"The Moral Landscape" – Sam Harris

Harris asks how empirical knowledge – in particular, new scientific knowledge in the field of neuroscience – can be used to answer moral questions.

He starts by showing the problems with a widely held view that this is an impossible task: that facts and values are entirely unrelated, that right and wrong are independent of the consequences of an action, that morality is entirely a function of the social context.

Unfortunately, Harris chooses to adopt a position as extreme as the one he attacks: he claims that facts and values are essentially indivisible, that right and wrong are dependent on nothing other than consequences of actions, which are in principle in the realm of the empirically knowable, and consequently that "science can determine human values". Yet his arguments, as far as I can see, do not warrant such an extreme position. What they do support is a much more common sense approach: that values are affected by facts AND by intuitions, interpretations, feelings and so on; that practical morality involves consideration of consequences AND other matters such as rules and motivations; and hence that science can indeed *contribute to* answers to moral questions without in general determining them. In fact, it turns out that it is only the latter "common sense" position that needs to be defended in order to validate the writer's investigations into the implications of experimental psychology and neuroscience.

Harris's starting point is that conscious experience is the only possible ultimate criterion of right and wrong – that which is not experienced by a conscious being must be morally neutral. This seems pretty convincing to me. More questionably, he then equates conscious experience exclusively with the *consequences* of an action to determine the morality of the action, and attempts to answer the usual objections to such "consequentialism" by saying that they define consequences too narrowly, and consideration of the widest consequences will meet the objections.

There are several problems glossed over here. Firstly, are we talking about the consequences of applying a proposed rule to a general class of actions or about the consequences of a particular action in a specific context? Harris mostly seems to mean the first but is never explicit. In this case the relationship between classes of action and typical consequences is certainly a matter for empirical study, rational evaluation and provisional conclusions. But in practice the context of every action is in some sense unique, so the question is always likely to arise whether conclusions about a general class of action are valid in the particular case - for example if it seems evident that the consequences in the current case will be different from those in the typical case. On the other hand - especially if each case is to be treated on its merits - the problem arises that the consequences may be in principle unknowable at the point of decision: to take just one example, even the "direct" eventual consequences may depend on some technical innovation that has not yet been made. Along with such objective uncertainty is subjective uncertainty about one's own capacity for judging consequences, given the human capacity for self-deception and rationalisation. Hence morality necessarily involves strategies for dealing with uncertainty. To put this another way, while consideration of the consequences of actions is a necessary input into moral debate, it is not sufficient: consideration of the sources of uncertainty and their implications is also required and often just as important.

Secondly, motivation also falls under the description of "conscious experience". If it can be morally desirable in its own right that certain consequences should be experienced, why cannot it be similarly desirable in its own right that certain motivations are experienced? This is particularly significant in

the case of actions such as democratic voting where the consequences of a single action or inaction may be of negligible consequence and the motivation is not so much "having a consequence" as "being part of something one approves of".

A third problem with Harris's defence of consequentialism arises because consequences very importantly include psychological consequences. Such consequences are inevitably a function of the moral beliefs of the person affected, hence an unavoidable circularity is introduced. For example, if the consequences of some action (or inaction) are believed to be shameful, then to avoid the experience of shame the action should be avoided. Since shame itself is a social construct which cannot be abolished by decree, it is on such self-justifying cases that belief in moral relativism seems often to rest.

Another example where Harris seems blind to the significance of circularity is in his attack on belief in "free will" - since, it seems to me, it is a conception of the possibility of choice that makes consciously willed choice possible. Again, however, the object of his attack seems to be a straw man, a concept of free will as a mysterious independent and influence-free agent that is the *cause* of thought, rather than as a decision-making process that includes conscious thought with choice in mind. The common sense view sees the notion of free will (in the latter sense) as accurately reflecting the situation from the point of view of somebody faced with a decision, and causality as an accurate description of the same situation looked at from the point of view of an outsider, *with no conflict between these accounts.* Harris's objection to this "compatibilism" seems to be that people don't really believe it but are using it to cover an inability to believe in causality as applied to themselves. But if I say (for example) that I decided not to take a short cut path because on reflection it looked too muddy, I am acknowledging that the sensory input from the muddy path "caused" the decision, while still properly regarding it as "my" decision because this and other considerations were only brought together in my conscious mind. In any case, the significant outcome of this discussion by Harris is to make the case that moral responsibility is a matter of degree and punishment should be driven by deterrence and rehabilitation rather than retribution - which can equally well be argued from a "compatibilist" viewpoint.

Having picked on "consequences" as the criterion for moral choices, it is necessary to distinguish good consequences from bad, and here Harris proposes "well-being" as a useful label for what is to be maximised. Compared with "the greatest good" this is certainly more helpfully focussed, without the narrowness of "the greatest happiness". It is then that the title of the book is explained: the peaks and troughs of a "landscape" provide a metaphor for the variations of well-being that are supposed to result from following different sets of ethical prescriptions. This is perhaps more convincing if envisaged on a geological time-scale, so that the peaks grow or decline as humanity's understanding of ourselves and technological knowledge develops over time. (The greatest well-being for a hunter-gatherer community is unlikely to follow from exactly the same rules as would be the case for an industrialised society).

Harris moves onto his strongest ground in considering some tentative insights into the nature of belief suggested by neurological research. This suggests that belief in what is right and wrong uses broadly the same bits of the brain as belief in what is true or false. Perhaps this should be no great surprise. In both cases we seek to adopt beliefs that can be presented to ourselves as free from internal contradiction, and as consistent with the remainder of our world-view. However, as I commented above, a close relationship between facts and values does not necessarily imply that they are indistinguishable.

In the final chapter, Harris addresses some of the difficulties of his thesis posed by recent research. What are the moral implications if differences between politically left and right values are truly a matter of inborn personality differences, as Jonathan Haidt claims? How can we hope to calibrate supposedly objective measures of well-being when people are so inconsistent in their own judgement of their welfare, as Daniel Gilbert reports? Do we seek to optimize human experience as it occurs to the "experiencing self" or as it is remembered by the "remembering self", as distinguished sharply by Daniel Kahneman? Harris is better at clarifying the questions and expressing his optimism about there being solutions than in finding a path towards the latter, but he is certainly persuasive in identifying research such as this as essential grist to the mill of ethical debate.