Ethics // Spring 2023

Handout 3

Ethics and neuroscience: Greene

Contrast with Haidt's project. Greene begins with the major premiss shared by Haidt that there is a dual-process cognitive system, where one process is 'reasoning' and the other 'intuition' (see the summary in Haidt's Table 1). Haidt's conclusion was that our moral judgement is largely determined by intuitions, even if occasional exceptions are possible—e.g., with moral philosophers. 35

Crucially, however, Haidt did not distinguish between deontological and consequentialist judgements: his conclusion was meant to be applied across the board. Greene will argue that deontological judgements are emotional and 'intuitive', whereas consequentialist (including utilitarian) judgements are rational and 'cognitive' (where this term is technical). So the two cognitive tracks are occupied by two different moral judgements. 36

Deontology and consequentialism. Greene has an interesting new suggestion. Normally, you'd think that deontology and consequentialism are two philosophical theories, in a sense *inventions*. So we might have: 37

> *Deontology:* the moral value of an action is determined by the agent's following certain unassailable principles.
> *Consequentialism:* the moral value of an action is determined by its consequences with regard to a certain good like universal welfare (in utilitarianism).

But Greene rejects this natural thought. Instead, deontology and consequentialism are two patterns of responses to a morally relevant situation. They are intrinsic to our cognitive mechanisms. Since our concern is not metaphysics, let's downplay Greene's talk of essences and natural kinds. Let's simply say this: Although we may define and use the terms 'deontological' and 'consequentialist' arbitrarily as we please, in fact the persistent appeal of deontology and consequentialism is explained by the presence of those two cognitive patterns. 38

To sum up: Just like Haidt, we define deontology and consequentialism in terms of their 'characteristic judgements', i.e. verbal responses to a morally relevant situation. This division is not arbitrary: it reflects the underlying 'essences' of deontological and consequentialist philosophies. 39 40

Cognition and emotion. As Greene uses 'cognition', it is a behaviourally neutral representation triggering no particular automatic response on the part of the agent. 'Emotion', on the other hand, does trigger it. Further, they can be distinguished by their location in the brain: pre-frontal cortex for cognition, amygdala for emotion. 40

So there are four possibilities to consider: 41

(*a*) Both deontological and consequentialist judgements are cognitive. (Rawls, Kohlberg)
(*b*) Both deontological and consequentialist judgements are emotional. (Haidt)
(*c*) Deontological judgements are cognitive, consequentialist judgements are emotional. (Partly a traditional view)
(*d*) Deontological judgements are emotional, consequentialist judgements are cognitive. (Greene)

Diagnosing moral dilemmas. The puzzle to explain is the divergence we observe in the consensus over the Trolley dilemma and the Footbridge dilemma (see the setup in the text). Why do people think that it is OK, if not required, to divert the trolley even if this causes a man to die, but not OK to push the man with the same result? A popular philosophical explanation is that in the trolley case the man dies merely as a side effect. But in the footbridge case he is used directly as a means to avoid harming others—something that is morally inadmissible. 42

Yet we should consider the Loop case (details in the text) where the person is, again, used as a means, yet the consensus is that diverting the trolley then *is* permissible. So the philosophical explanation can't be right.

Greene offers another explanation. The difference is the mode of violence. In the trolley and loop cases the violence is remote and impersonal. In the footbridge case the violence is salient. And when 43

it is salient, an emotional response against violence is triggered. That's why we tend to judge this action unacceptable. In the footbridge case this response is not triggered. Thus there, the judgement is formed by 'reasoning', with the typical consequentialist reasons taking over.

This claim can be tested empirically. When you think of the trolley case, the areas of your brain associated with 'cognition' should light up. When you think of the footbridge case, the areas of your brain associated with 'emotion' should light up. That is what indeed was observed. Similarly, it was shown that those who judged it permissible to push the man on the tracks took a longer time to respond than those who didn't. This matches the idea that the former relied on reasoning, whilst the latter on emotion, with the consequentialist and deontology judgements elicited as a result.

Reflections. Is this a convincing argument? There is no point in arguing with empirical results, but an alternative explanation may have to do with our intuitions about what constitutes an action. When I'm diverting the trolley, that's what I am doing—diverting the trolley. This is an intuition about what *action* is being performed. No violence is involved in the action itself. When I push a man on the tracks, my intuition is that this is a violent action. That's how I classify it (in my brain?). So the alternative explanation is that we respond negatively to violent actions (where violence is an integral part of the action), but not to actions that merely *lead* to violence, as in the trolley case. I have intuitions about what constitutes *my* actions, and I distinguish them, still intuitively, from their consequences. The contrast, according to this explanation, is between violent/non-violent actions, rather than between salient violence/remote violence.

YSB