Consciousness and Criterion: On Block’s Case for Unconscious Seeing

IAN PHILLIPS

St. Anne’s College, University of Oxford

Block (2012) highlights two experimental studies of neglect patients which, he contends, provide ‘dramatic evidence’ for unconscious seeing. In Block’s hands this is the highly non-trivial thesis that seeing of the same fundamental kind as ordinary conscious seeing can occur outside of phenomenal consciousness. Block’s case for it provides an excellent opportunity to consider a large body of research on clinical syndromes widely held to evidence unconscious perception. I begin by considering in detail the two studies of neglect to which Block appeals. I show why their interpretation as evidence of unconscious seeing faces a series of local difficulties. I then explain how, even bracketing these issues, a long-standing but overlooked problem concerning our criterion for consciousness problematizes the appeal to both studies. I explain why this problem is especially pressing for Block given his view that phenomenal consciousness overflows access consciousness. I further show that it is epidemic—not only affecting all report-based studies of unconscious seeing in neglect, but also analogous studies of the condition most often alleged to show unconscious seeing, namely blindsight.

1. Overview

From certain theoretical perspectives concerning consciousness and sight, there is no question that seeing occurs outside of consciousness. Take Rosenthal’s actualist higher-order thought theory of consciousness on which a first-order perceptual episode is conscious just if it is the object of a higher-order thought (Rosenthal 1986, 2005). Since on this theory the relationship between perceptual episode and higher-order thought is treated as merely causal, the theory predicts the occurrence of perceptual episodes in the absence of higher-order representation and so consciousness. Or consider the view—not uncommon amongst vision scientists—that seeing occurs whenever there is neural transduction of retinal input. Few would doubt that such processes can occur outside of consciousness (Stoerig 1997: §2). Hence few would doubt that seeing can. Neither of these perspectives is mandatory, however. Many theorists deny that consciousness requires
actual or even potential higher-order representation. And still more would balk at the suggestion that seeing of the same fundamental mental kind as ordinary conscious seeing occurs whenever there is neural transduction of retinal stimulation.

Block is just such a theorist. On the one hand, Block is perhaps the leading exponent of the view that an episode can be phenomenally conscious, not just in the absence of higher-order representation (Block 2011a), but despite not being access conscious, that is without its contents being available for use in reasoning, or for the rational control of speech and action (Block 1995: 231, 2005, 2011b).1 On the other hand, Block adopt a substantial and independently motivated conception of perception due to Burge. According to Burge, perception is a psychological kind constituted by ‘objective sensory representation by the individual’ (2010: 368, emphasis in original). Objective representation, for Burge, is representation of the distal physical environment as contrasted with proximal aspects of sensory stimulation. Such objective representation, Burge argues, requires exercise of the perceptual constancies: ‘capacities to represent environmental attributes, or environmental particulars, as the same, despite radically different proximal stimulations’ (2010: 114).2 This is a much more demanding characterisation of perception than mere neural transduction of sensory input. From Block’s perspective then we face a substantive and open question: is all seeing, i.e. objective visual representation by the individual, conscious?

Block answers negatively: seeing of the same fundamental psychological kind as conscious seeing can and does occur unconsciously.3 In support he cites ‘dramatic evidence’ from ‘cases [of neglect] in which a single perceptual state involves integration of both conscious and unconscious elements’ (2012: 11–2). This evidence is superficially compelling and representative of a large body of research on clinical syndromes widely held by philosophers and scientists to evidence unconscious perception. Block’s appeal to it provides an ideal opportunity to consider such work critically. My critique proceeds as follows. §2 examines the two studies of neglect patients to

---

1 For more on phenomenal and access consciousness, see §5 below.

2 Constancies do not suffice for perception by the individual on Burge’s picture. Objective representations are sometimes only attributable to sub-individual systems. Burge claims that we have genuine perception by the individual when such representations are appropriately connected to whole individual function, paradigmatically action-guidance (2010: 373–6). Though crucial in relation to the assessment of other alleged cases of unconscious perception (see Phillips MS), I do not rely on the gap between objective representation and individual-level perception in what follows.

3 Burge concurs (2010: 374–6) citing the dissociation of perception and consciousness (a) in certain arthropods, (b) in clinical patients with blindsight, unilateral neglect/extinction, and prosopagnosia, and (c) in ordinary subjects under conditions of masking or inattention. I consider Burge’s case as part of a wider examination of the consensus in favour of unconscious perception in Phillips (MS).
which Block refers, and explains why their interpretation as evidence of unconscious seeing faces several difficulties. §§3–4 argue that, even bracketing these local difficulties, an old but overlooked problem concerning what constitutes an adequate criterion of consciousness besets both studies. §5 explains why the problem is especially pressing for Block given his view that phenomenal consciousness overflows access consciousness. §§6–7 argue that the problem is epidemic, affecting all report-based studies of unconscious perception in neglect (§6), as well as analogous studies of the condition most often alleged to show unconscious perception, namely blindsight (§7). I end by considering a number of natural objections, thereby situating my case in the wider literature concerning the scientific study of consciousness (§8).

2. Illusions in Neglect: Dramatic Evidence for Unconscious Perception?

Unilateral neglect can be characterized as ‘the failure to report, respond, or orient to novel or meaningful stimuli presented to the side opposite a brain lesion, when this failure cannot be attributed to either [elementary] sensory or motor defects’ such as would be caused by lesions to primary visual or motor areas (Heilman et al. 1993: 279; for reviews see Buxbaum 2006, Bartolomeo 2007 and Husain 2008). Neglect is a complex and fractionated disorder but, in the prototypical form encountered in the present studies, it involves right parietal damage, and corresponding neglect of left egocentric space. I assume that neglect is both an attentional and a perceptual deficit: facts about perception being at least partly dependent on attentional processes.

The two studies to which Block appeals present neglect patients with stimuli which ordinarily induce illusions due to features on left and right-sides. Both studies aim to show that these illusions persist in neglect despite an absence of conscious awareness of the features on the subjects’ left-sides. Following Mattingley et al. (1995), Ro and Rafal (1996) investigate the Müller-Lyer and Judd illusions (Fig. 1). As Figs. 1(a) and (b) show, the fins of the Judd illusion shift the apparent centre of the horizontal shaft away from the true centre. This does not happen in the symmetric Müller-Lyer illusion where the fins instead affect apparent line length. By analyzing the separate effects of these illusions, Ro and Rafal argue that the bisection pattern of their subject, SD, shows the influence of ‘fins’ on her neglected left. For instance, whilst Figs. 1(a) and (c) have the same fin configuration on the right, SD exhibits a rightward shift in bisection in relation to Fig. 1(a) relative to 1(c) consistent with the standard Judd illusion in healthy subjects. Ro and Rafal further argue that SD is not conscious of the left fins and so must perceive them unconsciously. For Block this shows ‘integrated conscious and unconscious perception’ of right and left fins.
Such integration underpins his claim that this is perception of the same fundamental kind in both cases.

Vuilleumier and Landis (1998) exploit a rectangular Kanizsa figure (Fig. 2). To ordinary subjects this seemingly depicts four (amodally...
completed) black circles overlain by an illusory (modally completed) rectangular surface. Vuilleumier and Landis show that their neglect patients bisect the stimulus as they would a closed rectangular figure, and not as they would separated parallel lines. They conclude that their patients perceive the rectangular surface, and so the black ‘pac-men’ inducers on both sides. However, they go on to argue that their subjects are not conscious of the left inducers and so must perceive them unconsciously. As before, Block infers that ‘conscious and unconscious perception can be integrated into a single percept’ (13), evidencing their sameness of kind.

These studies appear elegantly to demonstrate unconscious seeing. However, on closer examination, numerous concerns emerge. First, although both studies show that left features are illusion inducing, it does not follow that they are perceived. For instance, the pattern of bisection behaviour in Vuilleumier and Landis’s study suggests that the subject’s visual system registers the inner edges of the inducers, interprets them as corners, and completes these into a closed rectangular figure. But, strictly speaking, such extraction of form from retinal stimulation does not suffice for perception in Burge’s sense. For Burge (2010: 351ff.), the registration of a luminance contour is a matter of pre-perceptual sensory registration. Moreover, Burge argues that there can be form completion in relation to such registration without genuine perception (ibid: 418). Perception requires more than merely going beyond sensory registration; it requires fully objective representation. Consequently, the bisection behaviour of patients may be explicable in terms of non-perceptual sensory registration of left-side features. Similar concerns can be raised in relation to Ro and Rafal’s study.

Second, Block appeals to the idea that left and right inducers are ‘integrated into a single percept’ to convince us that perception of the same fundamental kind occurs both consciously and unconsciously. However, it is obscure what such integration involves and whether it suffices to establish sameness of fundamental kind. People with strong imaginations can ‘integrate’ imagined elements within their visual experience. For example, Galton (1880: 322) describes ‘a power which is rare naturally, but can . . . be acquired without much difficulty, of projecting a mental picture upon a piece of paper, and of holding it fast there, so that it can be outlined with a pencil’. Few would suppose that this directly establishes that imagination and perception are of the same fundamental kind. Likewise, the naïve realist who denies that perception is of the same fundamental kind as hallucination is not confounded by partial (and so ‘integrated’) hallucinations (Martin 2004: 80–1). In the present case, moreover, we have yet to be given a reason not to think of the ‘integration’ as between pre-perceptual registration of features on left and right, registration which continues onto a perceptual stage on the right but not the left-side.
Let us bracket these concerns and assume that the illusory effects do demonstrate perception of the left inducers. Do the studies establish that such perception is unconscious? In both, the absence of consciousness is supported by data from a same/different (s/d) paradigm. Ro and Rafal sequentially present pairs of lines as in Fig. 1 and ask: ‘Did the first figure look the same as the second figure?’ Vuilleumier and Landis simultaneously present pairs of Kanizsa stimuli as in Fig. 3 and ask: ‘Is the top pair the same as the bottom pair?’ Both studies find close to no errors when the stimuli are the same on both sides, or differ only on the right, but many errors when the stimuli differ on the left. Does this show that subjects are not conscious of the left features?

Figure 3. Example of stimulus used in Vuilleumier and Landis’s s/d task. Reproduced with permission from Vuilleumier and Landis (1998: 2482), © 1998, Lippincott-Raven Publishers.

4 Note that it is not claimed by Vuilleumier and Landis, nor by Block, that the perception of the rectangular surface is unconscious. This is a further issue for which no direct evidence is offered. The crucial issue concerns the perception of the left inducers.
One major issue here is that the (two stimulus) s/d task places substantially greater demands on sustained attention and working memory than the (single stimulus) bisection task. Neglect patients have well-known problems with sustained attention and working memory, and are highly sensitive to changes in attentional demands. This raises two possibilities. First, that patients are conscious of the left-side inducers in the bisection task (where we have evidence of illusory effects) but not in the s/d task due to the higher task demands. Second, that patients are conscious of the inducers in both tasks but perform poorly in the s/d task—not because of differences in consciousness but because the comparison places too great a demand on working memory. Evidence of task demand effects on reported awareness in Ro and Rafal’s paradigm has in fact been found (Olk et al. 2001).

A further issue specifically troubles Vuilleumier and Landis’s methodology. Modal completion requires that our visual system detects the inducers’ inner edges. Yet the s/d task does not track differences in the inducers’ edges. The edges match in top and bottom figures in Fig. 3, producing modal completion in both cases. The s/d task instead tracks differences in how the surfaces are filled-in: either solid black, or with black and white lines. Thus, it is consistent with failing to detect such differences in filling-in (e.g. due to degraded contrast perception on the unattended left) that subjects nonetheless do detect the edges consciously.

These various concerns problematize Block’s appeal to these particular paradigms as direct evidence of unconscious perception. They also illustrate the difficulties of experimental work in this area. However, even setting these concerns aside, a more fundamental and general problem casts doubt on the experimental methodology in these and all structurally similar studies. To appreciate the issue we require some basic concepts from signal detection theory. I review these in the next section before putting them to work in §4.

3. Signal Detection Theory and Response Criteria

According to signal detection theory (SDT) (Tanner and Swets 1954, Green and Swets 1966), a subject’s responses in any given task are determined by two parameters: first, the subject’s underlying perceptual sensitivity to the relevant target stimulus; second, their response criterion or threshold. Consider a simple ‘yes/no’ (y/n) task by way of example. In such a task, the subject is asked to say whether or not a target stimulus was present on a given trial (or, alternatively, if they saw a stimulus). Here SDT models the situation in terms of two sensory distributions—one associated with noise (in system and environment), the other with a signal (i.e. stimulus presence), together with omnipresent noise. Making the (non-trivial) assumption that both are equal-variance Gaussian (normal) distributions, the distance between distributions in units of standard deviation is given by $d^\prime$ (‘d-
prime’). The parameter $d'$ characterizes the subject’s sensitivity to stimulus presence. However, this parameter is insufficient to determine whether or not a subject will respond ‘yes’ in any given trial. To determine how a subject will respond to sensory stimulation, we also need to know their response criterion ($c$)—effectively the threshold which stimulation must exceed to elicit a positive response (see Fig. 4).

Notoriously, a subject’s response criterion is not fixed. As Green and Swets write, an ‘observer can be induced to change his [response] criterion in any of several ways’ (1966: 87). The most obvious ways of manipulating a subject’s criterion are to vary the prior probability of a stimulus being presented, or by changing the relative costs and benefits of the four possible response pairings (saying ‘yes’ when there is/isn’t a stimulus, and saying ‘no’ when there is/isn’t a stimulus). But it is commonly recognized that response criteria are subject to the influence of many subtle and poorly understood factors including (but not limited to) experimental instructions, task design, motivation, fatigue and preconceptions about the experiment’s purpose or intended outcome (so-called ‘demand characteristics’).

The fact that responses are the joint product of sensitivity ($d'$) and criterion ($c$) means that it is impossible to determine a subject’s sensitivity from the percentage of correct answers they give in a y/n task (e.g. Azzopardi and Cowey 1998). For example, in a case of near-threshold perception, noise and signal distributions are close to one another. Accordingly, a subject who wishes to avoid false alarms (saying ‘yes’ when no stimulus is present) will adopt a significantly conservative response criterion ($c \gg 0$, where 0 is midway between signal and noise distributions). As a result, they will respond ‘no’ even when it is much more likely than not that their sensory response is due to a stimulus. At the extreme, a subject may adopt a criterion so conservative that they effectively say ‘no’ on every trial. If so, their percentage correct will utterly fail to reveal their underlying ability to

![Figure 4. SDT analysis of a simple y/n task showing a moderately conservative response criterion. The subject will indicate that a stimulus was present whenever sensory stimulation exceeds their response criterion. Otherwise they will indicate that no stimulus was present.](image-url)
detect a stimulus ($d'$). To determine this, a psychophysicist must either manipulate the subject’s response criterion, plotting a so-called receiver operating characteristic (ROC) curve whose shape suffices to calculate $d'$, or use a ‘bias-free’, multi-alternative forced-choice (mafc) task.\(^5\) A 2afc task involves presenting a subject with two intervals on each trial, and ‘forcing’ the subject to select which contains the target. The intervals might be temporal or spatial. One interval will always contain the target; the other might simply be a blank field, or it might contain a distractor. Standard examples include asking a subject to say which of two temporal intervals a Gabor patch is presented in, or presenting two patches next to each other and asking the subject to say whether the left or right patch is higher in contrast. Such tasks are effectively ‘bias-free’ as subjects naturally adopt a symmetrical criterion, simply choosing whichever interval corresponds to the strongest sensory stimulation. In consequence, 2afc tasks are used directly to estimate a subject’s perceptual sensitivity (Green and Swets 1966: 107–8, Macmillan and Creelman 2005: 179).

In itself, SDT is silent about consciousness. However, we can think of it as clarifying two thresholds: an objective threshold above which stimuli are discriminable by a subject ($d' > 0$); and a subjective threshold above which a subject will produce a positive response in a given task (e.g. indicate that a stimulus was presented).\(^6\) One possibility is that we should associate consciousness with the subjective threshold. If we do, then there is a strong case to be made for the existence of unconscious perception since it is easy to find subjects who consistently deny that anything has been presented or seen despite $d' \gg 0$.\(^7\) However, there are good reasons to reject any such strict association. To take an extreme example (cf. Merikle 1984: 450) if we pay a subject $1 for every time they correctly say that a stimulus was present but penalize them $100 for every time they incorrectly say a stimulus was present, that subject will naturally adopt a highly conservative criterion and so regularly deny that a stimulus is present even in a situation where they have significant perceptual sensitivity. Few will be tempted by the

\(^5\) Though usage varies, I here follow the traditional psychophysical practice which does not count y/n tasks or s/d tasks as forced-choice tasks precisely to restrict the term ‘forced-choice’ to naturally unbiased tasks.

\(^6\) This terminology originates in Cheesman and Merikle (1986) and Merikle and Cheesman (1986). Note that the distinction I am marking is not between environmental reports (e.g. that a stimulus was present) and introspective reports (e.g. that a stimulus was seen). Both environmental and introspective reports are affected by bias since they have the same non-symmetrical form. Thus both face the same fundamental concern pressed in the text. Though the environmental/introspective distinction may be important in some contexts, experimental and folk practice tends to move freely between such reports, and both are typically regarded as evidence of consciousness (cf. Shea 2012: 310, fn. 2).

\(^7\) Such cases do not strictly entail the existence of unconscious perception since it remains to be shown that $d' > 0$ is a sufficient condition for perception.
thought that such monetary incentives directly affect consciousness as opposed simply to affecting verbal responding.

More generally the recognition that responding can be biased, not just by monetary incentives, but by myriad hard-to-control factors means that ‘most investigators . . . reject any approach for distinguishing conscious from unconscious perceptual processes that is based solely on subjective reports’ (Reingold and Merikle 1990: 17–8). Some (e.g. Eriksen 1960 and Holender 1986) instead associate consciousness with the objective threshold i.e. $d^\prime > 0$. Others deny that there is any simple association between detection theoretic thresholds and consciousness, typically granting only that $d^\prime = 0$ is a sufficient condition for absence of consciousness. I return to these issues in §5. For now, the crucial point is this: where a subject performs well on an unbiased (e.g. 2afc) task but poorly (as judged by percent correct) on a biased task, this pattern of responding does not entail unconscious perception unless we make the very strong assumption that consciousness is found only above the subjective threshold. If we reject that assumption, we need to consider seriously the possibility that the response pattern reflects conscious perception unreported due to a conservative response criterion.

With this in mind, I now return to the two studies of neglect to which Block appeals. I argue that both are naturally interpreted as cases in which the subjects perform poorly (as judged by percent correct) on a biased s/d task due to conservative response bias and so despite having residual perceptual sensitivity for left-side features. Consistent with—though not required by—this interpretation is the hypothesis that the subjects have degraded conscious perception of features on the left.

4. Reinterpreting Block’s Cases

Although Ro and Rafal somewhat misleadingly describe theirs as a forced-choice task (1996: 974), both studies to which Block appeals use s/d tasks to establish lack of consciousness of the left inducers. The analysis of s/d tasks is complex (Macmillan and Creelman 2005, Petrov 2009, DeCarlo 2013). But for present purposes a simple point suffices, namely that ‘participants seem to naturally adopt strong response biases in same-different experiments. In particular, a preference for “same” is commonly observed for hard-to-discriminate stimuli’ (Macmillan and Creelman 2005: 218, Petrov 2009: 1014). In other words, s/d tasks exhibit strong biases towards ‘same’ responding in ordinary subjects near threshold.\(^8\)

\(^8\) In line with this, in their *Psychophysics: A Practical Introduction*, Kingdoms and Prins write: ‘Because the two discriminands (Same and Different) are not symmetric, this task is particularly prone to the effects of bias . . . Thus, it is advisable to analyze the data to take into account any bias’ (2010: 46).
Neglect patients have perceptual deficits on their left, meaning that stimuli presented there are plausibly near threshold. Accordingly, we should anticipate strong biases towards ‘same’ responding, and be extremely cautious in inferring from ‘poor’ performance as judged by a subject’s percentage correct to a lack of objective sensitivity. Moreover, there is additional theoretical reason why neglect patients may have conservative left-side criteria. If such subjects fail to update their pre-lesion (or current right-side) criterion in relation to their post-lesion, left-side field, then insofar as their lesion leads to a reduction in perceptual sensitivity, their previously unbiased criterion will become conservative.

Table 1 presents the data from Vuilleumier and Landis’ three subjects. They score perfectly when reporting whether stimuli are the same on both sides. Setting aside right-side differences with respect to which subjects are at ceiling, we have a striking contrast: subjects score 100% when the left-sides are the same, yet do poorly when they are different. This alone suggests subjects are biased toward ‘same’ responding. Table 2 presents Ro

<table>
<thead>
<tr>
<th>Patient</th>
<th>Same on both sides</th>
<th>Right-sided differences</th>
<th>Left-sided differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient 1</td>
<td>8/8</td>
<td>8/8</td>
<td>2/8</td>
</tr>
<tr>
<td>Patient 2</td>
<td>8/8</td>
<td>8/8</td>
<td>1/8</td>
</tr>
<tr>
<td>Patient 3</td>
<td>8/8</td>
<td>8/8</td>
<td>1/8</td>
</tr>
</tbody>
</table>

Table 2:

<table>
<thead>
<tr>
<th>Response</th>
<th>Left</th>
<th>Difference on Right</th>
<th>Neither</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same</td>
<td>12</td>
<td>1</td>
<td>23</td>
</tr>
<tr>
<td>Different</td>
<td>0</td>
<td>11</td>
<td>1</td>
</tr>
</tbody>
</table>

To test this hypothesis we might investigate the performance of ordinary subjects using degraded left inducers in both s/d tasks.

and Rafal’s similar results. Their subject, SD, scores close to perfectly (96%) when reporting whether stimuli are the same on both sides but answers ‘same’ in every case in which they differ on the left (and indeed in one case on the right). Again, this is extremely suggestive of a bias toward ‘same’ responding.

The limited data available preclude reliable estimates of $d'$ and $c$ for these subjects. Nonetheless, a qualitative assessment is clearly consistent with the hypothesis that subjects are strongly biased towards ‘same’ responding and thus that their s/d performances mask underlying perceptual sensitivity. Consequently, the interpretation of the s/d task data as indicating a lack of consciousness ignores the possibility that the subjects may in fact be conscious but performing ‘poorly’ on the s/d task due to conservative response bias. This possibility is not only plausible in the present instance but, as the next section explains, one which Block himself explicitly urges us to take seriously in similar cases.

5. Phenomenal Overflow and the Subjective Threshold

At this juncture, one way to press the case for unconscious perception would be to insist that a verbal or otherwise explicit detection response was the *sine qua non* of consciousness (cf. Weiskrantz 1997, Naccache 2006).\(^{11}\) In other words, we might insist on a subjective threshold for consciousness. However, for reasons noted in §3, such theorists are very much in the minority. A weaker, and much more popular, line is that consciousness requires only *accessibility* to verbal report (or to explicit response systems). This line is consistent with an episode being conscious, but unreported, since an episode can be accessible without actually being accessed. Neglect patients typically can be induced to report features on their left-sides with appropriate attentional direction. We might also reason that if their response criterion had been more liberal, then they would have reported the left-side features/differences.\(^ {12}\) As a result, a theorist committed to an accessibility constraint on consciousness need not feel precluded from embracing the hypothesis above, viz. that the patients in

---

11 In fact, even Weiskrantz does not insist on an actual report. His striking view is rather that ‘phenomenal awareness itself . . . results from the delivery or potential delivery of a report’ (1997: 76, my emphasis).

12 Cf. Snodgrass and Shevrin: ‘phenomenal consciousness can be ascertained . . . by utilizing objective threshold methods, which are intrinsically sensitivity-based. . . . participants can report phenomenal states that they are not (i.e. would not ordinarily be) reflectively aware of—namely, by performing forced-choice discrimination tasks, which compel participants to respond to stimuli that they would otherwise ignore (i.e. to include below-criterion stimuli in reflective consciousness)’ (2006: 75). The context is a discussion of subliminal perception in ordinary subjects—the authors go on to mention blindsight as a potential counter-example to this picture. But see below §7.
the studies to which Block appeals enjoy degraded conscious awareness of the left inducers.

Block himself famously rejects either access or accessibility as a constraint on consciousness. More specifically, he argues that phenomenal consciousness overflows (i.e. contains more than) access consciousness. For Block, a state is phenomenally conscious just if there is something it is like to be in that state from its subject’s point of view (Nagel 1974), and, roughly speaking, access conscious just if its content is available for reasoning or the rational control of speech and action.\(^{13}\) Block’s claim that there can be unconscious seeing is the claim that seeing can occur without phenomenal consciousness. Given his view that phenomenal consciousness overflows access consciousness, a subjective threshold is plainly inadequate: any case in which phenomenal consciousness is present without access consciousness will be a case in which the subjective threshold is not met.

These are points Block himself explicitly commits to. First, consider the following passage in which Block discusses extinction, a close cousin of neglect:

If we take seriously the idea of phenomenality without access, there is a theoretical option that should be on the table, one that I think is worth investigating—that ventral stream activation [which Block earlier says is intact in extinction and neglect (198f.)] is visual phenomenality and the search for X [the supposed ‘missing ingredient . . . which, added to ventral activation (of sufficient strength), constitutes conscious experience’ (199)] is the search for the neural basis of what makes visual phenomenality accessible. The idea would be that the claims of extinction patients not to see extinguished stimuli are in a sense wrong—they really do have phenomenal experience of these stimuli without knowing it. (2001: 203)

Second, consider Block’s discussion of the exclusion paradigm (Jacoby 1991, Debner and Jacoby 1994), in which he explicitly claims that SDT ‘gives us reason to think that experiential content . . . can be instantiated without . . . access’ (2005: 49).\(^{14}\) Block says this because he thinks that we should interpret the relevant shifts of response criteria as reflecting a difference in access, ‘rather than a difference in consciousness’ (ibid.). Yet, if Block is happy to associate a subjective threshold with access but not phenomenality here, we are owed an explanation of why the same interpretation is inappropriate for the studies of neglect discussed above. Similarly, if Block is willing to attribute consciousness to extinction patients in the face

---

\(^{13}\) The classic case for a distinction and potential dissociation between access and phenomenal consciousness is made in Block (1995). Block’s more recent work argues for the closely related conclusion that ‘the machinery of phenomenology contains more than the machinery of cognitive accessibility’ (2007: 487). For further empirical defence of overflow see also Block (2005, 2011b, and 2014). In addition, see Burge (2010) and Lamme (2003). For critical discussion see Phillips (2011 and forthcoming).

\(^{14}\) I say a little more about the exclusion paradigm in §8 below.
of their denials that they see the relevant stimuli, we are owed an explanation of why such a hypothesis should not be taken seriously in the neglect cases above. In sum, Block’s case for unconscious perception faces a very serious internal challenge. By Block’s own lights an interpretation of the studies which he cites which appeals to degraded conscious perception together with biased responding must be ‘on the table’.

6. The Bigger Picture: Neglect

Thus far I have considered just two (albeit striking) studies highlighted by Block. I argued that neither provides clear evidence of unconscious seeing since in both the measure of consciousness is problematic. Such concerns are far from local, however. As I now argue, they affect all report-based studies of unconscious perception in neglect. In §7, I argue that such concerns also arise in analogous studies of blindsight.

The single most cited study of unconscious perception in neglect is a single-patient report by Marshall and Halligan (1988). Their patient, PS, who had severe, prototypical neglect, was presented with pairs of cards showing line drawings of houses (Fig. 5). When asked, ‘Are the houses the same or different?’, she replied that both were the same. Asked, ‘Which house would you prefer to live in?’, she replied that this was a ‘silly question’ (766). However, forced to choose, she picked the non-burning house on nine out of eleven trials. In contrast, when cards were presented with flames on the right, PS ‘immediately noticed the flames’ (766) and preferred the non-burning house on every trial.

Marshall and Halligan concluded that PS saw the flames unconsciously, an interpretation which is now standard lore (Husain 2008). Subsequently, much controversy has focused on whether PS really perceived flames as opposed to a lower-level, pre-attentive feature such as asymmetry (Farah 1994). However, from the current perspective a more fundamental objection is that the study (and subsequent studies such as Bisiach and Rusconi 1990, Vallar et al. 1994, and Doricchi and Galati 2000) contrasts a s/d task (‘Are the houses the same or different?’)—which as we have seen is subject to strong near-threshold biases—with what is effectively an unbiased 2afc task (‘Which house would you prefer to live in?’, i.e. which spatial interval contains the preferable house?). Consequently, any inference that a patient lacks conscious awareness of left-side features based on her responding in the s/d task ignores the possibility that she has degraded conscious percep-

---

15 Doricchi and Galati (2000) attempt to show that higher-level features are perceived in such cases. However, as they themselves recognize, perception of asymmetry together with right-side categorical perception suffices to explain their results. This debate raises the issue of whether we in fact need to appeal to objective representation to explain the preference data. I set this issue aside here.
tion but that such perception falls below her response threshold which is naturally conservative, i.e. biased towards ‘same’ responding.\textsuperscript{16}

Another important body of work on unconscious perception in neglect comes from priming paradigms in which neglected primes facilitate subsequent responding. Details aside, a crucial question is always: can we be sure that the prime was not consciously perceived? The gold standard in work with ordinary subjects involves showing that $d' = 0$. Unfortunately, no neglect study comes close to meeting this standard. Typically biased measures are used to determine awareness without any attempt to calculate underlying sensitivity (e.g. Berti and Rizzolatti 1992, Ládavas et al. 1993, Della Sala et al. 2010) and in cases where $d'$ can be crudely estimated from the reported data it is well above zero. For example, in McGlinchey-Berroth et al. (1993), it pos-

\textsuperscript{16} Marshall and Halligan also asked a y/n question, viz. ‘anything wrong’ with either card? PS replied negatively. But, as discussed, this too is a biased task. As a result, her responses can equally be explained in terms of conservative response bias.
sible that $d'$ is as high as 1. In consequence, I know of no study of unconscious priming in neglect which convincingly rules out the possibility that the prime is consciously perceived. Moreover, given the additional difficulties of providing sufficient statistical power to demonstrate that $d' = 0$ as opposed to just low (e.g. Rouder et al. 2007) and heightened concerns about task artefacts arising from lack of motivation due to problems with sustained attention in neglect (cf. Pratte and Rouder 2009), it is unlikely a convincing demonstration will be forthcoming soon.

Neglect is of course far from the only alleged case appealed to in defence of unconscious perception. Any study which measures awareness using a biased s/d or y/n measure and contrasts this with an unbiased forced-choice measure or behavioural response, must consider the possibility that the alleged discovery of unconscious perception simply reflects conservative responding in the biased task. This possibility has long been a complaint raised in relation to work on subliminal perception. In the next section, I explain why precisely this issue arises in relation to many studies of blindsight.

7. The Bigger Picture: Blindsight

Blindsight is perhaps the clinical condition most widely appealed to in defence of unconscious seeing. Blindsight is best characterized as ‘residual visual processing after destruction of primary visual cortex’ (Cowey 2010: 3). This definition raises two questions. First, does this residual function constitute perception? Second, insofar as it does, is consciousness missing? I explore the first question elsewhere (Phillips MS). Here I simply note that it is far from obvious that residual processing in blindsight does constitute perception. In the present context, I focus instead on the issue of

---

17 McGlinchey-Berroth et al. used a two-alternative delayed match to sample task. One of the four patients should arguably have been excluded. Depending on whether this patient is included and on task strategy assumed, $d'$ can be estimated as between 0.7 and 1.0. Estimates calculated using Palamedes (Prins, N. and Kingdom, F. www.palamedestoolbox.org) and MATLAB (MathWorks, Natick, MA).

18 The locus classici here are Eriksen (1960) and Holender (1986). For further discussion see Reingold and Merikle (1990). More recent work on subliminal perception seeks to find evidence of perceptual effects at the objective threshold (i.e. $d' = 0$). I discuss whether such work evidences unconscious seeing in Block and Burge’s sense in Phillips (MS).


20 If we adopt Burge’s conception of perception, at least two separate concerns arise. First, do visual sensory responses in blindsight implicate constancy mechanisms and so involve objective representation? Pace Burge (2010: 189, 374) there is no clear evidence that blindsight does involve perceptual constancies. Second, does successful forced-choice responding in blindsight reflect a direct visual response? A long-mooted alternative is that such responding reflects indirect awareness of autonomic reflex responses to visual stimulation, including but not limited to pupillary responses, blink reflexes and optokinetic reflexes. For references and discussion see Phillips (MS).
consciousness. For it is widely assumed that the residual visual capacities in blindsight need not be conscious.21 This claim is of course crucial to the interpretation of blindsight as evidence of unconscious seeing.

Given the discussion above, it should immediately give us pause to realise that the claim that blindsight involves unconscious perception is largely based on a dissociation between responding in a biased task and performance in an unbiased forced-choice task. One salient biased task in blindsight is the perimetry test which is used to establish the subject’s field defect in the hemifield contralateral to their damaged striate cortex—that is, their ‘blindness’.22 This is essentially a y/n task in which the blindsighted subject is asked whether or not they see a light at various intensities and locations in their visual field. Negative responses indicate a field defect. The striking feature of blindsight is that such defects are consistent with success on forced-choice tasks. For example, the blindsighted subject may be able to indicate in which of two temporal or spatial intervals a stimulus is presented, or whether a given stimulus is an ‘X’ or an ‘O’. Famously, blindsighted subjects can achieve high-levels of performance in a narrow range of such tasks, indicating residual perceptual sensitivity or ‘sight’.

As already much emphasised, poor performance (as measured by percent correct) in a biased y/n task combined with above chance performance in a 2afc task is consistent with conscious perception unreported in the y/n task due to a conservative response threshold. If this is all that is happening in blindsight, blindsight might be no more striking than ‘the behaviour of normal subjects near the lower limit of conscious vision’ (Azzopardi and Cowey 1998: 292). This claim was first pressed in the early eighties by Campion et al. (1983) and Campion and Latto (1985) who argued that ‘blindsight reduces to no more than the effect of using different decision criteria with degraded vision’ (1983: 446).23 Campion et al. in fact push two thoughts. First, that subjective measures are inadequate, not least given issues of response bias. As they put it: ‘The view that proponents of blindsight appear to hold, that an appropriate behavioural metric [for phenomenal representation] can be found in the methodology of clinical perimetry or casual verbal report, seems to us so inadequate as not to warrant serious consideration’ (1985: 756). Second, Campion et al. adopt an objective threshold for measuring consciousness. They thus argue that because

21 Though he is more cautious than most, Block holds that it is ‘plausible’ that consciousness is missing in at least some cases of blindsight (1995: 232).

22 Another salient, and equally biased task, involves so-called ‘commentary key’ responses elicited following a forced-choice task, one key indicating that the subject was aware of some ‘visual aspect of the stimulus presentation’, the other indicating that they were not (Weiskrantz et al. 1995: 6122).

23 For further criticism of the ‘orthodox’ interpretation of blindsight see Zeki and ffytche (1998), Kroustallis (2005), and Overgaard et al. (2008).
forced-choice procedures reveal perceptual sensitivity \((d' > 0)\), this shows that ‘the stimuli are not unconscious’ and as a result that ‘all subjects in all studies are, by conventional criteria, aware of stimuli to some extent’ (1983: 480). This second move of course trivially blocks any interpretation of blindsight in terms of unconscious vision. However, whether or not we sympathize with the association of consciousness and the objective threshold, we should sympathize with Campion et al.’s initial concerns about the adequacy of subjective reports. As a result, a hypothesis needs considering: blindsight is in fact just degraded conscious vision unreported due to conservative response bias.

There are a number of related reasons to take this hypothesis seriously. First, as against a picturesque conception of blindsight as entirely preserved visual function with consciousness simply spliced away, blindsighted subjects not only have severely degraded visual function but do sometimes report conscious awareness of some kind (so-called type-II blindsight). Further this reported awareness has been shown to correlate well with perceptual sensitivity in certain conditions (Barbur et al. 1993, Stoerig et al. 2002). This pattern is, of course, what we would expect of normal, degraded vision. Reason to suspect that the apparently more striking dissociations of visual capacity and awareness may arise from variation in response criterion comes from findings that whether or not a blindsighted subject reports awareness is highly variable. This is evident from a study by Zeki and ffytche (1998) which found a broad correlation between GY’s reported awareness and detection rate, but that on some sessions GY’s reported awareness varied dramatically despite no change in stimulus conditions or detection rate (as measured by percent correct). Azzopardi and Cowey (1998) suggest that this variation may be matter of criterion fluctuation, perhaps due to boredom and fatigue. Kentridge (2015), offering a similar explanation in terms of criterion fluctuation, suggests that the variation may reflect how GY is interpreting his own condition as well as the experimenter’s expectations.

Whatever the explanation of Zeki and ffytche’s data, one thing that is known is that reported awareness can be affected by altering the experimental question. For example, Stoerig and Barth (2001) investigated GY’s perimetry results under fixed stimulus conditions, simply changing the question from ‘Did you see anything?’ to ‘Were you aware of anything?’. As can be seen by comparing the number of white circles (indicating positive

---

24 A number of controversies surround type-II blindsight—see the papers in Foley and Kentridge (2015), especially Brogaard (2015), Foley (2015), Kentridge (2015), and Balsdon and Azzopardi (2015) (which pertinently discusses how type-II blindsight might arise from response bias).

25 GY is probably the most extensively studied blindsight subject, and a much clearer case than Weiskrantz’s original patient DB.
responses) in Figs. 6(a) and (b), this simple change has dramatic results. A natural interpretation is that GY adopted a more liberal criterion with respect to the ‘awareness’ question.

Similarly, Overgaard et al. (2008) studied a blindsighted woman, GR, using a graded, four-point scale—the Perceptual Awareness Scale of Ramsøy and Overgaard (2004)—and found that, whereas she reported no awareness using a traditional y/n measure (and so met the traditional diagnostic criteria for type-I blindness), using this finer grained scale, ‘her blindsight seemingly “disappeared” in the sense that . . . [a]ll correctness above chance seemed related to vague yet conscious vision’ (Overgaard 2011: 477). One possible interpretation here is that GR enjoyed dim conscious awareness of stimuli in line with her perceptual sensitivity; exhibited strong conservative bias in the y/n task; but was more liberal in her responding in respect of the lower ratings of the Perceptual Awareness Scale. Of course, an alternative possibility here (and in relation to GY’s responses in Stoerig and Barth 2001) is that her criterion was systematically biased in a liberal direction such that her reports exaggerated the extent of her conscious awareness. Current evidence does not allow us to choose between these interpretations.

The hypothesis that blindsight results exclusively from conservative responding in y/n tasks was subject to a rigorous investigation by Azzopardi and Cowey (1997, 1998). Their work may seem to tell against that hypothesis—and so with it the idea that blindsight is degraded con-

Figure 6. Static perimetry results for GY. Figure 6(a) shows results for the instruction ‘press when you see something’; 6(b) shows results for the instruction ‘press when you are aware of something’. White circles indicate stimuli with respect to which a button press was made. Black circles (which form radii extending along the tested meridians) indicate that no button press was made. Reproduced with permission from Stoerig and Barth (2001: 575), © 2001 Elsevier Science (USA).
scious vision. However, as I now discuss, such evidence shows only that blindsight is not simply degraded normal vision.

Azzopardi and Cowey first establish a proof of concept, demonstrating that GY can be extremely biased in his y/n responding ($c = 1.87$ on average) even when he has good perceptual sensitivity to stimuli. In contrast, he is effectively unbiased in 2afc responding ($c = 0.01$ on average). As Azzopardi and Cowey comment, such shifts ‘could easily produce significant dissociations of performance when measured with percent correct’ (1998: 298). Azzopardi and Cowey then test GY’s precise discriminative sensitivities with respect to two basic classes of stimuli: (a) vertically moving bars or fields of random dots and (b) large ($20^\circ \times 20^\circ$) static, vertical square-wave gratings of high luminance and contrast (see 1998: 301 for details). In respect of either type of moving stimulus, GY’s ‘sensitivity did not differ significantly between yn and 2afc detection’ (302). This they conclude, ‘conforms with the argument put forward by Campion et al. (1983) that blindsight might be nothing more than [a] patient’s use of consistently different response criteria during clinical [y/n] and forced-choice testing’ (302). However, they found that ‘GY was significantly more sensitive to the static gratings in the 2afc task than in the yn task’ (299). As a result, the dissociation between the two types of responding with respect to static stimuli cannot simply be explained in terms of shifts of response criteria from unbiased in 2afc responding to highly conservative in y/n responding.

Does this refute the hypothesis that blindsight is conscious vision unreported due to response bias? It does not. First, that hypothesis is shown to be quite consistent with the data concerning moving stimuli. Second, the explanation, and so upshot, of the finding of differential sensitivity with respect to static stimuli is obscure. One conclusion we can draw is that blindsight, in respect of static stimuli, is not simply degraded normal vision. For no dissociation of y/n and 2afc responding is found with respect to normal vision with comparable stimuli at threshold. However, our question is not whether blindsight is normal vision, but whether it should be regarded as conscious vision—even if a form of vision radically unlike that found in

---

26 Weiskrantz (e.g. 2009a: 57–8) relies on Azzopardi and Cowey’s work to counter Campion et al.’s critique (and implicitly the interest of signal detection approaches). However, he rightly relies on their work only to make the limited claim that blindsight is not simply degraded normal vision. As I explain in the text, it cannot be appealed to in order to show that blindsight is blind (nor, surely, to dispute the interest of signal detection approaches!).

27 GY may not always be so biased in y/n responding. Azzopardi and Cowey encourage him towards conservative bias by the inclusion of an additional ‘yes or guess’ task. This encourages him not to guess in the y/n task.

ordinary subjects (cf. Overgaard 2011). A dissociation between \( y/n \) and \( 2afc \) responding does not settle this issue. Moreover, the dissociation may be explicable in ways which do not require a dramatic departure from the basic detection theoretic account outlined above. For instance, one idea proposed by Azzopardi and Cowey (2001; see also Ko and Lau 2012: 1405–6) is that the differential sensitivity between \( y/n \) and \( 2afc \) responding may result from the blindsighted subject’s criterion continually being updated (‘jittered’), perhaps because of standing problems adjusting to their post-lesion signal distribution. This can be shown to yield an increase in the effective variance of the trial-averaged signal and noise distributions, and in turn a reduction in sensitivity specifically affecting detection (\( y/n \)) as opposed to discrimination (\( 2afc \)) tasks. This is clearly an important area for further empirical investigation. However, accepting the interpretation offered by Azzopardi and Cowey, it is not at all obvious what inference we should draw concerning consciousness. Instead an added layer of complexity is added to the already vexed concern of relating detection theoretic thresholds to consciousness. In particular, one wants to know whether the factors (e.g. criterion jitter, lack of visual memory standards) which make \( y/n \) sensitivity poorer than \( 2afc \) sensitivity are factors which affect consciousness or not. Given these outstanding issues, it hard to see how blindsight can be relied upon as a clear case of unconscious vision.

No doubt some theorists will be inclined to respond by stamping their feet down and insisting that their subjects should be trusted when they deny that they are conscious. Thus, Weiskrantz bluntly insists: ‘I trust my subjects’ reports’ (2009b: 415; cf. Cowey 2004: 580, 586–8). Weiskrantz may ultimately be right to do so. But notice two points. First, as we saw above, Block urges us to take seriously the possibility that ‘the claims of extinction patients not to see extinguished stimuli are in a sense wrong’ and that ‘they really do have phenomenal experience of these stimuli without knowing it’ (2001: 203). Thus, Block cannot simply insist that we trust blindsight subjects’ reports. Second, and less ad hominem, Weiskrantz’s reliance on his subjects’ reports can be argued simply to ignore the issues about bias raised above, and in particular the long-standing plaint of many psychophysicists

---

29 One might doubt (as Kentridge 2015 does) whether blindsight should really be counted as visual even when it is associated with consciousness as in type-II blindsight. However, this is no help to someone who wishes to claim that blindsight is a case of unconscious vision since it presses the question of what constitutes vision proper, and plainly if blindsight does not constitute vision, it cannot be a case of unconscious vision. For reasons to doubt that blindsight does constitute visual perception proper by Burge’s lights see Phillips (MS). For an argument that type-II blindsight should be thought of as a form of abnormal but nonetheless genuinely visual awareness see Foley (2015).

30 Azzopardi and Cowey (1997: 14194) earlier suggested an explanation in terms of the role of visual memory standards specifically in relation to \( y/n \) responding.
that ‘the fundamental problem with using self-reports as a measure of awareness is that it transfers the responsibility for defining awareness from the investigator to the observer’ (Merikle and Reingold 1992: 59, Merikle 1983). This plaint is especially pressing in the case of blindsight where the nature of awareness is highly degraded and abnormal (compared to the subjects’ own good fields, as well as to that of ordinary subjects). Furthermore, as discussed above, the reporting of awareness in blindsight is highly variable and potentially sensitive to a range of factors: motivation, fatigue, experimental instructions, a subject’s interpretation of their own condition, and demand characteristics. For all these reasons we should agree with Kentridge (2015): ‘it is dangerous to rely on introspection’.

To close this section, I offer one further illustration of this issue. Persaud and Lau (2008), following Weiskrantz, but sharply departing from Merikle and Reingold, pursue ‘the idea that experiences should be related by the experimenter, not the experimenter’ (1046) in relation to GY’s experience (or lack thereof). However, they share the concern that GY subject might define or interpret ‘awareness’ (and cognates) in an idiosyncratic manner. To address this they seek to ensure that GY adopts the philosophically-relevant understanding by getting GY to read ‘definitions of qualia given by The Oxford Companion to the Mind [Gregory (1987)], The Stanford Encyclopaedia of Philosophy [Tye (2003)], Jackson (1982), and Dennett (1991)’ (1047) and then asking him whether he enjoys qualia in the defined sense. Since he (mostly) denies it, they conclude that he does not enjoy ‘qualia’ in his scotoma.

However, one only has to look at these ‘definitions’ to feel the force of Merikle’s worry about the transfer of responsibility, as well as more general concerns about response bias. For instance, The Oxford Companion to the Mind speaks of qualia in terms of ‘a distinctive phenomenological character which we have all experienced but which ... is very difficult to describe’ (1987: 666). But what does the word ‘distinctive’ here mean or add? Might it not be naturally taken to suggest a division between certain conscious experiences which have qualia and certain conscious experiences which don’t? This is further encouraged by the characteristic choice of examples of qualia in the relevant definitions. Thus, the reader is pointed to ‘the smell of freshly ground coffee or the taste of a pineapple’ (ibid.); to ‘the characteristic experience of tasting a lemon, smelling a rose, hearing a loud noise or seeing the sky’ (Jackson 1982: 127); or to ‘seeing green, hearing loud trumpets, tasting liquorice, smelling the sea air, handling a piece of fur’ (Tye 2003, viz. the ‘Qualia’ entry in The Stanford Encyclopedia of Philosophy). Such examples are naturally read as suggesting that experiences of loud noises are distinctive in possessing qualia. That hearing muffled sounds, or smelling freeze-dried coffee does not produce them; that whereas pineapples and lemon are qualia-productive, dry toast and water are not. To
the extent he has perception at all, GY has very severely degraded and abnormal perception in his blindfield. His experiences are the visual equivalent of the faintest of murmurs; he sees through a glass, darkly. On a natural reading of the ‘definitions’ of qualia which Lau and Persaud offered him (combined with his natural reading of their expectations), it is perhaps no surprise that GY says that he experiences qualia, ‘Only very rarely . . . on very easy trials, when the stimulus is very bright’ (Persaud and Lau 2008: 1048). It would be rash to conclude that GY lacks any form of conscious experience in other trials from this data.31

8. Objections and Conclusions

Block’s case for unconscious perception is not compelling, or so I have argued. Given his commitment to a distinction between phenomenal and access consciousness, the fundamental problem with it runs deep. It also runs wide, problematizing the appeal to all report-based studies of neglect and blindsight as evidence of unconscious perception. In this closing section, I consider various lines of response to the concerns raised thus far. The discussion is necessarily highly selective; its primary purpose is to locate my critique in the wider context of debates about the scientific study of consciousness.

A first natural reply is to insist that, even if subject to bias, subjective responses, or specifically verbal reports, surely remain a good guide to phenomenal consciousness. As a result, so the reply continues, their absence (in neglect or blindsight) does at least evidence the absence of consciousness, even if it does not outright demonstrate it. However, whilst it should be granted that subjective reports do correlate well with phenomenal consciousness in ordinary cases, perceptual sensitivity also correlates well with phenomenal consciousness in ordinary cases. The cases under current consideration are manifestly not ordinary—in particular they are cases where perceptual sensitivity and verbal reports appear to dissociate. Minimally, this means that in such cases we have evidence (from the presence of perceptual sensitivity) that consciousness is not missing. What is more, we must consider the hypothesis that consciousness tracks perceptual sensitivity which sometimes is, and sometimes is not, revealed in subjective reports. Given the availability of this hypothesis, it is doubtful that the absence of verbal reports in the presence of perceptual sensitivity provides discriminating evidence in favour of the absence of consciousness.

31 Compare a suggestion which Dennett makes in relation to a hypothetical ‘super blindsight’ subject: ‘if he complains of the absence of qualia, he might simply be noticing the relative paucity of information he now gets from his vision and misdescribing it’ (1991: 358–9).
It is natural to pursue this objection by suggesting that the neurological damage sustained in the relevant cases provides evidence that consciousness is indeed missing. Dennett suggests this in relation to blindsight:

The proof [that blindsight is unconscious] is more circumstantial: the subjects deny that they are conscious of any such events, and their verbal denials are supported by neurological evidence of brain damage on the one hand, and by the coherence of their denials on the other. So we believe them! (1991: 326; see also 327, fn. 3)

In response, note three things. First, to the extent that patients’ denials are coherent, this may be equally well explained by supposing that they are operating with a stable, conservative response criterion. Second, as Dennett himself adds in a note: ‘the details of neurological damage by themselves (without the denials) would prove nothing; it is only by matching up neurological damage with (credible) reports and behavioural evidence that we get any hypotheses about which parts of the brain are essential for which conscious phenomena’ (ibid: fn. 2). Yet evidence of neurological damage, even taken together with behavioural and report-based evidence, does not offer a direct route to claims about consciousness. For, as already much emphasised, an alternative understanding of the neurological damage in neglect and blindsight is that it has led to severely degraded or distorted—yet conscious—perception, together with a tendency towards conservative responding in biased tasks. Finally, and relatedly, Block himself appeals to neurological evidence, at least in the case of extinction, to undermine rather than support patient reports. Thus, contrasting ordinary subjects in a change detection experiment with an extinction patient, GK, Block writes: ‘there is no reason to distrust these [ordinary] subjects, whereas the subject GK . . . does have brain damage that prevents attention to the left side of space when there is a competing stimulus on the right, so there is a real question as to whether he might see something on the left that he cannot report’ (2008: 310–1).

No doubt if we knew the true neurological theory of consciousness, we could appeal to that theory to decide a particular case or condition. For example, Block (1995: 232) suggests that it is ‘independently plausible’ that phenomenal consciousness is missing in blindsight on the grounds that blindsight lesions disrupt the 40 Hz oscillations which Crick and Koch (1990) propose are the neural basis of perceptual consciousness. Proper engagement with Crick and Koch’s proposal (not to mention other competi-

---

32 Of course, a patient’s criterion may not be fully stable but then their reports will not be fully coherent across trials and conditions. As already noted, such an unstable pattern of data arguably tells in favour of a criterion-based account and against an account which postulates damage to certain essential neural circuits of consciousness.

33 It is unclear exactly what evidential status Block accords this theory.
tor theories of consciousness) is well beyond the present brief. In the specific case in point, suffice to note that Crick and Koch themselves now take a much more cautious approach to the significance of 40 Hz oscillations, deeming them neither necessary nor sufficient for consciousness (see, e.g. 2003: 123). But more generally, it is, I suggest, doubtful that we currently have any theory of consciousness which is sufficiently well-supported that it could be used to decide whether the cases currently in dispute involve consciousness or not. A further complication is that conditions such as blindsight and neglect often constitute a key part of the case in favour of a given theory of consciousness. Consequently, there is an additional concern with looking to any such theory for an independent assessment regarding the presence or absence of consciousness in these conditions. An obvious example is the recently popular theory that feedback to primary visual cortex is essential for visual perceptual consciousness—a theory to which Block (2005) is attracted—which was initially proposed partly on the basis of blindsight (Lamme 2001). Another salient example is the role of neglect and extinction in supporting ‘attentional’ theories of consciousness (e.g. Driver and Vuilleumier 2001 and Prinz 2011).

A second important line of reply to the problems raised above is to argue that psychophysicists have now developed ways of moving beyond ‘the problem of the criterion’ which lies at the heart of my critique. What is undoubtedly true is that psychophysicists have long recognized these issues and have made various attempts to overcome them, and so to distinguish between unconscious and conscious processes below the subjective threshold (e.g. Cheesman and Merikle 1986, and Merikle and Reingold 1992). However, no proposal has garnered consensus. In particular, the inevitable question arises as to why we should think that the feature alleged to distinguish conscious from unconscious processing is in fact a reliable, independent measure of phenomenal consciousness (Goldman 2000). An obvious example is the exclusion paradigm briefly alluded to already. In this task, a masked prime (e.g. SCALP) precedes a visible word stem (e.g. SCA—). Subjects are told to complete the stem with any word except the prime (e.g. SCARE or SCARF). Debner and Jacoby (1994) found that subjects often flouted the exclusion instruction (and at a rate higher than a baseline established using non-word primes) when primes were very briefly presented (e.g. 50ms) but not when they were presented for slightly longer durations (e.g. 500ms in Exp. 1 or 150ms in Exp. 2). A tempting inference is that is precisely the longer duration, and so consciously perceived, primes that sup-

34 Though, in fact, more recent studies of blindsight patients have done much to call it into question (e.g. ffytche and Zeki 2011).

35 In assuming that there is a problem to overcome such approaches reject objective threshold approaches such as Eriksen (1960) and Holender (1986).
port exclusion behaviour in contrast to their briefer and only *unconsciously* perceived counterparts (Debner and Jacoby 1994; Merikle and Joordens 1997). However, following Snodgrass (2002), Block (2005) argues that exclusion may not be a measure of phenomenal consciousness at all but rather a subjective threshold effect, or, as he puts it, a failure of a certain kind of access.\(^{36}\) The essential idea here is that the relevant qualitative difference, viz. exclusion behaviour, may be due to a response strategy only being adopted or available when stimuli reach a subject’s criterion—something which stimuli may fail to do even though they are phenomenally conscious. As discussed by Snodgrass (2002: 563ff.), a structurally similar concern applies to many other putative ‘converging’ qualitative differences between conscious and unconscious perception such as reverse priming effects (Merikle et al. 1995), false recognition effects (Jacoby and Whitehouse 1989), and divided attention effects (Merikle and Joordens 1997).

A more recently popular idea has been to focus on so-called type-2 (i.e. response-directed or meta-cognitive) measures instead of more traditional type-1 (i.e. stimulus-directed) measures. For example, a subject might be asked first to say whether a stimulus was present (a type-1 measure) and then to wager on their decision (a way of establishing how confident a subject is in their decision, a type-2 response; see Persaud et al. 2007). As with type-1 responses, type-2 responses can also be thought of as a joint product of sensitivity and criterion, hence it is crucial to establish properly bias-free measures of such meta-cognitive sensitivity (see Kunimoto et al. 2001, Galvin et al. 2003, Evans and Azzopardi 2007, and in particular Maniscalco and Lau 2012 on what they call meta-\(d^\)\)). However, whilst this is certainly an important line of investigation, it is again far from clear what exactly such measures are measures of. Whatever the answer, it is very doubtful that they provide a direct window onto consciousness (see e.g. Seth 2008 criticising Persaud et al. 2007). This is obvious if one is sceptical of higher-order constraints on phenomenal consciousness as Block of course is (Block 2011a). However, even theorists who are sympathetic to higher-order approaches urge caution in interpreting type-2 sensitivity as a direct measure of awareness. Thus, Maniscalco and Lau, partly on the (as we have seen, problematic) ground that the blindsight subject GY has above chance (type-2) sensitivity to his response accuracy in his blindfield (Persaud et al. 2007) argue for ‘a double dissociation between type 2 sensitivity and the contents of awareness’:

Reductions in type 2 sensitivity do not necessarily reflect reductions in phenomenalological stimulus awareness, and above-chance levels of type 2 sensitivity do not necessarily imply the presence of phenomenalological

\(^{36}\) For further critical discussion of the implications of exclusion effects see Irvine (2009).
stimulus awareness. Thus it does not seem tenable to use type 2 sensitivity as a hard and fast measure of awareness. (2012: 429)

These concerns with attempts to overcome the problem of the criterion are not intended to counsel despair, only that there is little reason to think that the problem has been overcome.37 My central claim has, in any case, been that it must not be forgotten. For in its light, we can see that the kinds of cases adduced by Block in order to establish that seeing of the same fundamental kind as ordinary conscious perception can occur outside of consciousness are not compelling. This, of course, does not show that all seeing is conscious, and that no case for unconscious seeing can or will be made. Only that more empirical and theoretical work is required before the traditional view that all seeing is conscious is taken off the table.38

References


37 Irvine (2013) contains an excellent and more comprehensive discussion of these and other issues concerning the scientific study of consciousness. Arguably she does counsel despair, arguing that the concept of consciousness should be abandoned as a scientific concept.

38 This paper was originally written for a workshop in honour of Ned Block organized by Frédérique de Vignemont in conjunction with his receipt of the Jean Nicod Prize. As I said then, I owe a huge debt to Ned. His work and example have been immensely influential on my own fledgling efforts to connect philosophical and empirical work on the mind. I am also extremely grateful for his personal support and encouragement, and his comments on that and other occasions. My understanding of both empirical and theoretical issues has been greatly aided by Paul Azzopardi who kindly allowed me to sit in on his advanced option seminars on the psychophysics and neuropsychology of conscious awareness. I have also learnt a great deal from conversations with Bob Kentridge, especially regarding blindsight. Sincere thanks to both. Thanks also for comments and discussion to Tim Bayne, Janet Bultitude, David Carmel, Nele Demeyere, Anil Gomes, Glyn Humphreys, Nick Jones, Rory Madden, Hanna Pickard, James Stazicker, Wayne Wu and Frédérique de Vignemont, and to an anonymous reviewer for this journal for an extremely helpful and constructive set of comments. This work was partially supported by a Leverhulme Trust Research Fellowship which I gratefully acknowledge.


— 2011a. ‘The higher order approach to consciousness is defunct’ Analysis 71(3): 419–431.


—— forthcoming. ‘No watershed for overflow: recent work on the richness of consciousness’ Philosophical Psychology.

—— MS. ‘Unconscious perception reconsidered’ Under review.


