Do We Have Evidence for Monkey Blindsight?

Nicholas Branigan

Blindsight is the presence of intelligent behavior in response to visual information, in a conscious creature, despite an absence of phenomenal vision. Following demonstrations of blindsight in human subjects, Cowey and Stoerig claim to show the existence of blindsight in monkeys. I argue that their case is modestly convincing. However, I also claim that their evidence fails to fully foreclose the alternate explanation that the monkeys are not entirely blind. I then propose modifications to their experimental design that could provide stronger evidence for blindsight.

Before I launch into the details of how Cowey and Stoerig's evidence supports the existence of blindsight, I address a confusion regarding different conceptions of consciousness. As Cohen and Dennett point out, some scientists believe that there can be phenomenal consciousness without access consciousness. I will not attempt to provide clearer definitions for these concepts than Cohen and Dennett. They define access consciousness as "conscious states that can be reported by virtue of high-level cognitive functions such as memory, attention and decision making." And phenomenal consciousness is "the subjective aspect of experiencing the world (e.g. the experience of seeing the color red)." I will refer to the idea that access and phenomenal consciousness can come apart as the Disjunction Thesis (DT). If DT is true, no experimental evidence of any sort could suggest the existence of blindsight in monkeys, or even in the easier case of humans. From my definition, an experimenter must show three things to establish blindsight: 1) intelligent behavior in response to visual information, 2) consciousness, and 3) the absence of phenomenal vision. If DT is true, 3 can never be established.

To see this, it's helpful to consider Dennett and Cohen's "perfect experiment." In this imagined scenario, a subject has the right connections severed such that information about the color of an object is correctly inferred in visual areas but is never transmitted to higher-level areas that deal for example with language. As a result, the color information is not accessible (it is not in access consciousness). The subject will insist that she is colorblind. Nonetheless, certain adherents to DT would claim that she has phenomenal experiences of color since the networks creating the color information have not themselves been damaged. The only damage is to projections from the visual areas to other functional areas. Under the DT scheme, just about any phenomenal claims may fly. A subject insisting that there is no phenomenal consciousness cannot be believed, because she can only report on what's in access consciousness, and the link between access and phenomenal consciousness may break. Under DT, investigating phenomenal consciousness appears scientifically hopeless.

If I believed in the truth of DT, the paper would end here. However, I reject the thesis. The unfalsifiability of DT is icky. Scientific theories that can't be disproved don't *feel* right. Of course, this is a minimal argument, but I think that saying more about why unfalsifiability should disqualify the thesis would require an argument from philosophy of science, and that is beyond the scope of the paper. My primary objection then is that I find DT to be conceptually implausible. When I reflect on my concept of consciousness, there is no separation between access and phenomenal. The idea that I may be conscious of something without being aware of

that is simply incompatible with how I think about consciousness. I can't bring myself to describe what's going on in a case like the "perfect experiment" as consciousness without access. Some questions that arise include, for whom or what is there a phenomenal experience? Are we to believe that the circuit is having an experience of the redness of red?

Having rejected DT, I can proceed to discussing Cowey and Stoerig's evidence. Moving forward, I will treat phenomenal and access consciousness as identical. I now address the criteria for blindsight sequentially. Part 1 is perhaps the most obviously implied by the evidence. The monkeys are exposed to flashing lights (visual information). Then, they do a very effective job of pressing the buttons that the lights direct them to press and so receive a reward (intelligent behavior in response). This is all neatly summarized in the paper's figure 2c. Part 2 is slightly less clear. Are monkeys conscious? I follow Nagel in considering something to be conscious if there is something that it is like to be that thing. I strongly suspect that monkeys are conscious. Moreover, all (nearly all?) modern neuroscientists believe that monkeys are conscious. Surely, the experiences of monkeys are incomparably different from our own. Nonetheless, it's hard to think that the lights are off. Human brains, which give rise to the conscious experiences we are so familiar with, share massive structural similarities with primate brains.

Part 3 is where the rubber meets the road. Are the monkeys succeeding because they are *seeing*? The authors attempt to show that no, in fact, the monkeys have no phenomenal vision in the right visual field. By seeing and phenomenal vision, I am talking about what it is like to see, the experience of vision. Their key evidence that this is absent in the right hemifield is that when the monkeys are presented flashing lights on the right side and are given the option of reporting having seen nothing (as opposed to being forced to choose where they saw something), they virtually always choose the seeing nothing option. However, there's one big problem. When the light flashes on the right, the authors reward the monkeys equally for picking the correct and seeing nothing options. This fails to rule out the possibility that the monkeys' vision is simply degraded with respect to the right visual field. When forced to choose, they are somewhat less competent at picking the correct light when the light on the right-side flashes. From figure 2c, it appears that, on average, the vision-damaged monkeys make the correct choice in the right-side case 90% of the time. On the other hand, those monkeys make the correct choice in the left-side case nearly 100% of the time. A smart monkey with somewhat degraded right hemifield vision (but who is not blind in the relevant sense) could simply learn to ignore the information coming from the right field in the second part of the experiment. The expected value of ignoring the right hemifield ("ignore") is strictly higher than that of paying attention to both hemifields ("paying attention"). There is nothing to gain by paying attention to both, but there is something to lose. By paying attention to both, for example, the monkeys may sometimes select the right side when no light flashes, losing out on reward.

The authors attempt to address this concern by dialing up the luminance on the right lights. The thinking is that if the monkeys' vision is simply somewhat degraded (rather than absent), the higher luminance would remove the disadvantage created by the degradation, removing the edge enjoyed by ignore over paying attention. Since the monkeys in the higher luminance case don't alter their behavior, the authors conclude that their right hemifield vision is not degraded but absent. I'm not entirely convinced by this trick. It's not obvious to me that higher

luminance would fully undo the deficit created by vision degradation. Another response to the payment problem might be that one of the vision-damaged monkeys, Dracula, was not payed at all in the second part and behaved just as the other vision-damaged monkeys. This helps the authors but doesn't fully resolve the issue. In the case where a monkey is not rewarded at all, any pattern of behavior would be rational.

Is there any way to answer this concern about reward? I think that there may be. My proposal for an alternate experiment has two parts. First, administer to the monkeys some assessment to gauge risk preference.[1] I expect that the monkeys would be slightly risk averse to risk neutral, as most human subjects are. Second, with this data in hand, progressively increase the reward associated with correctly picking the right-side. If you observed the exact same behavior as in Cowey and Stoerig but with 1) monkeys known to be a little risk averse to risk neutral and 2) substantially greater payouts on the right-side, I claim that that would constitute significant evidence for blindness. Blindness would be the most compelling explanation for the inconsistency in the monkeys' expressed risk preferences and their behavior. If they're nearly risk neutral and rational (i.e., consistent) agents with accurate information about their cognitive capacities, they should begin choosing the right-side. If they do not, the monkeys are either not rational or not operating with accurate information. Psychological evidence shows that rationality cannot be assumed in general. However, a significant deviation would be unexpected. The more plausible explanation, then, would be that the monkeys do not have accurate information about their cognitive abilities. They do not know that they have blindsight (or don't appreciate how good their blindsight is) but they do know that they are lacking in phenomenal vision! This combination biases whatever crude assessment they are making of the probability of "guessing" correctly.

So, where does this leave us? I have argued that the disjunction thesis is false and claimed that Cowey and Stoerig have successfully produced low-grade evidence for blindsight. They run, unsurprisingly, into the challenge of assessing consciousness in creatures that cannot talk to us and tell us what and whether they are experiencing. Their response to this challenge is admirable but incomplete. Their evidence for an absence of phenomenal vision is not a smoking gun. I would be pleased to see further studies that replicate the authors' findings with the modifications to the experimental design that I propose.

[1] These abound in the behavioral economics literature. A canonical example involves asking participants to choose between receiving $1 with certainty and a bet in which they receive $0 with some probability and $3 with the complement probability. Experimenters vary the bet probabilities and observe where subjects switch from choosing the sure thing to the risky proposition, or the other way around. I see no reason why a similar experiment could not be carried out with monkeys