On the Demonstration of Blindsight in Monkeys

CHRISTOPHER MOLE AND SEAN DORRANCE KELLY

Abstract: The work of Alan Cowey and Petra Stoerig is often taken to have shown that, following lesions analogous to those that cause blindsight in humans, there is blindsight in monkeys. The present paper reveals a problem in Cowey and Stoerig's case for blindsight in monkeys. The problem is that Cowey and Stoerig's results would only provide good evidence for blindsight if there is no difference between their two experimental paradigms with regard to the sorts of stimuli that are likely to come to consciousness. We show that the paradigms could differ in this respect, given the connections that have been shown to exist between working memory, perceptual load, attention, and consciousness.

1. Cowey and Stoerig's Claim

In 'Blindsight in Monkeys', a brief article published in *Nature* in 1995, Alan Cowey and Petra Stoerig described a pair of experiments in which three monkeys, following brain surgery, behaved in a way that was, they said, '*reminiscent of* patients with blindsight' (emphasis added). This conclusion is a modest one, and there is little here that might be objected to, even by a sceptical philosopher.

Later in that same year, however, in an article entitled 'Visual Perception and Phenomenal Consciousness', Cowey and Stoerig were writing as if they had established the *actual existence* of blindsight in monkeys, or at least provided strong evidence for it. Subsequent commentators have certainly taken them to have established this much stronger conclusion: That there really is blindsight in monkeys. Do Cowey and Stoerig's results support this strong conclusion, or do they only warrant the conclusion that lesioned monkeys behave in a way reminiscent of blindsight? The present paper makes the case for modesty here. We will suggest that Cowey and Stoerig's experiments do not provide good evidence for the strong conclusion that there really is blindsight in monkeys.

One reason why their interpreters may have overestimated what Cowey and Stoerig's work shows is that the interpreters are under a misapprehension as to the details of Cowey and Stoerig's experimental procedures. Andy Clark was certainly under a misapprehension of this sort when, in a paper from the year 2000, he wrote about 'the test used by Cowey and Stoerig to convince us of genuine blindsight in a monkey.' The description that he gives of Cowey and Stoerig's test

Address for correspondence: Christopher Mole, Philosophy Department, Washington University, One Brookings Drive, Campus Box 1073, St. Louis, MO 63130, USA.
Email: cmole@artsci.wustl.edu; sdkelly@fas.harvard.edu

Mind & Language, Vol. 21 No. 4 September 2006, pp. 475–483.

is only slightly wrong, but the place at which it goes wrong is an absolutely crucial one. Here's how Clark describes the experiment:

> [T]he monkey was trained to touch a screen in response to a visually presented target. When forced to respond to a target presented in the blind hemifield, the monkeys often succeeded. Yet they *also* indicated (by touching a different part of the screen) that on these forced trials they judged there to be no visual target present (Clark, 2000, p. 31).

Clark's mistake was to think that the monkeys successfully responded to a stimulus whilst *also* indicating that no visual target was present. What the monkeys actually did was to successfully locate visual stimuli, but then subsequently, and in a *different* experimental context, show that they failed to notice stimuli of the same sort. It was *not* that the monkeys responded to stimuli and then indicated that they couldn't see *those very stimuli*. It was, rather, that stimuli *of the same sort* were successfully responded to in one experiment but were treated as if absent in a second experiment. (We shall present Cowey and Stoerig's method in more detail in the next section.)

Cowey and Stoerig, and those who, like Clark, have drawn on their work, take this behaviour to be evidence of blindsight because they take the failure to notice the stimuli in the second experiment to be evidence that the successful performance in the first experiment wasn't guided by a conscious experience. But, of course, the behaviour of the monkeys in the second experiment only provides evidence of what the monkey could see in the first if the two experiments don't differ with regard to the sorts of stimuli that the monkey is likely to be conscious of. The case to be made for the existence of blindsight in monkeys on the basis of these experiments depends on the assumption that stimuli that aren't seen consciously in the second experiment couldn't have been seen consciously in the first.

In section three we shall provide a reason for thinking that the assumption on which Cowey and Stoerig's case for blindsight depends is at best questionable, and very possibly false. The two experiments may very well differ in respect of what the lesioned monkeys are likely to be conscious of.

2. Cowey and Stoerig's Experiments

Cowey and Stoerig removed the striate cortex of the left cerebral hemisphere in three macaque monkeys. The analogous lesion in humans is associated with blindsight in the right half of the visual field. Cowey and Stoerig's lesions did not cause their monkeys to lose the ability to locate visually presented stimuli in either side of space. The ability to locate stimuli was demonstrated in the following 'localization paradigm': The monkeys were trained to respond to the presentation of a stimulus in any of the four corners of a screen by touching the place where the stimulus had been flashed up. After the monkeys were lesioned the stimuli presented in the right half of the visual field became *slightly* less effective at eliciting responses, but only slightly. The monkeys were still well above chance at responding to visual stimuli, wherever they were presented in the visual field, correctly locating the stimuli more than 90% of the time. Cowey and Stoerig's hypothesis is that this above chance performance was blindsight. The possibility they are interested in (and that they are taken to have established as actual) is that the monkeys did not really have any *experience* of the stimuli—that although they responded to stimuli on the right hand side, they were not aware of them.

The lack of awareness that would qualify the monkeys' performance as a case of blindsight was purportedly established in a second experiment: As in the first experiment, the monkey had the task of touching the place at which a stimulus had been briefly shown. (Stimuli were presented for 750ms.) The experiments differed as to where the stimuli could appear: Stimuli in the second experiment were presented in any of five locations in the left half of the visual field—the half unaffected by the lesion. They were also presented, although much more rarely, in one location in the right half of the visual field—the half that *is* affected by the lesion. Stimuli were presented in this location on five percent of trials. The other difference between the two experimental procedures is that in the second experiment, unlike the first, no stimulus was presented at all in half of the trials. On those trials when no stimulus was given, the response from the monkey that was rewarded was the touching of a white area displayed on the left (unaffected) half of the visual field.

The conditions in this 'signal-detection' experiment may be summarized like this: The monkey was rewarded for touching the place where the stimulus had been shown, if there was such a place, and when no stimulus was shown, it was rewarded if it touched the white area. On five percent of the trials the stimulus was presented on the right hand side of the screen—in that part of the visual field affected by the lesion. The rest of the time the stimulus, if there was one, occurred in any one of five locations on the left.

It is the way in which the monkeys responded to the occasional targets on the right-hand side that led Cowey and Stoerig to conclude that they were never aware of the stimuli presented on that side, despite their success, in the previous experiments, in locating them. The monkeys with lesions touched the white area when the stimulus was presented on the right—that is, they made the response that they had been trained to make in cases where no stimulus was present for them to be aware of. A monkey that, although trained in the same way, *hadn't* been lesioned, treated the stimuli on the right, just as she treated the stimuli on the left, by touching the place in which they had been shown.

These observations tell us something about the nature of the impairment resulting from lesions analogous to those responsible for blindsight in humans. But, as we have already noted, they do not demonstrate the existence of blindsight in monkeys unless an assumption is made about the similarity of the two experimental paradigms with regard to which sorts of stimuli the lesioned monkeys will be conscious of. That assumption, though on the face of it plausible, is in fact questionable, and in the light of some empirical results that we shall review shortly, is quite possibly false.

3. The Confounding Effects of Attention

The kind of difference between paradigms that would undermine Cowey and Stoerig's method is a difference that would have an effect on the likelihood of a stimulus in the right half of the visual field being noticed. If stimuli in the right half of the visual field are noticed in the localization paradigm but are not noticed in the signal detection paradigm then, although the lesioned monkeys in the signal detection paradigm treat such stimuli as if they were absent, they might have been able to see stimuli of the same sort perfectly well when in the localization paradigm, in which case the successful performance in the localization paradigm is not performance in the absence of consciousness, and so is not a case of blindsight. One reason why a stimulus might come to consciousness in the localization paradigm while going unnoticed in the signal detection paradigm is that the monkey was *paying attention* to the relevant part of space in the one paradigm, but not in the other.

There are several factors that might cause the monkey to pay attention to a part of space in the localisation paradigm but not to pay attention to that part of space in the signal detection paradigm. Some of these factors are straightforward (and they have been taken into account in some of Cowey and Stoerig's later work). Other factors are less obvious (and these *have not* been taken into account). One of the straightforward factors is this: In the signal detection paradigm stimuli occur much less often on the right than on the left, but in the localization paradigm the stimuli are just as likely to appear on either side. That, one would have thought, might incline the monkey to attend to the left side of space in the signal detection paradigm, but not in the localization paradigm. Another straightforward factor that might produce a difference in attention is that in the signal detection paradigm 'null responses' are made in the white area, which is over on the left side of space. That also might incline the monkey to attend to the left side of space in the signal detection paradigm, but not in the localization paradigm. It has been shown, however, that these simple factors cannot provide the whole explanation of the behaviour that Cowey and Stoerig observed. These straightforward explanations for a difference in attention (and some others suggested in the work of Moore, Redman and Gross, 1998) have, to some extent, been controlled for in later work (Stoerig, Zontanou and Cowey, 2002). We favour a more complicated explanation for attention being more focussed in the signal detection paradigm.

Before we go into the details of our preferred explanation, it will help to have the explanation before us in outline: The factor that we suggest could produce a different deployment of attention in the two paradigms is *working memory load*. There is good empirical evidence to show that, in humans, attention is affected by changes in working memory load, and we know that Cowey and Stoerig's two paradigms differ in the load that they place on working memory (because the task to be kept in mind in the localization paradigm is just 'touch what flashes up', but the task to be kept in mind in the signal detection paradigm is a more complicated one: 'touch what flashes up if anything does, and otherwise touch the white area'). Keeping in mind the rules of the more complicated paradigm will place heavier demands on working memory, and these heavier demands may lead to a difference in the monkey's focusing of its attention. This difference in attention could in turn lead to a difference in which stimuli the monkey notices, and so the fact that the stimuli are unnoticed in the second experiment doesn't show that they were unnoticed in the first. If this is right then it follows that there is no reason to think that performance in the first paradigm was blindsight.

The next section goes through this line of thought in more detail, and, in particular, it explains and gives evidence for the connection between differences in attention and differences in working memory demands.

4. Factors Influencing the Deployment of Attention

A connection between the storage of information in working memory and the extent to which visual attention can be focused has been demonstrated by Jan de Fockert, Geraint Rees, Christopher Frith and Nilli Lavie. In a 2001 paper they show that human subjects are distracted more from a name-reading task by the concurrent presentation of pictures of faces if they are trying to keep in mind a hard-to-remember sequence of numbers, than if they are trying to keep in mind an easy-to-remember sequence. What this suggests, as de Fockert *et al.* emphasize, is that there is some way in which we use working memory in focusing our attention.

Their brain imaging data suggest part of the story about why it is that the faces are less distracting when the load on memory is low. The brain scans show that the trials in which subjects need only remember an easy sequence of number are trials in which visual cortex is less strongly activated by the presentation of faces.

When the subjects in de Fockert's experiment only needed to remember an easy sequence of numbers there was plenty of working memory left over to be given to the task of reading the names, and so attention was closely focused on the names, and the concurrently presented faces were therefore processed less. But when de Fockert *et al.* used up their subjects' working memory with the irrelevant business of number storage this left less working memory available to serve the word reading task. Because the reading task now had less memory serving it the subjects' attention was now less focused on that task, so peripheral stimuli were processed more and therefore provided a greater distraction.

The crucial lesson for our present purposes is just that peripheral stimuli are attended to less and processed less when the subject is engaged in a task that is served by more working-memory. With that in mind we can begin to see why it is that the monkeys who had the task of localization (in which the working memory load is just 'touch the place where a stimulus is presented') might have focused attention less than the monkeys who had the task of signal detection (in which the working memory load is 'touch the place where a stimulus is presented, if there is such a place, or else touch the white area').

We can infer, from the difference between paradigms as to the working memory load, that attention will be more focussed in the signal detection paradigm. From this, in turn, we infer that there will be a difference between the paradigms as to which stimuli will come to consciousness: Peripheral stimuli are more likely to *go unnoticed* in the signal detection paradigm. It is this last point that enables us to explain Cowey and Stoerig's data without crediting the monkeys with blindsight. The first of the steps towards this point—the step from a difference in memory load to a difference in attention—is, as we have just seen, an empirically supported one. For the second step—from a difference in attention to a difference in consciousness—we are again on empirically supported ground, and on ground that is more familiar. It is very likely that the difference in attention will lead to a difference as to which stimuli reach consciousness, especially if, as the de Fockert brain scans suggest, it is a difference that corresponds to a difference in the extent to which the presentation of stimuli activates relevant regions of sensory cortex.

Many empirical results support the idea that by increasing the focus of attention one will decrease the chance of peripheral stimuli being seen. There has been a recent flourishing of memorable experiments in which it is made vivid that the way in which one pays attention to the visual scene can make a big difference to what it is that one is conscious of. In Mack and Rock's inattentional blindness experiments, for example, the engagement of attention by a peripheral stimulus results in the failure to notice stimuli that are presented briefly in the middle of the visual field (Mack and Rock, 1998). In Simons and Levin's 1998 experiments subjects who are distracted in a naturalistic setting fail to notice a craftily executed change in the identity of the person to whom they are speaking, while most of the subjects in Simons and Chabris's 1999 experiments failed to notice the appearance of a man in a gorilla suit when they were attending to a game of catch that was taking place around him. These much discussed results show that what one is attending to can make a difference to what one is conscious of, and they show that the difference that attention makes to consciousness can be a surprisingly profound one.

Since, due to working memory differences, the monkeys in the signal detection paradigm might very well have focused *attention* more than the monkeys in the localization paradigm, they might very well be *conscious of fewer things*. It seems very likely that, if working memory affects attention in the way suggested, there will be stimuli that the monkeys can notice in the localization paradigm, but that they will fail to notice in the signal detection paradigm. The influence of working memory on attention can explain why attention is more focused in the signal detection paradigm than the localization paradigm. If the effects of attention focusing are sufficiently profound (as, in a monkey, they might be) then this focusing of attention can explain why stimuli are not noticed in the signal detection paradigm, even though similar stimuli are noticed in the localization paradigm. *Moreover*, the influence of *perceptual load* on attention can explain why this difference as to what is noticed is only found in the monkeys that have lesions to the visual cortex.

This last form of influence on attention is also empirically well established in work by Christopher Frith, Nilli Lavie and Geraint Rees. The influence that perceptual load has on attention is an influence on the extent to which unattended stimuli are neglected. When attention is focused on a central stimulus the processing of peripheral stimuli is always suppressed—that's just part of how attention is paid—but the *extent to which* peripheral processing is suppressed depends on the perceptual load of the attentively performed task. The effect of perceptual load is most vividly demonstrated in Rees, Frith and Lavie's 1997 paper in which they show (through psychophysical effects and through fMRI observations) that, in human subjects, the processing of irrelevant motion in the periphery of the visual field was reduced when the subject's task was changed from the low-load task of classifying the words presented in the centre of the visual field as upper or lower case to the high-load task of classifying them as mono- or bi-syllabic. The lesson to be learnt for our present purposes closely parallels the lesson that we took from Lavie's work with de Fockert *et al.*: As one increases the proportion of a subject's perceptual resources that an attentively performed task demands, one decreases the extent to which unattended stimuli will be processed.

Exactly what perceptual load *is* is a subject of ongoing enquiry, but we do not need a well-articulated notion of perceptual load in order to see that it is a factor that could come into play in Cowey and Stoerig's experiments. Perceptual processing resources in lesioned monkeys must be more limited than in unlesioned monkeys. They are, after all, missing a large part of striate cortex. In each of Cowey and Stoerig's experiments, therefore, the perceptual resources of the lesioned monkeys will be under a greater load than the perceptual resources of the unlesioned monkeys. This could explain why the influence of attention on working memory leads the lesioned monkeys, but not the unlesioned monkeys, to miss some of the stimuli in the signal detection paradigm. Here, then, is a complete, non blindsight-mentioning explanation of what Cowey and Stoerig observed: All of the monkeys successfully locate the stimuli in the localization paradigm because all of the monkeys have a conscious experience of the locations of these stimuli, although the perceptual load of locating these stimuli is higher for the lesioned monkeys than for the unlesioned monkeys. All the monkeys have more focused attention in the signal detection paradigm than in the localization paradigm due to the higher working memory load of the signal detection task. This focusing of attention does not prevent the unlesioned monkeys from seeing the stimuli that are presented on the right side of the screen in the signal detection task because stimuli that aren't attended are still processed when perceptual load is low, but for the *lesioned* monkeys the focusing of attention in the signal detection paradigm *does* prevent the stimuli on the right side of the screen from being seen, because for them the perceptual load is higher and so the processing of peripheral stimuli is reduced.

5. Strengthening Cowey and Stoerig's Case

The effects of working memory load and perceptual load could combine to provide us with a complete, non blindsight-based explanation of the behaviour that Cowey and Stoerig observed, and therefore show that, after Cowey and Stoerig's evidence is taken into account, the verdict on monkey blindsight must be that the existence of the phenomenon remains unproven. It need not remain unproven forever. Our rival explanation is susceptible to empirical refutation, as all good explanations should be, and it may well be that a case for the existence of monkey blindsight can be made by a strengthened version of the argument that Cowey and Stoerig offer.

One way to strengthen the case would be to show that changes in the monkeys' attention do not affect the likelihood of stimuli like those in Cowey and Stoerig's experiments being noticed. It could be that differences in a monkey's attention only make a difference to the likelihood of a stimulus being noticed when the stimulus in question is a peripheral one, or is not relevant to the task at hand, or occurs in a crowded visual environment. The stimuli in Cowey and Stoerig's experiments are none of these things. Perhaps it could be shown that their stimuli are *so salient* as to be exempt from attentional effects on the likelihood of their being noticed. If that were so then we would no longer have a rival explanation for the lesioned monkey's failure to detect stimuli in the signal detection paradigm and the case for monkey blindsight could be made persuasive. There are some reasons to think that Cowey and Stoerig's results could successfully be strengthened in this way. Experiments in which people fail to notice unattended stimuli are typically those in which the unnoticed stimuli lack the task-relevance and perceptual salience of Cowey and Stoerig's stimuli, and it is also true that when the task is just the simple detection of a target's location, as it is in Cowey and Stoerig's experiments, then it makes only a little difference to the performance of human subjects when attention is focused elsewhere (Braun and Julesz 1998). But these are experiments conducted on normal, unlesioned adult humans, and not on brain-damaged monkeys. It could well be that the effects of attention in lesioned monkeys are not limited in this way. It is an open question how the focussing of attention affects the performance of brain damaged monkeys, and so it is an open question whether Cowey and Stoerig's case could be strengthened in this way.

If the case is to be strengthened in this way then it must be strengthened *by some further empirical results*. One cannot just rule our rival explanation out of court on grounds of far-fetchedness or biological implausibility. Stoerig, Zontanou and Cowey admit that their claims cannot be proved 'beyond a shadow of a sceptic's doubt' (Stoerig, Zontanou and Cowey, 2002), and it is true that scientific practice does not require radical sceptical scenarios to be ruled out, but our doubts are not the doubts of a hyperbolic sceptic, and nor is our rival explanation an unparsimonious one. Our explanation does not require the postulation of any effects beyond those that are already known to exist.

Nor do we require the postulation of an implausible discontinuity between men and monkeys. One cannot rule out our explanation by reference to considerations of ‘evolutionary continuity [and] similarity between the nervous systems’ (Stoerig, Zontanou and Cowey, 2002). Monkey brains are like human brains in many respects, but there are many other respects, size being the most obvious, in which the two brains are different. In the absence of a good theory of consciousness any one of these differences is a candidate for grounding a difference between monkey and human in the way that brain damage affects conscious experience. One of the reasons for pursuing research into blindsight, and into the parallel effects of lesions in animals, is that this might cast light on the question of which features of the brain are relevant to the production of consciousness. To appeal to considerations of similarity between monkey brain and human brain when arguing that the lesions producing blindsight in humans will produce blindsight in monkeys begs the question against the relevance to consciousness of all of those features that differentiate monkey brain from human brain. It therefore begs precisely the questions that we want research on blindsight to answer.

Philosophy Department
Washington University;
Philosophy Department
Harvard University

References

Braun, J. and Julesz, B. 1998: Withdrawing attention at little or no cost: Detection and discrimination tasks. *Perception and Psychophysics*, 60, 1, 1–23.

Clark, A. 2000: A case where access implies qualia? *Analysis*, 60, 1, 30–38.

Cowey, A. and Stoerig, P. 1995a: Blindsight in monkeys. *Nature*, 373, 247–249.

Cowey, A. and Stoerig, P. 1995b: Visual perception and phenomenal consciousness. *Behavioural Brain Research*, 71, 147–156.

De Fockert, J., Rees, G., Frith, C. and Lavie, N. 2001: The role of working memory in visual selective attention. *Science*, 291, 1803–1807.

Mack, A. and Rock, I. 1998: *Inattentional Blindness*. Cambridge MA: MIT Press.

Moore, T., Redman, H. R. and Gross, C. G. 1998: Man, monkey and blindsight. *Neuroscientist*, 4, 227–230.

Rees, G., Frith, C. and Lavie, N. 1997: Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science*, 278, 1616–1618.

Simons, D. and Levin, D. 1998: Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*, 5, 4, 644–649.

Simons, D. and Chabris, C. 1999: Gorillas in our midst: sustained inattentional blindness for dynamic events. *Perception*, 28, 1059–1074.

Stoerig, P., Zontanou, A. and Cowey, A. 2002: Aware or unaware: Assessment of cortical blindness in four men and a monkey. *Cerebral Cortex*, 12, 565–574.