtaining of counter (cognitive) evidence (presumably, evidence for the absence of another relevant cognitive attitude) would *weakly defeat* the evidential power of that (positively testifying) cognitive attitude, in the sense that we will not be justified in preferring the assumption that the relevant phenomenal state occurs over its negation. The argument allows that the evidential power of the required cognitive attitudes be weakly defeasible in this sense, since as long as the occurrence of a (positively testifying) cognitive attitude guarantees that counter evidence cannot put us in a position in which we would be justified in believing that the relevant content is not phenomenally conscious, it would be impossible to isolate the neural correlate of the (positively testifying) cognitive attitude.[21]

I mentioned examples of two kinds of cognitive attitudes suitable for fulfilling these requirements: reflective beliefs—which constitute what Block calls "reflective consciousness"—and attitudes of the sort which constitute what he calls "access consciousness". It is very plausible that in order to be justified in viewing a content as phenomenally conscious we must be justified in ascribing to the subject cognitive attitudes of one of these two kinds. For what, besides evidence for a subject's believing herself to be conscious of a content and her freely using it in various cognitive processes, can count as evidence that she is really phenomenally conscious of this content (and not

[21] Examples of negatively testifying cognitive evidence are a negative reflective belief, and the ("free") use of a content that clashes with the content of the alleged phenomenal state.

only as evidence that her cognitive system received and processed the relevant information without its reaching the level of phenomenal awareness)?[22] It is not only that the occurrence of either reflective consciousness or access consciousness constitutes a *prima facie* reason for attributing phenomenal consciousness; rather, each of these forms of cognitive consciousness also satisfies the requirement of not being strongly defeasible. As to access consciousness, if a subject is access conscious of a certain content, the strongest possible evidence against her being phenomenally conscious of that content is not simply her lacking the appropriate reflective belief (it is hard to accept that the absence of such a belief even weakly defeats the evidential power of full access consciousness), but her believing that she is not phenomenally conscious of that content (i.e., that she does not see the stimulus). But even such a belief on her part cannot strongly defeat the evidential power of her exhibiting full access conscious of the relevant content. As the discussion of super-duper-blindsight suggests, we simply cannot know in such strange cases whether the subject has the relevant phenomenal state or not. Since the evidential power of reflective consciousness is certainly not weaker than that of access consciousness, reflective consciousness too cannot be strongly defeated. (And, of course, if *neither* is strongly defeasible, then *both* are weakly defeasible; for when one of them is negatively testifying, they may clash.) Thus, the epistemic requirements that the line of reasoning presented in this paper sets appear to be satisfied.

The issue of the relations between different kinds of cognitive evidence for (or against) the obtaining of phenomenally conscious states might raise an objection to the argument I have advanced. Since cognitive attitudes of different kinds can serve as evidence for the occurrence of type-identical phenomenally conscious states, it might be argued that we can compare two such cases and see what neural features the processes which lead to these cognitive attitudes share. These features, the objection proceeds, constitute the neural correlate of the *common* phenomenal feature.

Perhaps this is the way for discovering the neural correlate of phenomenal consciousness Block has in mind. Replying to the question, "[I]f our evidence always concerns experiential contents that are actually accessed, how can the Phenomenal and Access NCC ever be empirically distinguished?" (Block unpublished, p. 10), he says: "The answer is that it is not true that our evidence always concerns experiential contents that are accessed. There are a variety of paradigms in which we can use convergent evidence involving varying degrees of access to try to separate the Phenomenal from Access NCC." (*ibid.*) One such paradigm, according to Block, is signal detection theory (SDT), which examines subjects' reactions to stimuli at around threshold level, and model their behavior in terms of two factors: the extent to which the subject has the relevant visual experiences and the subject's (implicit) criterion for reporting seeing the stimuli. The subjects' reports have been shown to be influenced not just by their experiences, but also by various other factors like

[22] In fact, the argument advanced here does not depend on the assumption that cognitive attitudes of no other kind may serve as evidence for phenomenal consciousness (and as I said (see note 14 above), nothing hinges on whether the cognitive patterns I speak of exactly match Block's characterizations). The suggestion that cognitive attitudes of these two kinds are the required ones might seem more plausible when we bear in mind that "access consciousness" covers several sub-kinds of cognitive attitudes.

expectations and motivation (which are controllable to a certain extent in experiments). Further, there is evidence that what subjects *believe* about their experiences is influenced by various factors other than their experiences themselves (see Snodgrass 2002). Such findings indeed suggest that phenomenal consciousness can be separated from some kinds of access consciousness. As Block says, "SDT gives us reason to think that experiential content—based on the Phenomenal NCC—can be instantiated without the kind of access that is based in the Access NCC." (*ibid.*, p. 11)

This is still not enough for ensuring the separation sought after, since it remains the case that we cannot have evidence for phenomenal consciousness that is not evidence for *some kind* of cognitive consciousness. Indeed, Block goes as far as to say that SDT suggests that experiential contents need not be accompanied by "the kind of access required for report, planning, decision-making evaluation of alternatives, memory and voluntary direction of attention" (*ibid.*) I do not think that Block takes himself to have shown that the exclusion *of all these kinds* of access is warranted by SDI, but this is not essential to our concern, since he does not argue that we can have evidence for phenomenal consciousness that is not evidence for *any kind* of cognitive consciousness—since, that is, he does not reject the very notion of epistemic linkage between phenomenal consciousness and cognitive consciousness (of some kind). In speaking of kinds of access and of using convergent evidence involving varying degrees of access Block plausibly intends to suggest how the difficulty posed by the obtaining of this linkage for finding the neural correlate of phenomenal consciousness can be overcome. The only way in which the notion of evidence involving different degrees (or kinds) of access—in contrast with the notion of evidence involving no access (in the wide sense)—may be thought to be employed in the search for the neural correlate of phenomenal consciousness is this: we can compare cases involving the same phenomenal content but different degrees (or kinds) of access; the neural features that are shared by the processes which lead to the different accessing cognitive attitudes are supposed to constitute the neural correlate of the common phenomenal content. This is, in fact, the objection I raised to the main argument of this paper two paragraphs ago.

This objection, however, cannot work. It presupposes that apart from sharing the neural correlates of the relevant phenomenally conscious states, the processes in questions have nothing (non-trivial) in common at the neural level. Otherwise, of course, we would not be able to isolate the neural correlate of the phenomenally conscious states. But this presupposition seems unrealistic, for there are strong constraints on the cognitive attitudes that can serve as evidence for the same phenomenal state and thus also on the processes leading to them. These processes must all be such that they make use of the same representational content (the representational content of the phenomenal state), so it is hard to see how eliciting the relevant cognitive attitudes from the phenomenal state would not involve neural stages that have much in common.[23] Block claims that various empirical findings regarding visual detection are better explained by assuming that what underlies the

[23] I write "neural stages" in the plural, since a process which involves a digitalization of analogue-perceptual information would most plausibly not be a simple one-stage process. (This fact makes the indeterminacy problem even more severe.)

subjects' behavior that indicates their having various perceptual phenomenal states is their having environmental beliefs—beliefs about what is in front of them—rather than introspective beliefs (which, as noted, are influenced by various factors other than the phenomenal states themselves). Since I am willing to grant this claim, we need not get into the details. What is important is that beliefs about environmental features that are represented by phenomenal states—beliefs which as far as their evidential role vis a vis phenomenal states is concerned are a kind of cognitive access—share elements of their representational contents with introspective beliefs about the phenomenal states. These shared elements must be reflected at the neural level. This illustrates that the presupposition that different kinds of cognitive consciousness do not share elements at the neural level is untenable. So this strategy too would fail to isolate the neural correlates of the phenomenally conscious states.

It is important to emphasize that there is nothing anti-scientific in my argument for the untraceability of phenomenal non-cognitive consciousness. Specifically, I certainly do not purport to undermine any scientific hypothesis regarding mental–neural correlations. My point, rather, concerns the way such hypotheses should be interpreted: they should be interpreted as hypotheses concerning correlations between neural properties on the one hand and cognitive properties, or cognitive properties plus phenomenal non-cognitive properties, on the other hand. They should *not* be interpreted as hypotheses concerning correlations between neural properties on the one hand and phenomenal *non-cognitive* properties on the other.

The Psychological Untraceability of Phenomenal Consciousness

Some philosophers who defend the notion of non-cognitive phenomenal consciousness insist that this phenomenon can be subject to scientific investigation (see, e.g., Block 1994b; Chalmers 1996, 1998). According to my view, it is impossible to have it both ways: taking phenomenal consciousness to be non-cognitive implies that this phenomenon evades scientific investigation. I hope to have shown that non-cognitive phenomenal consciousness evades neuroscientific investigation; I shall now briefly present an argument to the effect that non-cognitive phenomenal consciousness lies beyond the reach of psychology. If this is indeed the case, we may say that it lies beyond the reach of science in general. The argument, as presented here, certainly deserves a wider discussion, but I believe that the following brief presentation nevertheless delivers the main idea.

What underlies the phenomenal–cognitive indeterminacy I have argued for is the fact that if phenomenal non-cognitive conscious states exist, they operate through cognitive mediators, and these cognitive mediators are also epistemic mediators, the ones through which we learn about these phenomenal states. It is *only* through these cognitive mediators that we can come to learn about the obtaining of phenomenal conscious states.[24] If so, then studying the regularities of the (mediating) cognitive

[24] This is true, at any rate, with respect to the phenomenal states of others. But this is the important thing as far as scientific findings regarding consciousness are concerned.

attitudes can suffice for the purposes of psychology, and there is no scientific value in postulating non-cognitive phenomenally conscious states. To see this, consider, first, paradigmatic or normal cases of phenomenal consciousness. In paradigmatic or normal cases, attributing (non-cognitive) phenomenally conscious states to a subject may play some role in predicting and explaining the subject's behavior, but this role is dispensable, for the same predictions and explanations can be arrived at by studying the connections between the relevant cognitive attitudes—those which serve as evidence for the occurrence of phenomenally conscious states, whatever other cognitive attitudes, stimuli and behavior.[25] This claim by no means presupposes that (non-cognitive) phenomenal properties are epiphenomenal;[26] the point is rather that since their instantiations are inferred from the occurrences of certain cognitive attitudes to which they give rise—so that their instantiations are epistemically determined by these occurrences—then, in the paradigmatic cases, we can infer the relevant behavior directly from the occurrences of those attitudes (and whatever other relevant factors), without appealing to those phenomenal properties.

In non-paradigmatic or pathological cases, however, which are the truly interesting cases, the attribution of phenomenal properties to a subject may play no role in predicting or explaining the subject's behavior. If we are interested in explaining or predicting the behavior of a subject with a certain deficiency in employing perceptual information, we have no choice but to study her unique perceptual–cognitive patterns, i.e., what cognitive attitudes she is likely to have under circumstances of various kinds (and then study the connections between these attitudes and behavior, which may or may not be normal). The introduction of phenomenally conscious states cannot be of any help in predicting or explaining the cognitive attitudes the subject is likely to have in such abnormal cases, since the phenomenally conscious states are not accompanied in these cases by their (standard) cognitive accompaniments. In general (that is, in both normal and pathological cases), then, as far as psychological predictions and explanations are concerned, the ascription of phenomenally conscious states yields nothing over and above what the ascription of the relevant cognitive attitudes yields. So in an important sense, phenomenal properties lie beyond the reach of psychology. If one wishes to defend the assumption that non-cognitive phenomenal properties exit, it

[25] For convenience, I speak of prediction and explanation of behavior, but the points argued for here equally apply to the prediction and explanation of various mental states that ensue from the "mediating" cognitive attitudes (and indirectly from the phenomenal states).

[26] In his debate with Shoemaker about absent qualia, Block (1980) argues that the possibility of functionally identical states only one of which is qualitative does not imply that qualia are epiphenomenal, since the causal efficacy of a qualitative character is compatible with the possibility that a state with a qualitative character and a state which lacks a qualitative character play the same causal role. The epistemic linkage I highlight is compatible with this view and it does not depend on qualia being epiphenomenal. For the very possibility that Block mentions prevents us from taking evidence for the obtaining of a state with the causal profile characteristic of a certain qualitative state to indicate that it is the qualitative state, rather than causally identical non-qualitative state, that obtains.

seems that one cannot rely on considerations which concern the predictive or explanatory value of the postulation of such properties.[27]

Since neither psychology nor neuroscience can provide reasons for postulating non-cognitive phenomenal states, it seems reasonable to say that science in general can provide no such reasons. I do not proceed to argue here that there can be no reason for postulating such states (or that the idea of their existence makes no sense)—I do not pursue what may be viewed as the ontological implication of the epistemological and methodological claims I have advanced. I think, though, that if the arguments presented here are on the right track, we are at least entitled to say that the burden of proof on those who wish to defend the existence of such elusive properties is anything but trivial.

[27] A related issue which I can only briefly mention is the following. One might argue for the existence of non-cognitive phenomenal properties on the grounds that postulating them best explains our *reflective beliefs* concerning our phenomenal states. Indeed, it seems plausible to claim that introspective beliefs are ultimately causally anchored in, and hence explainable by the postulation of, some "intrinsic" features of brain events which correlate with phenomenally conscious events. But this claim falls short of vindicating the existence of non-cognitive phenomenal properties, since it is significantly weaker than the claim that those intrinsic features are identical with phenomenal properties. (Further, one can deny the latter claim while still accepting that phenomenal properties are causally relevant for the existence of those reflective beliefs. For example, it is arguable that if phenomenal properties are identified with (segments of) the functional profiles of those brain events, they are causally relevant for the existence of reflective beliefs.) The same rationale would undermine the objection that postulating non-cognitive phenomenal properties best explains various other cognitive processes (other than those which involve reflective beliefs). A mistaken line of reasoning similar to the one which moves from the first (and weaker) of those claims to the second is exemplified by Flanagan's argument to the effect that propositional attitudes have characteristic qualitative feels. According to Flanagan, the fact that there is almost never any doubt that a person who takes herself to be in a state of belief is indeed in a state of belief rather than in a state of desire, is best explained by the assumption "that there really is a certain way it feels to be believing something and that way of feeling is different from the way it feels to be desiring something" (Flanagan 1992, p. 68). But it isn't clear why this explanation is the best one. What advantage does it have over the explanation according to which it is simply part of the causal role of beliefs and desires to produce characteristic second-order beliefs? What explanatory insight is gained by identifying the internal features in virtue of which beliefs and desires cause characteristic second-order beliefs with qualitative feels?

References

Baars, B. (1988). *A cognitive theory of consciousness*. Cambridge: Cambridge University Press.

Block, N. (1980). Are absent qualia impossible? *Philosophical Review*, *89*, 257–274.

Block, N. (1993). Review of Dennett: Consciousness explained. *The Journal of Philosophy*, *90*, 181–193.

Block, N. (1994a). Consciousness. In S. Guttenplan (Ed.) *A companion to the philosophy of mind*. Oxford: Blackwell.

Block, N. (1994b). Qualia. In S. Guttenplan (Ed.) *A companion to the philosophy of mind*. Oxford: Blackwell.

Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, *18*, 227–287.

Block, N. (1996). Mental paint and mental latex. In E. Villanueva (Ed.) *Philosophical Issues*, *7*: *Perception*. Atascadero: Ridgeview.

Block, N. (1998). How not to find the neural correlate of consciousness. In S. A. Hameroff, A. W. Kaszniak, & A. C. Scott (Eds.) *Toward a science of consciousness II (The Second Tuscon Discussions and Debates)*. Cambridge: MIT Press.

Block, N. (2003). Mental paint. In M. Hahn, & B. Ramberg (Eds.) *Reflections and replies*: *Essays on the philosophy of Tyler Burge*. Cambridge: MIT Press.

Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Science*, *9*, 46–52.

Bogen, J. E. (1998). Locating the subjective pump: The thalamic intralaminar nuclei. In S. A. Hameroff, A. W. Kaszniak, & A. C. Scott (Eds.) *Toward a science of consciousness II (The Second Tuscon Discussions and Debates)*. Cambridge: MIT Press.

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

Chalmers, D. (1998). On the search for the neural correlate of consciousness. In S. A. Hameroff, A. W. Kaszniak, & A. C. Scott (Eds.) *Toward a science of consciousness II* (*The Second Tuscon Discussions and Debates*). Cambridge: MIT.

Chalmers, D. (2000). What is a neural correlate of consciousness? In T. Metzinger (Ed.) *Neural correlates of consciousness*: *Conceptual and empirical questions*. Cambridge: MIT.

Chalmers, D. (2004). The representational character of experience. In B. Leiter (Ed.) *The future for philosophy.* Oxford: Oxford University Press.

Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in the Neuroscience*, *2*, 263–275.

Crick, F., & Koch, C. (1995). Are we aware of neural activity in primary visual cortex? *Nature*, *375*, 121–123.

Davies, M., & Humphreys, G. W. (1993). Introduction. In M. Davies, & G. W. Humphreys (Eds.) *Consciousness*. Oxford: Blackwell.

Dehaene, S., & Changeux, J.-P. (2004). Neural mechanisms for access to consciousness. In M. Gazzaniga (Ed.) *The cognitive neurosciences III*. Cambridge: MIT.

Dennett, D. (1991). *Consciousness explained*. Boston: Little, Brown.

Dennett, D., & Kinsbourne, M. (1992). The where and when of consciousness in the brain. *Behavioral and Brain Sciences*, *15*, 183–247.

Edelman, G. M. (1989). *The remembered present*: *Biological theory of consciousness*. New York: Basic Books.

Fehrer, E., & Raab, D. (1962). Reaction time to stimuli masked by metacontrast.*Journal of Experimental Psychology*, *63*, 143–147.

Flanagan, O. (1992). *Consciousness reconsidered*. Cambridge: MIT.

Jackendoff, R. (1987). *Consciousness and the computational mind*. Cambridge: MIT.

Jacobson-Horowitz, H. (2003). *What matters in phenomenal consciousness: a conative-evaluative account*. Doctoral Dissertation Harvard University.

Llinas, R. R., Ribary, U., Joliot, M., & Wang, X.-J. (1994). Content and context in temporal thalamocortical binding. In G. Buzsaci, R. R. Llinas, & W. Singer (Eds.) *Temporal coding in the brain*. Berlin: Springer.

Logothetis, N., & Schall, J. (1989). Neuronal correlates of subjective visual perception. *Science*, *245*, 761–763.

Noe, A., & Thompson, E. (2004). Are there neural correlates of consciousness. *Journal of Consciousness Studies*, *11*, 3–28.

Newman, J. (1997). Putting the puzzle together: Towards a general theory of the neural correlate of consciousness. *Journal of Consciousness Studies*, *4*, 47–66.
Penefield, D. (1975). *The mystery of the mind*: *A critical study of consciousness and the human brain*. Princeton: Princeton University Press.
Poeppel, D. (1996). A critical review of PET studies of phonological processing. *Brain and Language*, *54*, 317–430.
Sartori, G., & Umilta, C. A. (2000). How to avoid the fallacies of cognitive subtraction in brain imaging. *Brain and Language*, *74*, 191–212.
Searle, J. (1992). *The rediscovery of the mind*. Cambridge: MIT.
Snodgrass, M. (2002). Disambiguating conscious and unconscious influences: Do exclusion paradigms demonstrate unconscious perception? *American Journal of Psychology*, *115*, 545–580.
Van Gulick, R. (1989). What difference does consciousness make? *Philosophical Topics*, *17*, 211–230.
Weiskrants, L. (1986). *Blindsight*: *A case study and implications*. Oxford: Oxford University Press.