PETER RAILTON

REPLY TO JUSTIN D'ARMS

1. INTRODUCTION

It was a pleasure to receive Justin D'Arms' thoughtful comments on Part I of *Facts, Values, and Norms*, and it will be a challenge to reply. D'Arms calls first for clarification of my view about how empirical theory bears on moral theory, and then for an explanation of how the relationalist form of moral realism I defend can avoid relativism.

2. NATURALISM AND MORAL THEORY

D'Arms does not doubt that the answers to questions about the nature and authority of morality depend in some measure on psychological and social facts – he accepts the method of broad reflective equilibrium, which calls for compatibility between moral theory and our best available empirical theories. He doubts, however, that fine-grained empirical considerations – such as the specific character of purported explanations involving moral properties – can contribute much to the resolution of important issues in moral theory. As a first step in responding, let me point to some influential *critiques* of substantive positions in moral theory that appear to depend upon rather definite claims about the explanation of moral phenomena.

2.1. *Epistemic Critiques*

Many philosophers hold that an individual's belief that p, even if true and justified, will not constitute *knowledge* that p

unless the fact that *p* plays an appropriate role in explaining it. More ambitiously, some philosophers have argued that *all* legitimate reasoning, including reasoning in the moral sphere, is a form of "inference to the best explanation." Such doctrines have been used to criticize realism and objectivity in ethics, by arguing that explanations adverting to moral facts are always pre-empted by, or inferior to, explanations citing only non-moral phenomena.[1]

2.2. *Semantic Critiques*

Mental or linguistic access to a domain of facts and properties, some philosophers claim, requires that these facts and properties have exerted an appropriate shaping force on the evolution of our thought and discourse. If this is right, then a defense of morality against an error theory based on referential failure must provide a credible account of how moral thought and language could track putative moral facts and properties.[2]

2.3. *Empirical Critiques*

Recent work in psychology has called into question common-sense and philosophical conceptions of moral judgment and sentiment (see Haidt, 2001), rational agency (see Bargh and Chartrand, 1999; Railton, 2004), traits of character (see Doris, 2002; Vranas 2005; Sabini and Silver, 2005), the relation of moral judgment to motivation (see Roskies, 2003), and the nature of subjective well-being and its place in the psychology of motivation and action (see Kahneman *et al.*, 2004). It is too early to say whether the empirical theories lying behind these particular challenges will stand the test of time, but their message for moral theory is already plain: we philosophers are not on such safe ground as we suppose in treating broad reflective equilibrium as a relatively undemanding standard.[3]

To assess the force and purported implications of these critiques of substantive positions in moral theory, we have no alternative but to investigate in some explanatory depth and

detail the genesis and character of moral thought, language, and practice. This claim would come as no surprise to Hume or Nietzsche.

3. NATURALISM AND NORMATIVE ETHICS

Still, D'Arms is surely right that appeal to empirical considerations can go only so far – we must connect with straightforwardly normative questions. Elsewhere I have attempted to make the case that naturalistic accounts of intrinsic value, moral rightness, and moral value can sustain appropriate normative roles, at least up to a good approximation (see Railton, 1984, 1989, 1992). Rather than repeat those arguments here, I would like to consider a particular way in which empirical considerations matter even when we engage directly in reflective normative judgment.

Let me follow D'Arms and discuss moral value. He appears to find this concept serviceable enough, and is worried less about its meaning than its measure – or rather, about what role empirical theory could play in *establishing* which measure of moral value, if any, is correct. Is it morally better to maximize overall good, or to distribute goods equally? Do all increases in well-being count in favor of an outcome, or only those that satisfy needs? The resolution of such questions is a straightforwardly normative matter, if any is.

How do we proceed with reflecting on such questions? Typically, we begin (where else?) by consulting our intuitive moral judgments. We count as evidence for a normative view that it fits our considered moral judgments, especially those widely shared and highly confident. However, recent work in cognitive social psychology and behavioral economics invites us to reflect self-critically on the epistemic authority we normally accord to intuitive judgments. Framing effects, anchoring effects, cognitive dissonance, heuristics, affect-priming, and the like play an uncomfortably large role in explaining our intuitive judgments. Merely by changing the mode of description of a given choice situation, for example, experimenters can induce subjects to switch their intuitive

judgments.[4] So when we engage in reflective normative judg-
ment, we should add to the evidence of our normative intu-
itions whatever evidence we have about how intuitive
judgment works and how this can shape or color our reac-
tions in ways we do not deem morally relevant.

Consider then the choice mentioned above, between opti-
mific *vs.* need-based conceptions of moral value. In making
this choice, we reflect on what matters most to us. Given this
framing, it seems plausible enough to say that we are chiefly
concerned that our needs be met. And yet if we look at the
actual goods at stake, and at the psychological and historical
evidence concerning human motivation, a different picture
emerges. Individuals regularly forgo meeting needs in order
to advance other goals, often symbolic in nature, and some-
times at great risk to their lives. Moreover, a sizable body of
psychological literature indicates that we often hold mistaken
views about what we really want and how much we want it.
This is especially striking in the case of individual concep-
tions of what will make us happy.[5] Such evidence, too, be-
longs with the intuitions upon which we reflect when asking
what really matters in life.

4. REALISM AND RELATIVISM

One ground for doubting whether the sort of naturalistic real-
ism I propose could afford a *vindicatory* explanation of moral
thought and practice is that it might turn out to be too sub-
jective or relativistic to capture the objective, universal pur-
port of moral judgment. D'Arms raises such an objection in
discussing an example I use to suggest that color properties do
not afford a good model for moral value, since (among other
things) color predicates appear to have their extension rigidi-
fied with respect to normal human color responses as they
actually are. Applied to the moral case, this could have the
result of projecting our actual, normal moral responses and
judgments into other possible worlds, including those where
individual psychology or the conditions of life are signifi-
cantly different. I see this as a parochialism contrary to the

objectivity to which morality aspires, and defend instead a functional relationalism that analogizes 'good' to functionally-characterized concepts such as "sweet-tasting" or "nutritious".

D'Arms is unsure whether actual humans accord intrinsic moral import to genetic kinship, as my own example supposes, and so proposes an alternative example. Take it as given that Present-Day Humans embrace *optimific ethics* – we are disposed to see social outcomes as morally valuable strictly as a function of net overall well-being. Future Humans, by contrast, are assumed to embrace *sufficiency ethics* – they are disposed to judge the moral value of outcomes strictly in terms of whether basic needs are met. He then asks: Does my relationalism commit me to the view that this change in moral sensibilities makes it the case that questions of moral value among New Humans are properly governed by sufficiency ethics rather than optimific ethics?

I appear to face something of a dilemma. If I say 'yes', then my relationalism would be at risk of becoming relativism. The authoritative moral code for a society is too dependent upon its own moral sensibilities.[6] If, on the other hand, I say 'no', then what would be the point of my original example? – After all, it, too, involved a change in moral sensibilities.

My response depends upon the nature of the changes that resulted in, or accompanied, the emergence of New Human moral sensibilities. At least three possibilities are relevant. (1) Changes in social life and individual psychology have altered the features that explain why Present-Day Humans attribute moral significance to issues about well-being beyond sufficiency. (2) Social life and individual psychology remain largely unchanged, but empirical or philosophical discoveries correct erroneous beliefs sustaining Present-Day Humanity's sense that differences in well-being beyond sufficiency have moral relevance. (3) Social life, individual psychology, and non-moral beliefs remain largely unchanged, except for the changes of thought and feeling involved in the shift of moral sensibilities itself.

Case (1) is closest to my original example. Unlike case (2), that example does not attribute any mistake of fact or

philosophy to those people who currently see genetic related-ness as inherently morally relevant. And unlike case (3), my example assumed significant changes in social practices and individual psychological development, which in turn ex-plained the changes in moral sensibility. Concocting examples in which normative principles other than our own apply is no easy matter. Some would say it is impossible, since a great deal in the normative realm strikes us as necessary truth. My hope was that intrinsic concern with genetic kinship was a sufficiently self-standing component of current moral sensibil-ities that a description of a society in which it does not have inherent moral authority is at least *prima facie* intelligible. My implicit model was the way in which spouses, and adop-ted children and adoptive parents and siblings, can develop deep mutual bonds and attitudes of special concern unrelated to genetic kinship. The example simply generalized these cases. The idea was that a shared "way of life" could emerge in which the specific relationship of sharing DNA simply lost any special significance in shaping how we grow up and age, the identities we form, and the ties of unconditional affection and concern we sustain.[7]

The difference between relationalism and the worrisome sort of relativism mentioned above is thus easy to state: rela-tionalism does not accord a society's prevailing moral sensi-bilities a society privileged status. But more plausible forms of relativism permit a gap between a society's actual norms and the norms appropriate to it, much as informed preference theory permits a gap between actual and authoritative prefer-ences. How does my relationalism differ from such a view?

Relationalism holds that the appropriate morality for a society depends upon its ways of life, the psyche and lived experiences of its members, and the psychological, social, and technical possibilities open to them.[8] D'Arms' example asks us to imagine a society of New Humans who share our views about well-being, but who differ markedly in moral sensibili-ties. According to them, morality has nothing to say about whether (for example) a wealthy parent ought to do more to encourage his children's aspirations and promote their

well-being than to insure that their basic needs are met, or whether the social distribution of well-being and goods above the level of sufficiency should reflect individual contribution and equal access. So dramatic a change in what counts as morally relevant could only arise from very substantial changes in human motivational psychology, the importance to us of our goals and aspirations, the character of our social relations, and the possible psychological, social, and technical alternatives. If no such changes accompany the changed sensibility, then I am not at all tempted to say that the New Human's moral sensibility is appropriately regulative for them.

In making this claim, I am drawing upon the higher-order Aristotelian principle according to which the proper regulative principle for a thing depends upon the nature of that thing (cf. Rawls 1971, 424 ff). I also draw upon characteristics of the moral point of view and associated principles concerning moral relevance and reasons. Does holding this much fixed when considering other possible worlds betray a rigidity that makes the analogy between moral judgment and color judgment appropriate after all? My response in the original paper was that the characteristics upon which I draw – impartiality, generality, benevolence, treating like cases alike, etc. – are not substantive moral principles but rather higher-order principles constitutive of the moral point of view. To use another helpful Rawlsian terminology, they belong to the *concept* of morality rather than any particular normative or meta-ethical *conception* of it – they are characteristics in virtue of which judgments count as moral rather than, say, prudential, politically expedient, epistemic, or aesthetic. I am not free to alter these characteristics if I wish to provide an example in which norms different from our own nonetheless strike us as *moral* norms.[9]

At the same time, however, a would-be naturalist must be able to explain how to account for such relatively *a priori* constitutive principles within the constraints of an empirical theory. A breezy answer can be given by gesturing toward a Quinean or Wittgensteinian account in terms of distinctive patterns of linguistic and inferential use (e.g., relative immunity from revision; see

Railton, 2001). But a fuller response calls for something no one currently has, namely, a convincing account of the normativity of meaning. Here the naturalist has the poor consolation that his non-naturalistic competitors enjoy no advantage in the face of this deep puzzle.

NOTES

[1] For questions of explanatory pre-emption and supervenience in the moral case, see Gibbard (2003). Harman (1996) makes quality of explanation the lynchpin of his moral relativism. I discuss moral explanation and address Harman's view in Railton (1998).

[2] See for example Boyd (1988) as well as the recent debate over "Moral Twin Earth" (Horgan and Timmons, 1991).

[3] John Rawls, for example, writes concerning constraints empirical theories of human nature might impose on normative theory:

... beyond the lessons of historical experience and such bits of wisdom as not relying too much on scarce motives and abilities (say, altruism and high intelligence), there is not much to go on. [Rawls (1993, p. 87)]

[4] See, e.g., Kahneman and Tversky (2000).

[5] See Gilbert et al. (2003).

[6] An exception is Harman's realist relativism (Harman, 1996).

[7] It is worth noting that in various societies the notion of kin embraces individuals without biological kinship, and that such symbolic kin relationships can take on primary importance in determining identity.

[8] Compare Rawls' discussion of the transition from the general to the special conception of justice (Rawls 1971, 62f).

[9] This, in itself, is not a normative claim on their behalf. Perhaps mankind has more reason in some global sense to pay attention to schmoral norms than moral norms ("schmoral norms" are like moral norms except not benevolent).

REFERENCES

Bargh, J.A. and Chartrand, T.L. (1999): 'The Unbearable Automaticity of Being', *American Psychologist* 54, 462–479.

Boyd, R (1988): 'How to Be a Moral Realist', in G. Sayre-McCord (ed.), *Essays on Moral Realism*, Ithaca: Cornell University Press.

Doris, J. (2003): *Lack of Character: Personality and Moral Behavior*, Cambridge: Cambridge University Press.

Gibbard, A. (1990): *Wise Choices, Apt Feelings*, Cambridge: Harvard University Press.

Gibbard, A. (2003): *Thinking How to Live*, Cambridge: Harvard University Press.

Gilbert, *et al.* (2003): 'Immune Neglect: A Source of Durability Bias in Affective Forecasting', *Journal of Personality and Social Psychology* 61, 617–638.

Haidt, J. (2001): 'The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment', *Psychological Review* 108, 814–834.

Harman, G. (1996): 'Moral Relativism', in G. Harman and J.J. Thomson, *Moral Relativism and Moral Objectivity*, Oxford: Blackwell.

Horgan, T. and Timmons, M. (1991): 'New Wave Moral Realism Meets Moral Twin Earth', *Journal of Philosophical Research* 16, 447–465.

Hume, D. (1888): *Treatise of Human Nature*, Ed. L.A. Selby-Bigge, Oxford: Oxford University Press.

Kahneman, D. and Tversky, A. (eds.) (2000): *Choices, Values, and Frames*, Cambridge: Cambridge University Press.

Kahneman, D., *et al.* (eds.) (2004): *Well-Being: Foundations of Hedonic Psychology*, New York: Russell Sage Foundation.

Railton, P. (1984): 'Moral Realism', *Philosophical Review* 95, 163–207.

Railton, P. (1989): 'Naturalism and Prescriptivity', *Social Philosophy and Policy* 7, 151–174.

Railton, P. (1992): 'Some Questions about the Status of Morality', *Philosophical Perspectives* 6, 27–53.

Railton, P. (1998): 'Moral Explanation and Moral Objectivity', *Philosophy and Phenomenological Research* 58, 175–182.

Railton, P. (2001): '*A Priori* Rules: Wittgenstein on the Normativity of Logic', in P. Boghossian and C. Peacocke (eds.), *New Essays on the A Priori*, Oxford: Oxford University Press.

Railton, P. (2004): 'How to Engage Reason: The Problem of Regress'. In R.J. Wallace, *et al.* (eds.), *Reason and Value*, New York: Oxford University Press.

Rawls, J. (1971): *A Theory of Justice*, Cambridge: Harvard University Press.

Rawls, J. (1993): *Political Liberalism*, New York: Columbia University Press.

Roskies, A. (2003): 'Are Ethical Judgments Intrinsically Motivational? Lessons From Acquired Sociopathy', *Philosophical Psychology* 16, 51–66.

Sabini, J. and Silver, M. (2005): 'Lack of Character? Situationalism Critiqued', *Ethics* 115, 535–562.

Vranas, P. (2005): 'The Indeterminacy Paradox: Character Evaluations and Human Psychology', *Noûs* 39, 1–42.

Department of Philosophy
University of Michigan
435 South state street
Ann Arbor, MI, 48109-1003
USA
E-mail: prailton@umich.edu