

## Paradigm Lost: Review of Lawrence Weiskrantz, *Consciousness Lost and Found*\*

ANTHONY I. JACK

---

For thirty years, Lawrence Weiskrantz has been at the forefront of experimental research into neurological patients who have 'lost' awareness. This book provides a history and an overview of that research; which has focused on 'blind-sight' patients, who report no visual awareness in part of their visual field, and 'amnesic' patients, who have no experience of remembering past events. Yet, the book aims to be much more than a review. Using findings from his patients, and taking in a great deal of other research along the way, Weiskrantz addresses some fundamental issues. He calls these the 'What?', 'Whether?', 'Why?' and 'How?' of consciousness. What is consciousness? When, if at all, can we attribute it to animals? Why did it evolve? How does the brain accomplish it? The book is an empirical enquiry into the nature of consciousness, and it is written in a consistently engaging and accessible style.

Neuropsychological research has clearly had a major impact on thinking about consciousness. Largely because of this work, few now regard the topic as an issue beyond scientific enquiry (McGinn, 1989; Nagel, 1974). This research also paved the way for the recent proliferation of 'bold' scientific hypotheses, most of which presently offer far less insight into the phenomenon. Perhaps the best indicator of this is that philosophers with widely varying theoretical approaches to consciousness, including those that remain sceptical about current scientific approaches, choose neuropsychological research to frame and illustrate the arguments they wish to present (e.g. Block, 1995; Dennett, 1991).

This influence was not easily won. In order to achieve recognition of the significance of these disorders, Weiskrantz had to struggle against some deeply held, yet ultimately unjustified, convictions. It is worth listing four of the most theoretically important of these: (i) The belief that awareness is necessary for

---

\*Lawrence Weiskrantz, *Consciousness Lost and Found: A Neuropsychological Exploration*. OUP, 1997. Pp. vi + 294. ISBN: 0-19-852301-7.

**Address for correspondence** Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London WC1N 3AR, UK.

**Email:** a.jack@ucl.ac.uk

goal-directed action (discussed in chapters 1 and 4). (ii) The belief that subjective reports (first person commentaries) are intrinsically unreliable and non-scientific (chapter 3). (iii) The belief that awareness judgements reflect nothing more than the strength of neural activity induced by a stimulus—and therefore that perception without awareness is just a weakened form of perception with awareness (chapters 2 and 5). (iv) The resistance of experimental psychologists to the supposition that anatomically separable pathways carry out different functions, and thus their exaggerated scepticism of the logic of dissociation that is central to neuropsychological investigation (chapters 2 and 6).

Weiskrantz provides convincing evidence and argument that these convictions are misguided. For instance, his well-written and accessible discussions of brain anatomy and function, in chapter 6 and 8, present a weight of evidence supporting functional differentiation. This draws largely on experimental work carried out by others. Weiskrantz's own research has been instrumental in shifting opinion against the other three. For instance, he has shown that responses to stimuli in the 'blind' field can be as accurate as responses to 'seen' stimuli. Blindsight is clearly not weakened normal perception, nor is it equivalent to peripheral vision (see p. 60).

Weiskrantz does not lay out his theoretical points in this way. There is not one list in the book. Nor does he structure his writing around a rigid theoretical framework. His most important chapters are headed with the large and ambitious issues that he aims to tackle, e.g. the 'What', 'Whether', 'Why' and 'How' of consciousness. These form a set of overlapping narratives, each of which could be read alone. Within chapters, Weiskrantz keeps the narrative grounded by sticking close to empirical data. He moves deftly from one experiment to the next, describing and responding to it. Often, he completely avoids drawing out any explicit moral. He simply allows the example to do its work, and moves on. Along the way, there are numerous digressions and anecdotes—some of which provide valuable historical context. Yet, Weiskrantz does not get lost in these conversational narratives. Weiskrantz knows where he is going, he is just taking a scenic route.

This style makes for stimulating and informative reading. It also provides an unusually close insight into the processes of reasoning and weighing of evidence that guide an accomplished experimental scientist. Weiskrantz appears to be writing in the same voice that he would use to design his experiments. However, the approach does have its difficulties. The reader may struggle to keep a clear picture of Weiskrantz's theoretical position. Different chapters contain somewhat different perspectives on the same issues. Further, there are points at which Weiskrantz's responses are simply too quick—causing him to gloss over important issues. Two related examples can be seen to have deep theoretical repercussions.

The first concerns his attitude towards subjective reports (first person commentaries), such as reported awareness of a stimulus. Initially Weiskrantz wishes to establish that this form of evidence is no different from other forms

of behavioural evidence. He tells us, 'If the results are systematic and lawful, they have precisely the same status as in normal psychophysics' (p. 53). This is not quite right. For all empirical measures, we can make a distinction between *reliability* and *validity*. The replicable and lawful patterns of results obtained from blindsight subjects establish that subjective report measures can be reliable. Further, Weiskrantz is right to stress the point. Until recently experimental psychologists, still influenced by behaviourism, thought of these measures as essentially unreliable. Where defended, this claim has not been supported by empirical evidence, but rather by questionable theoretical argument. Validity, the claim that a measure accurately reflects an hypothesised internal state, variable or process, is typically much more difficult to establish. Yet, Weiskrantz overlooks the fact that establishing validity presents a special problem for subjective measures. The problem comes down to what the measure is taken as evidence for. Objective behavioural measures are usually interpreted as providing evidence about functional states or processes. Consequently, it is possible, at least in principle, to validate these measures using converging evidence from other objective measures—both behavioural and physiological. Subjective measures, in contrast, provide evidence about subjective states. From one perspective, the whole theoretical debate about consciousness centres on one question. What objective evidence should be taken as evidence for *awareness*? It isn't even clear that we *could* answer this question. Conceivably, subjective states could bear no law-like relation to the physical or functional states of an organism. There is a problem *in principle* about validating subjective measures.

Weiskrantz is blind to this point, and this gives rise to inconsistencies in his position. In chapter 3, Weiskrantz (p. 66) quotes Dennett (1991, p. 326) as remarking 'the subjects' responses provide evidence of a startling phenomenon *only* when they are interpreted as speech acts.' Weiskrantz wants to resist this point. He responds by claiming that verbal reports are awkward, that collecting them is a nuisance, and that manual responses can be substituted. This move is important for Weiskrantz, because he wants to defend the claim that we can measure awareness in animals using his 'commentary-key paradigm' (the 'Whether' of consciousness, chapter 4). However, Weiskrantz is missing the point, and the correct reading of Dennett has deep consequences for his commentary-key paradigm. Dennett's main concern is not that the responses should be verbal, but rather that they only provide evidence for awareness when they are interpreted in a particular way: i.e. when they are interpreted as *communicating* the subject's state of awareness. (Weiskrantz comes close to acknowledging this on p. 168). This is an important point to appreciate because it is clear that it is not always appropriate to interpret commentary-key responses as communications about awareness.

In an amusing aside, Weiskrantz notes that even verbal commentary responses are not always indicators of awareness, such as when the English discuss their feelings about the weather (p. 75). He also mentions in passing

that commentary-key responses must be voluntary (p. 238). Yet, he fails to pull these strands together. Animals would need to be trained, perhaps for weeks or months, to use the commentary-key paradigm (a point which Weiskrantz repeatedly stresses in chapter 4). Unlike humans, they can't follow verbal instructions specifying that they respond only when they are aware. Nor can they be questioned afterwards about their basis for response. Thus, it is effectively impossible to establish whether their responses are voluntary. Certainly, animals can be trained to give a presence/absence response, and this *may* be a reliable indicator of awareness (Cowey & Stoerig, 1995, reviewed on p. 87). Indeed, the similar patterns of performance shown by blindsight patients and animals with striate cortex lesions provides reasonably good evidence that presence/absence is a good measure of awareness for these animals. Yet, Weiskrantz's supposition that it should be permissible to *directly* interpret trained commentary-key responses as evidence for awareness is stretching the argument from analogy too far. Animal studies can tell us about brain organisation and function, but we need human studies to reliably relate these findings to *awareness*.

Weiskrantz claims that commentary responses are valid if they are voluntary. This sounds right, but it merely raises a second problem: How can we tell when a response is voluntary? On the back cover, Patricia Smith Churchland's review reads 'The revolutionary blindsight results knocked the stuffing out of the "obvious" assumption that awareness of a signal is necessary for an intentional response to that signal.' Presumably, Weiskrantz holds that discrimination responses are not 'intentional' or 'voluntary'. (Otherwise, his position would appear to be inconsistent.) Yet, Churchland's review clearly indicates that intuitions differ on this point.

Weiskrantz's failure to tackle this issue is his second major theoretical oversight. He twice dismisses the idea that consciousness can be inferred from goal-oriented behaviour (p. 8–9 and p. 78–82), concluding with the claim 'Complexity of behaviour, as a criterion, sometimes despairingly seems to depend on a blend of intuition and literary skill.' Yet, he is willing to accept Dickenson's behavioural criterion for intentional action, and to use it as evidence for awareness (e.g. sensitivity to changes in the motivational value of the outcome of an instrumental response, see p. 98–99). This is a mistake. Voluntary action can't be *defined* behaviourally, as Dickenson believes. This is obvious. It would be possible to program a computer to perform in the same way as Dickenson's rats using a vast variety of different algorithms. At least *some* of these won't look anything like the processes that are involved when humans perform intentional actions. We can have behavioural *markers* of intentional action in humans, such as sensitivity to dual task interference, but we can't have a behavioural *definition*. Just like 'awareness', the 'voluntary' or 'willed' status of an action is something we know about from the inside. 'Voluntary action' is, first and foremost, a subjective concept. Consequently, the problem of validity applies to measures of voluntary action, just as it does to measures

of awareness. There is no escaping from the fact that attributions of consciousness involve interpreting utterances and behaviours in a particular way. Further, there is no escape from the fact that intuitions will differ as to which interpretations are correct. To ignore this problem, or to advocate reliance on intuition, is simply not good enough. Science shouldn't be conducted using measures that we know, from the outset, are likely to suffer from systematic errors.

There appears to be only one solution to this problem of validity. This is to use the subject's own interpretations of their subjective mental states. By attempting to use 'objective measures' of awareness, or quasi-subjective measures such as Weiskrantz's 'commentary-key paradigm', we are only making it more likely that systematic errors creep in. These methods only move the decision point, the juncture at which the attribution is made, further away from the only valid source of data—the subject's actual experience. Of course, there are still problems. These largely concern whether we have understood the subject's self-attribution correctly. Language is the best tool for investigating this. The 'awkwardness' that blindsight subjects display when they are asked about their basis for response may be a nuisance, but it provides the strongest evidence that they lack awareness of the stimuli.

Weiskrantz's attempt to 'found' an empirical basis for investigating consciousness in animals (the 'Whether' of consciousness) is, therefore, not a great success. Similarly, his discussions of the 'Why' and 'How' of consciousness (its evolution and how the brain accomplishes it) are highly speculative and inconclusive. However, Weiskrantz's theoretical work is much more valuable when he focuses directly on understanding the neurological disorders he has studied. Here, Weiskrantz acknowledges that there is more to awareness than would be required to provide a simple 'commentary-key' response. In chapter 3, 'The "What?" of consciousness', he develops the idea that awareness arises when information reaches a 'commentary stage'. Weiskrantz writes (p. 227) 'the commentary stage is not assumed typically to be a simple "yes-aware" or "no-not aware" state. An affirmation of awareness also entails the accessing of a large storehouse of relevant, contextual knowledge that is implicit in the affirmation.' Weiskrantz argues that 'to be conscious is to be able to manipulate items in thought and imagery' (p. 171). He speculates that this may be achieved by the 'accessing of semantic knowledge' that occurs when information reaches the commentary stage. More specifically, the accessing of semantic knowledge allows the subject to make an explicit comparison. 'When we speak of memory in a sense that also conveys the phenomenal property of "*having* a memory", the argument is that it requires a specific comparison between the current input and the substrate' (p. 122).

Weiskrantz's analysis looks attractive for two reasons. First, it allows us to see that the ability to report on our own states of awareness ('meta-awareness') may be an extension of a more basic capacity for comparing perceptual stimuli with other, remembered, stimuli. Meta-awareness need only require us to compare our current *experience* with other, remembered, experiences. This is

attractive because it avoids the uncomfortable idea, prevalent in much philosophical writing, that 'reflexive' awareness is a fundamentally different sort of phenomenon from some more basic form of awareness, often dubbed 'phenomenal' awareness. If 'phenomenal' awareness isn't closely related to 'reflexive' awareness, then it is difficult to imagine what it might be.

Second, the idea that awareness arises from the integration of current inputs with stored semantic knowledge, fits comfortably with the notion that 'voluntary' action serves as an indicator of awareness. Integration of conscious information with stored knowledge would be necessary for an organism to respond in a 'willed' manner, i.e. in a manner congruent with the organisms beliefs and desires. Awareness may have evolved as a mechanism for the flexible use of perceptual information.

Weiskrantz bases his theorising about conscious processes directly on his work with patients: 'If you want to know what something might be good for, examine the situation where it is no longer present' (p. 6). At the end of the book, he suggests that further progress may be achieved through continued investigations of the neurological disorders in which subjects lack awareness. Weiskrantz wants to claim that he has 'founded' a method for studying consciousness. This is doubtful. The patients he considers have not specifically lost awareness. They have lost function in part of their brains. It is unlikely that loss of awareness will be the only consequence of this damage. This raises a difficult question: How can we differentiate between those deficits that are associated with their loss of awareness, and those that are not? Weiskrantz has never made much of the observation that blindsight subjects must be cued to respond to a stimulus in the blind field, although it has been discussed at more length by others (e.g. Dennett, 1991, p. 326). Perhaps this is because Weiskrantz thinks that the detection or 'noticing' of a stimulus does not require awareness. Alternatively, it might be one of the most basic processes that do require awareness. Theoretically, it is important to know. Unfortunately, studies of blindsight cannot give us the answer.

Weiskrantz's analysis of the processes that give rise to awareness is a promising start. Yet, at present this analysis remains as 'frankly philosophical' as the alternative positions he critiques. It is clear that a great deal more work would need to be done to understand his process of 'accessing semantic knowledge'. Progress will require the investigation of the processes that underlie voluntary action, tasks requiring the explicit comparison of stimuli, and tasks requiring the production of commentaries. Critically, this can only be achieved by looking at cases in which awareness is present, although it may prove useful to look at cases in which it is partial or disturbed. On p. 56, Weiskrantz explains the rationale behind his comparison between the capacities that neurological patients retain, and those that they have lost. 'This is an exercise in concept formation, and concepts might ease the way to more venturesome nettle grasping'. Weiskrantz has pioneered an important approach to consciousness. Allowing himself to be led by his observations of patients, he has managed to

avoid the theoretical preconceptions that distort the vision of most who work on consciousness. Yet, in truth, Weiskrantz's subject matter has only allowed him to make one secure contribution. He has validated the concept of unconscious processing. In doing so, he has drawn our attention to the issue of real interest and increased our confidence that consciousness is a subject for scientific investigation. He has made a major contribution. Yet, we are some way from finding the concepts that will allow us to grasp the nettle of consciousness.

*Institute of Cognitive Neuroscience,  
University College London*

## **References**

- Block, N. 1995: On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 18, 227–287.
- Dennett, D.C. 1991: *Consciousness Explained*. London: Penguin.
- McGinn, C. 1989: Can we solve the mind-body problem? *Mind*, 98, 349–66.
- Nagel, T. 1974: What is it like to be a bat? *Philosophical Review*, 83, 435–450.