

MIT OpenCourseWare
<http://ocw.mit.edu>

24.120 Moral Psychology
Spring 2009

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.

XXII The Metaethical Implications of Empirical Work

META-ETHICAL BACKGROUND

Some very broad (and horribly simplified) distinctions:

Realist (or objectivist) theories take moral claims (and therefore moral judgements) to be factual claims, broadly similar to the claims of science, or, at least, our claims about ordinary medium sized objects.

Anti-realist theories take moral claims to be very different to scientific claims. It comes in various forms: expressivism, subjectivism, response-dependence etc.

There is much debate about how to draw the boundaries between the two.

Kantian theories take moral understanding to be a standard form of intellectual understanding (in terms of the categorical imperative). The motivation to act morally is now derived from this intellectual understanding. The immoral person acts irrationally; our motivation to act morally is, in some important sense, akin to our motivation to act rationally.

Humean theories understand moral judgement as basically driven, not by the intellect, but by sentiment. That is, we have certain *feelings* upon which our judgments are based. For Hume himself the central feeling is sympathy: the suffering, or happiness, of others produces similar suffering or happiness in us, in proportion to how closely we identify with them. Moral motivation is now understood as desire, not reason, based: we act morally in order to quieten these feelings.

Both accounts require an empathetic element, though in very different ways: the former to implement the categorical imperative, at least in one of its formulations (act only on maxims that you could will to be universal law); the latter to generate the sympathetic response. Kantian theories tend to be realist, and Humean theories tend to be anti-realist, but this falls short of any straightforward entailment.

Contrast meta-ethical theories with normative theories, e.g. consequentialism (e.g. utilitarianism) or deontology (e.g. follow the ten commandments). These are not theories about the *nature* of morality, but rather about what one ought to do in certain situations.

JUDGEMENTS AND MOTIVATION

A further complexity: a full moral sensibility requires not just the ability to make moral judgements, but also the tendency to be motivated by them. These might be thought of as separate faculties; or as somehow combined (so that someone who isn't motivated by a moral judgement hasn't really grasped it).

ONE POSSIBLE MODEL FOR ETHICS: MATHEMATICS

There is such a subject as moral philosophy for much the same reason that there is such a subject as the philosophy of mathematics.

Scanlon, 'Contractualism and Utilitarianism'

No one doubts that the psychology of mathematics is interesting. It shows how we do things, and the mistakes we are liable to make; for instance, the literature on street mathematics. Compare probabilistic reasoning, and conditionals. (Even here though, controversy, for instance, different readings of the conditionals in the Wason selection task.)

But: there is no expectation that work on the psychology of mathematics will lead to new mathematical results. Nor is there any sense that mathematics consists in giving an account of intuitive mathematical reasoning.

ANOTHER MODEL: LINGUISTICS

Hauser. His parallel concerns the nativism and modularity of ethics. But if we take the parallel seriously there is a further upshot: the linguistic intuitions (responses) of normal speakers are the very subject matter of linguistics

WHAT'S IMPLAUSIBLE ABOUT THE LINGUISTIC MODEL?

The nature of moral reasoning; the possibility of moral progress.

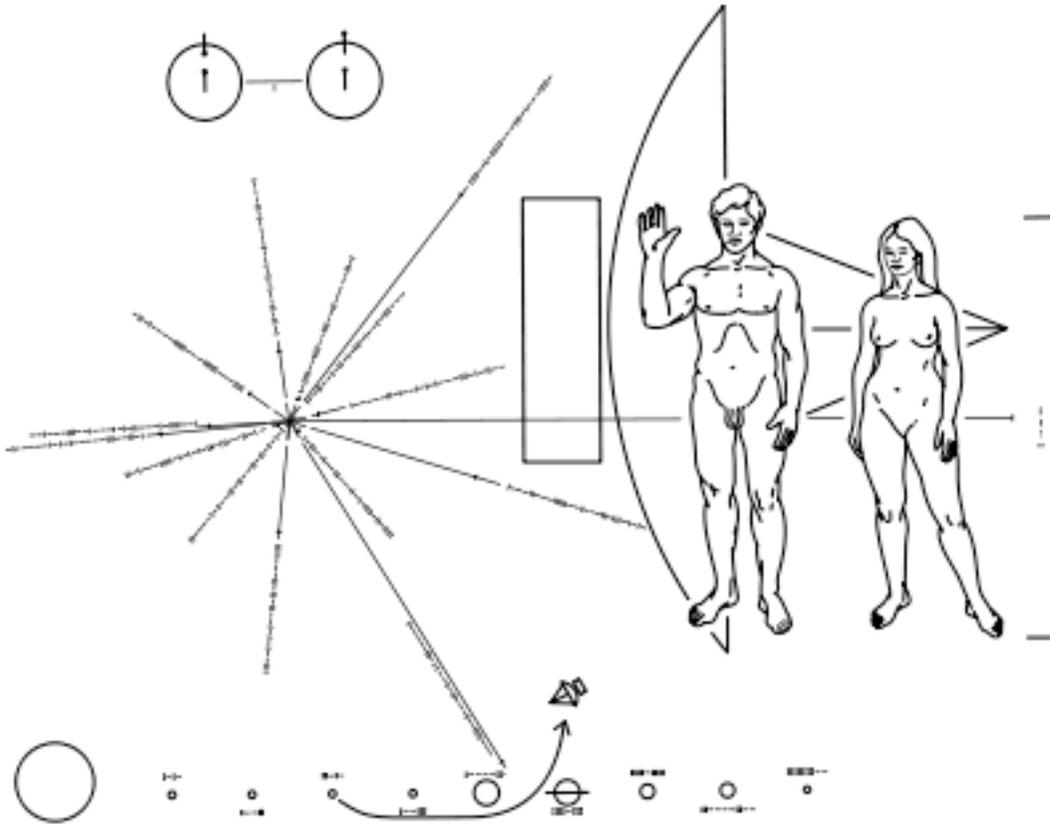
If ethics is a theoretical inquiry that can be approached by rational methods, and that has internal standards of justification and criticism, the attempt to understand it from outside by means of biology will be much less valuable. This is true for the same reason that the search for a biological explanation of mathematical or physical theories, or biological theories for that matter, would be relatively futile.

Nagel, 'Ethics without Biology'

Reflected in one philosophical notion of intuition: not fast, parallel, unconscious cognitions, but a faculty that gives us access to morality in something like the way that mathematical intuition gives us access to mathematics. The propositions whose truth is grasped in this way are typically seen as not in need of justification, rather like axioms. Nonetheless judgments made in this way are revisable in the light of greater expertise, again like mathematics. So intuitions can be educated.

WHAT'S IMPLAUSIBLE ABOUT THE MATHEMATICAL MODEL?

First reason: Carl Sagan's Pioneer 10 plaque.



Various mathematical data (in binary). Plausible. Human figures, man with 'hand raised in a gesture of goodwill'. Not so plausible.

Second reason: motivational connection. Harder to identify. Both premise and validity are controversial.

Linguistics is clearly Protagorean; plausibly morality is too.

REFINING THE PROTAGOREAN POSITION

Response dependent account: the Protagorean concepts are those whose extension is set by human responses. More precisely, a concept F is broadly response-dependent iff the basic equation is *a priori* true:

x is F iff x elicits response R from subjects S in conditions C

This is very broad. Amongst moral theories it captures two kinds of Protagoreanism: (i) metaethical: the nature of the moral concepts is determined by human attitudes (e.g. ideal observer theories); (ii) normative: the content of the moral concepts concerns human attitudes (e.g. utilitarianism).

A restriction that the subjects be *observers* (rather than *participants*) and that the conditions be standard, restricts us to the metaethical and gives the more common:

x is F iff x elicits response R from competent observers in standard conditions

A special case, judgment dependence:

x is F iff x is *judged* to be F by competent observers in standard conditions

Contrast linguistics and mathematics. The analogous equations for the linguistics are arguably *a priori* (a sentence is grammatical iff it is judged grammatical) whereas those for the latter are not. In this respect, morality, on this Protagorean view, is like linguistics. Unlike linguistics though, if we want to make sense of the nature of moral reasoning, we will want to be pretty fussy about the competent observers and the standard conditions.

But rather than thinking that this cuts against empirical research, this is the first place that it can help. Compare even color. Ideal observers and standard conditions are discovered empirically.

A second and greater worry concerns whether responses other than judgments are relevant, especially since these responses might pull in different directions to the judgments. This is so even for linguistics. There is the performance/competence distinction, where one tends to favor judgments; but also cases where one favors performance:

Since 1972 we have collected 12 cases of speakers who use positive *anymore* quite freely though their introspective judgments were entirely negative. Faced with a sentence like *John is smoking a lot anymore* they said they had never heard it before, did not recognize it as English, thought it might mean ‘not smoking’, and showed the same signs of bewilderment that we get from Northern speakers outside the dialect area. This describes the behavior of Jack Greenberg, a 58-year-old builder raised in West Philadelphia. His introspective reactions were so convincing that we felt at first that we had to accept them as valid descriptions of his grammar. Yet two weeks later, he was overheard to say to a plumber, “Do you know what’s a lousy show anymore? Johnny Carson.” (Labov, *What is a Linguistic Fact* (1975), pp. 34-5)

So for linguistics we have to choose between people’s beliefs about what is grammatical, and their actual linguistic behavior. For ethics there are even more possibilities:

- what you judge is right under imaginary conditions;
- what you judge that you would do under imaginary conditions;
- what you judge is right under real conditions;
- what you judge that you will do under real conditions;

- what you actually do under real conditions;

or some other response under imaginary or real conditions (e.g. an emotional or sympathetic response).

Compare the humorous: is this what one judges funny; or laughs at; or what moves one in some other way?

Here is part of the argument between those who take the analogy of morality with linguistics seriously, and the sentimentalists. For the sentimentalists our emotional responses are central. So even if these don't get through to our judgments, we might still accord them some weight.

Another example: Military literature on soldiers' refusal (inability?) to kill, especially when this is up close (main source: Dave Grossman's *On Killing*). People don't know this by introspection; so asking about hypothetical cases won't take us very far.

Two considerations:

- (i) we might worry that so much work has gone into the trolley case when this involves killing; reactions in other cases (e.g. Nichols and Mallon's teacup case) are somewhat different. It would be good to have other experiments that are morally serious, but don't involve killing.
- (ii) who is the ideally placed agent in these situations? A person who would have to do the killing; or a bystander. And in each case, how far away should they be?

SO WHAT MIGHT WE LEARN FROM EMPIRICAL WORK?

- (i) There is all kinds of complexity in our reactions, and it is only going to be untangled after a lot of empirical investigation.
- (ii) Whether a response-dependent account is plausible—and if so, one involving which responses, by which agents, under which circumstances—will only come clear once we are clearer on that complexity.
- (iii) So any basic equations can't be *a priori*. We need to get at the constitutive point differently. It is not even *a priori* that ethical facts supervene on human responses. Perhaps we should talk instead in terms of necessary *a posteriori* truths; the reasons for not taking that path now look less compelling than they once did. But it is not clear that we can even individuate the individual equations.
- (iv) If we end up building a theory around reactions other than judgments, and if the relevant subjects are participants rather than observers, then the distinction between metaethical and normative theories may become blurred.
- (v) We might get clearer on the motivational force of moral judgments. It might not be terribly philosophically interesting. Reluctance to see oneself as behaving badly. Developmental evidence. Findings from empirical economics.