ON NEEDING A MORAL THEORY: RATIONALITY, CONSIDERED JUDGEMENTS AND THE GROUNDING OF MORALITY

KAI NIELSEN

I

Meta-ethics, as traditionally conceived, is an attempt to elucidate the meaning of moral concepts or terms, to specify the logical status of moral utterances and to characterize and elucidate the structure of moral reasoning. In the past decade, there has been a healthy tendency to move away from such stultifying preoccupations on the part of moral philosophers.[1] Without sacrificing rigor, a broader understanding of what a responsible doing of moral philosophy might come to has quietly emerged. Even when, as with J. L. Mackie and G. Harman, there is a return to such analytically traditional concerns, it is in a wider and more fruitful context. I shall construe my topic broadly and conceive of meta-ethics as an inquiry into the 'foundations' of morality and into what Rawls has characterized as moral methodology.

What I shall say here will be mainly negative. Where, in the latter half of my essay, I do say something more positive concerning wide reflective equilibrium (WRE) my remarks are both tentative and extremely ambivalent. But, for the most part, particularly in the first sections, what I say will consist of dicta and some arguments about what we cannot do and indeed now should no longer try to do, rather than a positive account giving new directions. In part this is generated by a certain reaction to the work of John Rawls. Like many others, I see it as the most careful, systematic and reflective development of moral theory in our time. It is self-consciously aware of alternatives, profoundly aware of the history of the subject and, in spite of its ubiquitously moderate and almost self-effacing tone, it is a bold work while remaining thoroughly architectonic. Notwithstanding all this sophistication, integrity and years of careful reflection (reflection which was hardly carried on in cultural isolation), it seems to me that the upshot of the extensive and varied criticism that Rawls's work has received is to show that his endeavour fails. Most centrally he has not shown that his two principles of justice can be derived from an original position that has not already been so skewed as tendentiously and question-beggingly to produce this result. There can be several, perhaps many, contractarianly conceived original positions from which a variety of quite different principles of justice could be derived. And we have not been shown why we should accept Rawls's characterization of the original position. We have not been given a characterization of the original

---

[1] I have attempted to characterize the traditionally central problems and positions in my "Problems of Ethics" and my "History of Ethics" in *The Encyclopedia of Philosophy*, Vol. III, Paul Edwards (ed.), (Collier-Macmillan Ltd.: London, 1967), pp. 100–134.

position that any rational agent simply must accept. Even the position Rawls
is most concerned to defeat, namely utilitarianism, can be given a reasonable
contractarian articulation. In fine, Rawls has not shown that rational con-
tractors or reflective and impartial agents, reasoning in accordance with the
moral point of view, must opt for his principles or anything like his principles.
Pluralism is still strongly in the field as are various versions of utilitarianism.
Even perfectionism and natural rights accounts are not decisively routed.

In turning to Rawls's account, we could initially have reasonably hoped
for a narrowing of our historical options. What we discover, as a result of the
intensive and often acute discussions that have emerged in reaction to Rawls's
work, is that, in spite of the power and the probity of Rawls's account, the
old options are all there before us. They have been refined and we have
hopefully a refined methodology with which to examine them, but they all
are still with us. Similar things can be said for his methodology. His very
contractarian method is very much up for grabs and his appeal to considered
judgements in reflective equilibrium, something which I take very seriously
indeed, has been widely rejected, mistakenly I believe, as a refurbished version
of intuitionism.

Rawls sought to develop a systematic normative ethics which would afford
us an Archimedian point. That is to say, construing social justice as the first
virtue of institutions, it would give us a set of moral principles in virtue of
which we could objectively assess moral practices, various social arrangements
and institutions and indeed whole moralities and societies. It is surely not
surprising, given the history of the subject, that such a grand enterprize
should fail. What we need to ask, given the state of development of philosophy
and our enhanced understanding of the human sciences, and given the failure
historically of such grand endeavours, is whether we should continue to try to
do such things in moral philosophy. Rawls is often compared to Sidgwick,
and with good reason, but in one important respect at least they are very
different. Rawls articulates one complicated method of ethics and constructs,
using that method, a systematic normative ethics. Sidgwick displays several
methods of ethics and attempts to show how they are in intractable conflict.
Both Rawls and Sidgwick self-consciously develop their accounts with their
predecessors in mind, but Rawls, like Kant and Mill, articulates a distinctive
method of ethics and normative ethical theory, while Sidgwick, by contrast,
argues that there is a dualism in practical reason and that no single method of
ethics establishes itself. In this respect, though surely not in others, Sidgwick
is closer to the dominant modern temper. Rawls's attempt, like Kant's and
Mill's, is both ambitious and demanding but neither an understanding of the
history of the subject nor an awareness of the upshot of the critical reception
of Rawls's work should make us sanguine concerning the success of such
architectonic endeavours.

To make such contention does not require, though it is plainly compatible
with, an acceptance of ethical skepticism or non-cognitivism. However, it is
not unreasonable, even given the above, also to believe that there are certain

moral truisms or normative commonplaces with which any normative ethic must square. I have in mind such things as that it is wrong to torture the innocent, to lie, to fail to keep one's promises or to break faith with people. It isn't that these things can never be done, no matter what the circumstances. Rather what is the case is that to do any of these things is *ceteris paribus* wrong. There is always a presumption against doing them and in the case of torturing the innocent that presumption is very stringent indeed. There are also quite different commonplaces such as the recognition that pleasure is good and pain bad, that it is a good thing to develop one's powers and to have meaningful work and meaningful human relationships. Ethical skepticism and its country cousins aside, all moral theories accept such truisms.[2] They, of course, give them somewhat different readings and a different placement or weight in their theories, but all of them accept them. Moreover, against an ethical skeptic or subjectivist who would not accept these truisms, a parallel argument to a Moorean one against epistemological skepticism is perfectly in place. If there is a philosophical account of morality that requires us to reject these truisms, it would be more reasonable to reject that account and accept the truisms than to reject the truisms and accept the philosophical account which required us to reject the moral truisms.

However, such a "critical commonsensism" will not carry us very far against the moderate skeptic or against plausible versions of moral relativism. The various historically influential and plausible normative ethical theories, as I noted, all accept these truisms and, in the great ideological debates about society, it is not the case that some groups accept some of them and others reject them. Something similar obtains for the extant moralities of at least most societies. The dispute here is about the overall import, reading and placement of these truisms and about the inability of the various philosophical, theological or scientific accounts of morality to provide a cogent Archimedian point for the assessment of rivial moralities, ways of life and social theories. This, in a way Moore failed to see, was Westermarck's starting point and it should be the starting point of any ethical skeptic or relativist.

Repeatedly moral philosophers have claimed a systematic knowledge of good and evil or have claimed that we have, or can come to have, cogent grounds for accepting a given comprehensive normative ethical theory. The claim, typically, isn't just that such an account is *consistent* with reason but that reason *requires* a certain theory; there would be some failure in rationality for a properly informed, conceptually sophisticated person not to accept such a systematic account. But the point — a point the moral skeptic can drive home — is that no one has yet made good on the objectivist's claim to provide such an account.

Given past historical failures, Rawls's account is a watershed here. If his

[2] Given the vagueness of "accept" it may even be correct to say that in some weak sense the ethical skeptic or subjectivist accepts them too. But what is relevant to my remarks in the text is that they could hardly accept them as something they take to be true or as beliefs that a rational person properly informed must have and act in accordance with.

account fails, and fails to the extent I have claimed, given his very reasonable contraints on theory construction and acceptance, its sophistication, care and probity, the state of the philosophical art, the development of the social sciences and given the cultural ambiance in which Rawls's account was developed, is it very reasonable to believe that some future systematic norma- tive ethicist, constructing a normative ethical theory on the traditional model, is likely to succeed? Even after Rawls's efforts, and the efforts of other contemporary analytically oriented philosophers, all the traditional accounts appear at least to remain in the field. This is hardly — or so it appears — like the development of science and it seems at least to give considerable force to the ethical skeptic's or moral relativist's case. It isn't that we never know what is right or wrong, good or bad or that in morals it is as reasonable to believe one thing as another. Rather what is the case is that we lack a system- atic knowledge of good and evil. We have no sound account of the development of morals such that we are justified in claiming that there is a normative account of morality which we can rightly claim is the correct account or the most adequate approximation of a correct account.[3]

Perhaps the disquietude that is likely to be engendered by the realization that we do not have, and are not likely to get, such an Archimedian point is misplaced. It does not take an inordinate amount of political savvy and human understanding to recognize that there is much that is just plain rotten about our societies and that there are alternative ways of conceiving of society

---

[3] Here I should perhaps add a skeptical note concerning the strength of my claim stated above and developed in my essay — a claim I remain ambivalent about. If there is much to be said for some of the claims of Kuhn and Feyerabend, there is perhaps not as great a difference between science and ethics as I have given to understand and perhaps there is more of a modest rational development in ethical theory paralleling that of science, than I have acknowledged. All the traditional accounts, it might be argued, remain in the running both in ethics and in science, but in both we get progressively refined theory construction. These accounts give us increasingly more adequate conceptualizations of the domains of ethics and science respectively. There is rational development in both, though no account is through and through victorious. Rawls's metaphor of an Archimedian point, on such an understanding, is an unfortunate one, but we can see, both from the example of his own theory, arising as an alternative to utilitarianism, after the steady refinement of utilitarianism in our period, and from his account of wide reflective equilib- rium, where and how the development of moral theory occurs. Neither Rawls's normative theory, nor some rational reconstruction of it, anymore than any other carefully crafted and historically important theory, wins the day, but he moves our understanding of morality and its justificatory base forward. Surely this is a possible view and an attractive one. I think the most serious doubts concerning such a view are (a) the sort that I bring up in the latter part of my text and (b) worries about the closeness of the analogy between ethics and science. It seems to me dubious, to say the least, to claim that the traditional scientific accounts (even given some plausible rational reconstruction of the more archaic ones) are all still in the running, while this is not an implausible thing to say about ethical theories. It is certainly not clear that we have anything sufficiently unproblematical and hierarchically ordered to give us anything of much value as a systematic normative ethic. For a suggestive account, cutting, I believe, in the other direction, see Thomas Nagel, *Mortal Questions*, (Cambridge University Press; London, 1979), pp. 128–146.

and our lives together in which life would not be so rotten.[4] While the unflinch-ing seeing of this, and the taking of it to heart, require considerable moral integrity and an ability to overcome ideological mystification, there is no need, in order to see steadfastly the ills of our condition and to gain a conception of liberating alternatives, to have a developed moral theory. To gain such a perspective, we do not have to set out a complete and sound moral theory. Almost any of the traditional moral theories — the various varieties of utilitarianism, pluralistic deontology, justice as fairness, perfection-ism, even some natural rights theories — can provide sufficient "foundations" for critiquing such states of affairs, such conditions of society.[5] There can, of course, be no critiquing of society without some conception of what a good society would look like. But it is not clear that this requires anything as complicated as a normative ethical theory. In critiquing society and setting out the requisite conception of a good society, what is humanly most important to try to sort out is what we can know or reasonably believe about our society and ourselves, what we can know or reasonably believe about our own situation, its class formations, patterns of dominance, conditions of living and the like and what objective possibilities there are for a change in our condition. In social critique, it is these things we very much need a good understanding of. Moral theory here is ancillary. Here we need a sophisticated political sociology and political economy, not, at least as a matter with a claim to any priority or emphasis, a more developed moral philosophy.

However, we should also recognize that significant work of a reasonably extensive sort in political sociology and political economy will not be *wertfrei*. An important question is whether the value judgements that are crucial here are not such that we can find, for the problem at hand, sufficient support in what I have called moral truisms and in a rather theoretically unramified moral understanding or whether they need for their support the sharpening of moral theory. Over social policy issues, questions will emerge about the ranking and the perspicuous setting out of values. There may be a modest place here for moral theory.

How important to such questions of ranking is the fact that no overall account of morality has showed anything even approximating signs of winning out? We should not, over such questions, be *a prioristic*. Until we have actually worked out in detail what we want to say about these policy issues, using a sophisticated political sociology and political economy, we cannot know how much, if any, moral theory is required here; until that work has actually been done, until, more generally, there actually have been extended attempts to develop a normatively oriented social critique which is built up out of a thorough understanding of the human sciences, we will not know what kind of a moral theory, if any, we will need as one of the instruments of social

[4] Noam Chomsky shows this brilliantly and unforgettably in his various political writings.
[5] This has been powerfully urged in an important essay by Michael Teitelman, "On The Theory of The Practice of The Theory of Justice", *Journal of Chinese Philosophy*, Vol. 5 (1978), pp. 217–247.

critique or how extended and systematic it needs to be. Certain things in our
societies about alienation, certain extreme and degrading inequalities, power-
lessness, exploitation and the lack of autonomy are perfectly evident to
anyone not thoroughly ideologically bamboozled. And they do not require
for their detection and suggested remedy a sophisticated moral theory.[6] It
may very well be that to acknowledge them and to acknowledge and come to
understand their wrongness, we do not require any moral theory at all. We, of
course, require some understanding of what society is like and what the human
possibilities are and we need some moral understanding and sensitivity. But
that is very different from having a moral theory. Moreover, it is chastening
for philosophers to recognize that the various even remotely plausible philo-
sophical candidates for a plausible moral theory all provide us with sufficient
grounds for such condemnation. What remains unclear is whether we need
any moral theory at all for rational versions of such critiquing or for specifying
the content of a good society. Even on the assumption that we do, just how
fine grained a moral theory do we need to say what we need to say in such
contexts concerning justice and its relation to equality and liberty and about
a characterization of human good? It is not evident that we need much in the
way of a moral theory to provide an adequate rationale for such a social
critique and a delineation of a good society.[7] However, what is crucial to
reiterate here is that until we have actually carried out such social analyses,
we will not know how much we need in the way of a moral theory. It is quite
possible that we will not need much of anything. If we keep a sure sight on
the moral truisms, the detailed social analysis, utilizing accounts of political
economy and political sociology, may do all the work. However, unless we
content ourselves with waxing rhetorical about what the design of a truly
human society would look like, it may be the case that we need, together
with an in-depth social analysis, to articulate something of a moral theory. If
that is so, the failure to articulate a defensible version of an objective norma-
tive ethics is not without its human consequences. But what should not just
be assumed is that we need such an account and that without it we will have
no rational basis for social critique or positive moral statement. In a whole
range of cases it very much looks like that, along with the moral truisms,
almost any of the historically influential non-skeptical theories will do. What
we need is a more developed and ideologically untrammeled social theory not
more and better moral theory.

III

I want now to shift the direction of my discussion. There is a tradition of
moral philosophy, culminating in Kant and finding a contemporary expression

[6] The anthology, *The Capitalist System*, Richard C. Edwards, *et al.* (eds.), (Second
Edition), (Prentice-Hall, Inc.: Englewood Cliffs, New Jersey, 1978) brings out the central
considerations here and the bibliography in that volume gives the crucial sources here.
[7] Teitelman, *op. cit.*

in the work of Alan Donagan, which makes a tight link between rationality and morality.[8] Practical reason, properly understood, issues in *categorical prescription*. There are, this tradition gives us to understand, some general moral principles at the foundation of a uniquely rational system of ethics which any, reasonably informed and conceptually un-befuddled rational agent just must accept and act in accordance with. But the search for such principles has been the search for a will-o'-the-wisp. Either, as with Aquinas's first principle of the natural law or Kant's categorical imperative, we get something devoid of content or we get something with content which is not required by reason, where "reason" is conceived in morally neutral terms. Where it is so conceived, it is crucial to realize that rationality under-determines morality. There are different moral codes and indeed even normative ethical theories which are *consistent* with reason and there is no single normative ethical theory or closely related subset of such theories which is *required* by reason. Indeed, Sidgwick to the contrary notwithstanding, there are no tolerably determinate non-vacuous normative principles of sufficient scope to serve as general principles of morality, which are such that it would be self-contradictory to deny them.

No amount of fiddling with prisoner's dilemma situations is going to make it the case that there cannot be consistent individual amoralists, adept free-riders, who can be just as rational as the person of moral principle. The same thing is even more obviously true for class amoralists.[9] Racists, slave-owners, members of the ruling class have no doubt had their share of members who have in various ways suffered from self-deception or ideological bamboozlement, but there have been tough-minded, clear-headed members of such groups, or at least there is nothing conceptually untoward in their being such people, who take a thoroughly class point of view and use moral discourse manipulatively as moral ideology to aid them in their domination of other classes or races.[10] They can do this with a vengeance without its being the case that they must suffer any diminishment in their rationality. It is hardly plausible to believe that all slaveowners suffered from some form of irrationality.

Amoralists aside, rationality underdetermines morality in another way. Different moral theories are consistent with the various principles of rational action and rational belief where those principles are construed, as is the case with instrumental rationality, in a normatively neutral manner. Purely rational

[8] Alan Donagan, *The Theory of Morality*, (The University of Chicago Press: Chicago, Illinois, 1977). See in relation to my discussion of Donagan and such a tight link between rationality and morality, L. W. Sumner's perceptive discussion of Donagan's account. L. W. Sumner, "Critical Notice of *The Theory of Morality*", *Canadian Journal of Philosophy*, Vol. IX, No. 1, (March, 1979), pp. 185–194.

[9] See my "Rationality and the Moral Sentiments", *Philosophica*, Vol. 22, No. 2 (1978), pp. 167–191 and my "Rawls and Classist Amoralism", *Mind*, Vol. LXXXVI, No. 341 (January, 1977).

[10] See my "Morality, Marxism and The Ideological Functions of Morality", *The Occasional Review*, Issue 8/9 (Autumn, 1978), pp. 165–182.

contractors, forced to reason impartially, but without a sense of justice or any moral motivation (a moral psychology), will perhaps not opt for general egoism, but, as much of the discussion around Rawls's theory has brought out, there are a number of different moral options open to such rational contractors. Rawls's contractors are not actually such *purely* rational contractors because they are imbued with a sense of justice. However, if they were such purely rational contractors without such a moral sense, as certain critics of Rawls have portrayed them as being, it still would not be the case that, even with their perfect rationality and general empirical knowledge, that they could deduce, from an original position that they would simply see as rationally manditory, a distinctive moral code or set of principles of morality or (more specifically) principles of justice such that they could determine, without any appeal to considered judgements, how it is that human beings should live and organize their social lives.[11] It is a Kantian illusion, an illusion shared by some who are not Kantians, that there is a morality, if only we can unearth or (perhaps) invent it, which is required by reason. There are many moral points of view that can be equally compatible with or in accordance with the principles of rational action and rational belief.

I have been assuming in my above remarks that some version of instrumental rationality (means-rationality) is being appealed to. Instrumental rationality, in its various versions, remains an anti-septic, morally neutral conception of rationality. In my above remarks about rationality that is the conception I have been assuming. If alterntively a substantive conception of rationality — an ends-rationality — is adopted, after the fashion of Habermas or the Frankfurt School, there is still no deriving morality from principles of rationality, for now the very principles of rationality are themselves being conceived partially in moral terms.[12] It may be a more adequate conception of reason than the one extant in the Anglo-American tradition of analytic philosophy — indeed I am inclined to think that it is — but it is not a reading of rationality which will enable us to determine what our moral principles are to be.[13]

IV

It may be that expecting so much of reason is extravagant. Neither pure practical reason nor pure any other kind of reason is sufficient to enable us to

[11] See here Robert Paul Wolff, *Understanding Rawls*, (Princeton University Press: Princeton, New Jersey, 1977). But see, as well, Brian Barry, "Critical Notice of Wolff: Understanding Rawls", *Canadian Journal of Philosophy*, Vol. VIII No. 4 (December, 1978), pp. 753—783.

[12] See my "Principles of Rationality", *Philosophical Papers*, Vol. III No. 2, (October, 1974), pp. 55—89, my "The Embeddedness of Conceptual Relativism", *Dialogos*, Vol. XI, No. 29—30 (November, 1977) and my "Reason and Sentiment" in *Rationality Today*, Theodore F. Geraets (ed.) (The University of Ottawa Press: Ottawa, Ontario, 1979), pp. 249—279.

[13] See my "Rationality and Sentiment".

determine how we are to live our lives or even what morality requires of us. Why expect reason to be such a Santa Claus? It is one thing to construct a moral theory which is *consistent* with the demands of practical rationality, it is another to give an account which is *required* by them. Both Hobbes and Kant, with their quite different conceptions of rationality, tried to establish that their moralities were *required* by reason. This is an honorable philosophical tradition which has been repeatedly defeated. It, as L. W. Sumner has observed, is perhaps comparable in its extravagance to the patently absurd claim of maintaining that theoretical reason determines the choice of scientific theories. It is absurd to believe that "where several rival scientific theories contend for attention all but one of them are not merely false or uneconomical or inelegant but inconsistent with the principles of logic".[14] As Sumner well puts it "Any deduction of morality from rationality must be either an *ignoratio elenchi* or a *petitio*: the former if the content of practical rationality is kept to a minimum and the latter if it is enhanced in order to yield the result which one wishes".[15] It is a difficult lesson for philosophers to learn, but we need firmly to recognize, if we want to avoid building castles in the air, that "rationality is too weak a foundation for any morality".[16]

It is then our fate to be moral skeptics or in one sense or another to be ethical relativists or subjectivists? Is it the case that there is both in reflective moral practices or ways of living and among moral theories, a plurality of practices and theories which are equally reasonable, equally well-grounded or at least that there is a plurality of ways of life or normative ethical theories none of which can be known to be superior candidates for being required by reason than the others? It is tempting to believe that this is the only thoroughly sane view to take of the matter. If to require that morality be so grounded in reason is not to ask for the color of heat, it is to ask for the moon. Again Sumner has put the matter well:

> All of the major theories among which we may choose are consistent with the minimal demands of practical reason. Each can be readily buttressed with an appropriate view of human and non-human nature. None, therefore can be decisively refuted by any of the others, nor by appeal to some deeper ground.[17]

I think, if we keep firmly before our minds, the phrase "decisively refuted", it would be difficult to demur at that statement. But then again, formal contexts apart, there is precious little that can be decisively refuted. Can we, in some reasonable manner, supplement practical rationality, or practical rationality and theoretical rationality taken in tandem, and, dropping the extremely stringent requirement of *decisive* refutation, sufficiently narrow

[14] Sumner, *op. cit.*, p. 194.
[15] *Ibid.*
[16] *Ibid.*
[17] *Ibid.*

the field so that we will have only one leading contender that all reasonable, properly informed and philosophically sophisticated people will agree is the leading contender? Is it reasonable to expect such a consensus? Is it even reasonable to hope that by determined, reflective and rigorous philosophical work, we can attain such an Archimedian point?

What are we to supplement reason with? We have already seen that it will not do simply to supplement it with our moral truisms. Suppose we add our common morality as reflected in our considered judgements in reflective equilibrium. But we must go very cautiously here. If "our common morality" comes to mean, as it does for Alan Donegan, something he calls, with a touch of Anscombe, "Hebrew-Christian morality", then we have, again as Sumner has pointed out, very poor grounds indeed for not being skeptical. Surely, if a "common morality" is to mean much of anything, it must be a nearly universal morality, reflecting a wide moral consensus. But, if we look at the content of Donagan's Hebrew-Christian morality, we will see how far we are from attaining anything like a moral consensus concerning it. We do not have a moral consensus over whether suicide, abortion or euthanasia are categorically impermissible. We surely do not agree that voluntary sterilization is impermissible because it is a form of self-mutilation or that casual sex must be exploitative or that we have duties to self-improvement. If we look at morality with even a remotely sociological eye, we will immediately recognize that there is nothing like a consensus about such matters. Perhaps this shows, as Anscombe believes, that a not inconsiderable number of us have "corrupt minds" but without some rather firm criteria for detecting such corruption, it is better to conclude that this is fanatical and arrogant arm waving. In trotting out such considered judgements, as our common morality, Donagan is simply being arbitrary. He is foisting off a conservative traditionalist moral point of view as something which is our common morality, in effect arbitrarily persuasively defining "our common morality".

We are much more comfortable with Rawls's examples that disapproval of racial discrimination and religious intolerance are firmly fixed parts of our common morality. They are things concerning which we do have a moral consensus. We do not feel the same disquietude that we do about Donagan's and Anscombe's claimed common morality. There is, however, a confusion on our part here, perhaps even a form of self-deception. As Steven Lukes points out, our feeling comfortable here may be due to no more than the fortuitous fact that most present day readers of such philosophical literature have a roughly similar, more or less liberal, cultural orientation.[18] But we, who constitute the dominant present day Anglo-American intellectual culture, are a rather shakey dominant majority. We are but a tiny minority of the peoples of the world and a minority with its own distinctive culture-specific consensus.

[18] Steven Lukes, *Essays in Social Theory*, (The Macmillan Press Ltd.: London, 1977), pp. 154–174 and 187–190.

How can we reasonably claim — given that the appeal is just to the consensus — to speak for humankind when our consensus is both a historically and culturally specific consensus? Indeed, our problem is worse than that, for even within our own societies at this particular period of time there is no such consensus. There are among us racial separatists, conservative traditionalists, St. Simonian technocrats, libertarians, Marxists and orthodox Jews and Christians. These people are people who would typically demur at one or another of the deeply embedded considered judgements to which Rawls appeals. Rawls's consensus appears to be as much a distinctively *liberal* consensus as Donegan's consensus appears to be a conservative Hebrew-Christian traditionalist consensus. But how can any such culturally specific fortuitous consensus serve as our supplement to reason? Shouldn't such appeals to consensus in considered judgements simply deepen our sense of the relativity of these things and strengthen what I have characterized as a moderate skepticism?

V

Perhaps? But perhaps not? It may also be the case that if we take very seriously the method of reflective equilibrium we can come to see how we can move from an initially culturally skewed appeal to considered judgements, even firmly fixed considered judgements, to something less parochial. The possibility of this is enhanced if we move from a *narrow* reflective equilibrium to the *wider* reflective equilibrium suggested by Rawls but incisively developed by Norman Daniels in an important series of articles.[29]

I should briefly characterize the distinction between *narrow* reflective equilibrium (NRE) and *wide* reflective equilibrium (WRE). It is the latter not the former, Daniels argues, that is of interest to the moral philosopher. NRE "consists of an ordered pair of (a) a set of considered moral judgements accéptable to a given person at a given time, and (b) a set of moral principles that economically systematizes (a)."[20] We start in NRE with our initial moral judgements and pare them down by using a two step procedure. (This procedure is also used as the two initial steps in WRE.) First, we set aside initial moral judgements of which we are not altogether confident. They may have been made in a condition of some agitation or confusion — an agitation or confusion that would distort our judgements — or, while they may have been made without any agitation or confusion being present, they may have been made without adequate factual information about the situation. In attaining a set of considered judgements we first eliminate all moral judgements which

[19] Norman Daniels, "Wide Reflective Equilibrium and Theory Acceptance in Ethics", *The Journal of Philosophy*, Vol. LXXVI, No. 5 (May, 1979), pp. 256–282, "Moral Theory and Plasticity of Persons", *The Monist*, Vol. 62, No. 3 (July 1979), pp. 265–287. "Reflective Equilibrium and Archimedean Points", *Canadian Journal of Philosophy*, Vol. 10, No. 1 (March 1980), pp. 83–104, and "On Some Methods of Ethics and Linguistics", *Philosophical Studies*, Vol. 37 (1980), pp. 21–36.
[20] Daniels, "Reflective Equilibrium and Archimedian Points".

suffer from any of the above mentioned *maladies*. The resulting set of moral judgements are further pruned by eliminating those which do not fit with the desired set of moral principles, though sometimes too, if there are many conflicting or very firmly embedded conflicting considered judgements, we may modify or even abandon a moral principle. This is all the pruning of initial moral judgements that goes on in NRE. Though we should, of course, note that this is a two way street, for moral principles will be rejected which fail to match with our firmest, most deeply embedded considered judgements.

WRE, by contrast, consists not just in an ordered pair (a) and (b) but, as well, in a set of moral theories and a set of background social and psychological theories. As Daniels puts it, WRE "consists of an ordered triple of sets of beliefs: (a) a set of considered moral judgements; (b) a set of moral principles and (c) a set of background theories".[21]

In wide reflective equilibrium, we do not simply settle for those moral principles which give us the best fit with our considered moral judgements and which most economically systematize the considered judgements that survived the first stage of pruning. Instead, as is clear in Rawls' actual practice, and which marks it as something which is a clear advance over pluralistic intuitionism, in wide reflective equilibrium (WRE), we launch philosophical arguments designed to bring out the relative strength and weakness of the various moral principles proposed as principles which match our considered judgements in reflective equilibrium. Daniels construes these arguments "as inferences from some sets of relevant background theories".[22] When one such argument wins, we are, if we utilize WRE, persuaded that one set of abstract moral principles is superior to the alternative sets of principles. Armed with such a set of principles, we very well may have to eliminate or revise some of our initial set of considered judgements. We shuttle back and forth in employing WRE within the ordered triple, sometimes making adjustments to our considered judgements, at others to our preferred set of moral principles and in still others to our preferred background theories, until we arrive at an equilibrium point which consists of an ordered triple of (a), (b) and (c).

This wide reflective equilibrium (WRE) will provide an account of moral theory acceptance which, while being thoroughly fallibilistic, will yield a justificational structure which is parallel to the justificational structure in science and which is in some sense objectivistic.[23] It is important to recognize that there is in such an account a rejection of foundationalism; not even our most firmly embedded moral judgements are regarded as necessary truths or as unquestionable, even in principle unrevisable, moral judgements. This even includes what I have called "moral truisms".

It is also a holistic view of theory acceptance; no "data" are taken to be unrevisable or rigidly foundational. Considered judgement are appealed to at

---

[21] Daniels, "On Some Methods of Ethics and Linguistics".
[22] *Ibid.*
[23] Jane English, "Ethics and Science", *Proceedings of the XVI Congress of Philosophy*, (1979). Kai Nielsen, "Considered Judgments Again", *Human Studies*, Vol. 5, No. 2 (1982).

several distinct points, but none are held to be foundational such that in no circumstances are they thought not, in principle at least, to be up for possible revision. There are, of course, certain considered judgements, at different levels of generality, that are "firmed up" and function as provisional fixed points in justificatory moral reasoning, but they are still held firm only relative to the rest of the equilibrium. For Rawls and for Daniels, there are no considered judgements which serve as fixed data to check all moral theories against and there are no considered judgements, serving as data, which are so impregnable that no moral theories or background social theories could ever force their revision or their abandonment. There is no more in ethics, than there is in science, such privileged data which could disconfirm any theoretical claim and which any account must square with to be viable. Indeed WRE, in a way NRE does not, stresses the importance of theory construction as a basis for evaluating considered judgements. Where we have a considered judgement, or even a cluster of related considered judgements, which plainly conflict with a well worked out normative ethical theory, buttressed, on the one side, both by attractive moral principles and by many and varied considered judgements and, on the other hand, by carefully developed social theories, in turn supported by corroborated empirical data and further considered judgements, at least some of which are independent of the considered moral judgements appealed to in confirming the moral theories, we have good grounds for rejecting the considered judgement, or even a cluster of related considered judgements, which conflict with this normative ethical theory. If, on the other hand, the conflicting intuitions (considered judgements) become rather numerous and varied, then there is good reason, particularly if there is a viable alternative moral theory buttressed by a carefully developed social theory, for either abandoning or radically modifying that moral theory. We have here a scope for reason similar to that in science. Starting at either end — either with specific considered judgements or with social theories (and their associated considered judgements) — we have a way of correcting our account that is not — or so the argument goes — tied to the initial, culturally specific consensus.

So even if we start with considered judgements which are culturally skewed, say distinctively liberal considered judgements, we still have a way of correcting them in terms of the moral theories and social background theories we develop. The latter in particular are plainly amenable to empirical constraints. Nothing is sacrosanct and we have a method for correcting intuitions (considered judgements). We do not need just to accept our considered judgements as a kind of incorrigible given which then determines the whole structure of the theory.

To this it surely could be responded that still at every major juncture or level in the WRE there is an appeal to considered judgements. This is true for the abstract moral principles themselves and for the background social theories. All along the line there is an appeal to considered judgements and what considered judgements will seem plausible is not unaffected by our distinctive

enculturation. We indeed seek to match principles and specific considered judgements, sometimes altering our general principles in accordance with these judgements. Here the culturally specific rather particular considered judgements deeply effect what principles are accepted. Moreover, as Rawls points out in his Presidential address, in implicit response to Peter Singer's confusions about these matters, sometimes the appeal to intuition is at the abstract level of principle, abstract in the way Sidgwick's fundamental intuitions are abstract, but again, in that context as well, culturally conditioned responses play a determinate role.[24] The weight that some people give to desert, the weight that they give to the abstract dictum that people must be rewarded according to their deserts, is, though abstract, an example of a deeply culturally conditioned response.

A similar thing obtains at the level of social theory. The conception of personal identity, to take an example developed by Daniels, chosen in such a context, in developing such a holistic moral-cum-social theory, is deeply affected by considered moral judgements which are again culturally conditioned. What begins to emerge is that while various empirical constraints and general theoretical and systemic considerations affect deeply the structure of the theory, the structure of the theory is also deeply and distinctively affected by culturally specific moral convictions. The circle is a big one and has many diverse intermeshing elements and perhaps is in many ways benign but still there is no way, as another tradition would put it, of breaking out of that hermeneutical circle and testing or assessing the data by judgements which are not already in part generated by judgements which reflect a distinctive cultural milieu which in turn find just such principles *simpatico*.[25] Such a circularity is not characteristic only of morality but obtains for many domains, including the different domains of science. Indeed it may be, in the nature of the case, unavoidable, but still, characteristic or not, if this is the structure of our reasoning here it should, for anyone on the quest for an Archimedean point, hardly generate great confidence with its repeated appeal to culturally specific considered judgements. When we realize that WRE, though a method for developing a theory, is still so deeply culture bound, so deeply reflecting the influence of a particular culture, we can hardly be anything but skeptical that it has given us the Archimedean point that Rawls seeks. It is not clear how moving from NRE to WRE enables us to set aside Lukes' point about Rawls's appeals to our considered judgements being skewed by a determinate cultural perspective, a perspective hardly dictated by reason, logic or a good knowledge of the empirical facts.

[24] John Rawls, "The Independence of Moral Theory", *Proceedings and Addresses of The American Philosophical Association*, Vol. XLVIII (1974–75), pp. 5–22. Peter Singer, "Sidgwick and Reflective Equilibrium", *The Monist*, Vol. 58, No. 3 (July, 1974), pp. 490–517.
[25] Charles Taylor, "Interpretation and the Sciences of Man", *The Review of Metaphysics*, Vol. XXV, No. 1 (September, 1971) and my "Social Science and American Foreign Policy" in *Philosophy, Morality and International Affairs*, Virginia Held, *et al.* (eds.) (New York: Oxford University Press, 1974), pp. 286–319.

VI

Both Rawls and Daniels are not unaware of the pervasiveness and depth of human enculturation, the extent of our distinctive social indoctrination and the persistent problems concerning imposed consciousness. "Social institutions", as Daniels well puts it, "affect not only the manner in which we realize our desires but also what our desires, ambitions, and hopes are. They affect not just the kind of persons we are, but the kind we want to be."[26] Moreover, while it is plausible enough to evoke a background social theory which contains, as a central element, the assertion that human beings have "a highest order interest in how social institutions shape their abilities, talents, desires and selves", what form that interest will take, what they will take as a legitimate institutional shaping of their abilities, talents, desires and selves, will vary considerably over cultural space and time.[27] All, for example, may very well, in circumstances of moderate scarcity, prize autonomy but a) what understanding of it they will have will vary not inconsiderably and b) what weight they will give to it in relation to other values may very well also vary not inconsiderably between contractors of equal intelligence and with an equal general knowledge of man and society. Given what we know about less well-informed and impartial people (people like ourselves), we know that this, as a matter of sociological fact, is how it is and we have been given no good grounds for believing it would be otherwise with rational contractors. To make the subject of justice take cognizance of the way in which the basic structures of societies, including most particularly our own societies, shape our persons in various ways is in effect to acknowledge the problem but not to show us how we can surmount it.

The device of the thick veil of ignorance is designed to exclude factors affected by the basic structure of some particular society. But the very conception of what such 'human beings' are like, what they, as rational contractors, would choose, what schedulings they would give to the various primary social goods, already reflects culturally specific beliefs. Indeed it reflects belief-systems to which there are recognizably human alternatives.

WRE makes perspicuous how much there is a coherence account of moral justification at work in such an account. No considered moral judgements, at any level, are foundational. There are no givens that must just be accepted; there are instead many mutually reinforcing considerations, some of them considered judgements, appealed to at various levels, which together, as part of a whole structure, provide the justification for a moral theory.

Rawls's own contract argument itself should be viewed as a feature of a particularly wide equilibrium. It is not self-evidently acceptable. Rather it appeals to a variety of background theories, a theory of the role of morality in society (including a conception of what would constitute a well-ordered

[26] Daniels, "Reflective Equilibrium and Archimedian Points".
[27] Ibid.

society), an account of procedural justice and a theory of the person. All
these theories are to some degree controversial and involve an appeal to
considered judgements but it is also the case that in Rawls's account they
constitute a coherent structure with mutually reinforcing elements. But what
in effect is Lukes's criticism persistently returns like the repressed. There can be
different WRE. It is neither Rawls's nor Daniels' expectation that there will be
only one. And right here emerge certain of the traditional difficulties with co-
herence accounts. If no considered judgements are foundational, we can have
several mutually reinforcing WRE with nothing to choose from between them
but particular, and sometimes incommensurate, cultural preferences.

Suppose we respond that it is not all that culturally relative, that there is a
deep cross-cultural *de facto* consensus here. This alleged fact can be plausibly
challenged. But putting that aside for a moment, even if there were such
de facto agreement, by using thought-experiments we could readily come up
with different clusters of possible considered judgements, leading to different
equilibria, again as mutually supporting in the systems of their respective
thought-experiments, as the *de facto* patterns of coherence. If actual considered
judgements are not treated as foundational, it leaves the account relying
crucially, without a foundational claim, on what our own convictions just
happen to be. But these convictions now come to seem at least to have a
status not unlike that of preferences. We have, or at least seem to have, no
reason to prefer them to the coherently imagined considered judgements of
the thought-experiment.

More than that, if we take this "thought-experiment move" and do not
challenge the "factuality" of the *de facto* consensus, it can in turn be
responded that we are in effect requiring moral principles to hold in all
possible worlds. By allowing the domain of the moral "to include all conceiv-
able possibilities", we may very well, as Rawls has argued, "be condemning
moral theory to futility from the very start."[28] If human beings, say biologi-
cally, were radically different than they are, then we would have — and
reasonably so — very different norms than we in fact have. To try to cater for
such possibilities in our moral theories is to render them totally useless as
guides for how we might live our lives and construct a truly human society.

However, there may be, some would even say there must be, some middle
ground between such a 'possible worlds approach' and accepting a *de facto*
consensus. The trouble with the latter is that it is held captive to our ideologi-
cally skewed and imposed consciousness. If we must rely so heavily at various
key points in our wide reflective equilibria on actual considered judgements,
which in turn are considered judgements of such a culturally distinctive *de
facto* consensus, with no device for in turn assessing their viability beyond the
confines of these admittedly wide patterns of coherence, the worry naturally
arises that what we are reflecting here is 'sacred culture' to which there could
very well be reasonable alternatives.

[28] Rawls., "The Independence of Moral Theory", p. 11.

Rawls's paradigms of provisional fixed points among our considered judgements are judgements that have, when we view them sociologically, a distinct liberal flavor, yet they certainly seem to us, children that we are of a liberal culture, thoroughly reasonable. With these provisional fixed points, namely the belief that religious intolerance and racial discrimination are impermissable, we have judgements that require, if we were to try to defend them, some appeal, on the one hand, to philosophical theology, and, on the other, to physical anthropology and biological theory. For someone who really believes, as a Roman Catholic must, that there is only the 'one true faith' and that heresy will lead to eternal damnation, religious tolerance has a Janus face. But there are people, strange as it may seem, of sophistication and intellectual probity — some of them are even distinguished analytical philosophers and some others (more surprisingly still) are distinguished social anthropologists — who do believe just that. And, if people do believe that some races are in some non-trivial ways inferior, not because of anything in the facts of enculturation but because of some biological story, then they can reasonably believe, relative to that belief-system, that *some* forms of racial discrimination are not such evident evils. I am not suggesting there is any case to be made for either of these illiberal views. I am instead pointing to the fact that there is not a consensus concerning even such considered judgements and that, in defending them, an appeal needs to be made to matters taking us a long way from moral theories or even from their supporting background social theories. It begins to appear that there must be some considerable agreement, as in reality there isn't, on world-views, e.g., Christian versus humanist, for an Archimedian point to be established by the method of wide reflective equilibrium.

If we turn instead to the *de facto* universal consensus of what I have called moral truisms, we will not have a rich enough base to discriminate between competing normative ethical theories and moralities. People with radically differing moral views and overall perspectives on morality can and do agree in accepting these moral truisms, though they will, as I have noted, give them differing weight, scope and different readings.

VII

The notion of objectivity in ethics is a multiply ambiguous notion. But one thing that is meant is that objectivity in a given domain requires some significant degree of intersubjective agreement in that domain. I have questioned whether we actually have it in a wide enough manner and I have further questioned whether we have good reasons for believing that if we rigorously apply WRE we are likely to get it.

Daniels believes that WRE does give us the requisite intersubjective agreement.[29] We will see, if we apply WRE carefully, that many disagreements turn on disagreements in background theories and this, I agree, makes the

[29] Daniels, "Wide Reflective Equilibrium and Theory Acceptance in Ethics" p. 274.

disputes somewhat more tractable, but only somewhat, for there is, (as we have seen,) still an appeal to considered judgements in these background theories and there is a lack of accord there about these considered judgements. Perhaps WRE could lessen it in a non-arbitrary way but it is not clear that it does so.

Moreover, as Daniels stresses himself, we could gain consensus without gaining truth. It is not the case that every time we have consensus, we have truth or even a *rational* consensus. This is particularly worrisome when we reflect on the ideological roots and the ideological overlay of many moral beliefs. We surely should be beware of saying that such intersubjective agreement constitutes moral truth or even that it is a firm ground for believing that there are moral truths.[30] It is true enough that "divergence among wide reflective equilibria does not imply that there are no such things as objective moral truths". However, if we have no grounds for believing there are such truths, then persistent divergence here gives us a strong additional reason for believing that moral truth is a Holmesless Watson. This is particularly true if we have no tolerably clear conception of what could count as "a moral truth".

However, we should also keep in mind the fact that such convergence if we were to obtain it, would not establish that there are objective moral truths. Such convergence doesn't establish truth and divergence doesn't establish falsity. However, we should also keep in mind that if we drop the idea that what accords with the moral point of view or what is agreed on in wide reflective equilibrium constitutes moral truth, we have no tolerably clear idea of what truth or falsity could come to in ethics. There are no other clear models for truth in this domain.

Wide reflective equilibrium does indeed embody "coherence constraints on theory acceptance or justification" in ethics. That is, I believe, a very useful idea indeed and perhaps we can go a long way in theory acceptance without developing an account of truth or committing ourselves at all about whether truth claims can be coherently made in ethics. Be that as it may, the repeated and ineliminatable use of "our considered judgements" all along the line in such a method renders suspect the very idea that we can attain a sufficient degree of objectivity to give us the Archimedian point that moral philosophers, who have thought of themselves as objectivists, have understandably sought. Still, it may be the case that they have asked for too much. Perhaps WRE gives us the only kind of objectivity that we can reasonably expect in ethics.

VIII

Rationality, as I have argued, underdetermines morality. Even when it is supplemented by an appeal to considered judgements, in or out of a narrow or a wide reflective equilibrium, it still appears at least to be the case both that many different moral codes and normative ethical theories are, or can be

---

[30] *Ibid.*, pp. 275–6.

made, *consistent* with reason and that we cannot select out any one of them
as *required* by reason. I am not, however, giving to understand that we can, in
thinking about how we justify moral claims, replace the appeal to considered
judgements with anything else. There is not, I believe, anything else which is a
more adequate replacement. Neither an appeal to preferences nor to human
interests provides a more adequate "data base". Such appeals either involve a
hidden but unacknowledged appeal to considered moral judgements or they
allow elements which are more subjective than the considered judgements
themselves and appear at least to be morally arbitrary.[31] If it is doubted, as I
believe Daniels does, that there are any persuasive examples of central
considered moral judgements over which there are significant class or cultural
differences which would not be extinguished by a rigorous application of
WRE, and if this doubt is actually justified, then WRE would be a very
powerful instrument indeed. But is this doubt justified? Isn't Lukes justified
in pointing out that Rawls's firmest paradigms of considered judgements
(judgements concerning the wrongness of religious intolerance or racial
discrimination) are not judgements concerning which there is wide inter-
subjective agreement? Has it been shown, independently of what considered
convictions we just happen to have, that, if we use WRE, such convictions
will be extinguished? I do not think that it has. However, it is still reasonable
to *hope* that there would be such agreement if WRE were conscientiously
applied. It would, however, as I have argued, have to be very wide WRE
indeed, for it would have to be wide enough to cover agreement in world-
views.

Still, even if we do have *some* convergence here, and indeed a convergence
which looks like a rational convergence, do we have the base in considered
judgements in WRE, together with a clear application of whatever firm
principles of rational action and rational belief we have at hand, to resolve the
deep, partially morally based disputes between libertarianism, liberalism and
socialism/communism?[32] If we do not, and if there is little prospect of the
method providing a way to achieve such a resolution, it is reasonable to be
skeptical about its power to provide the basis for a sound objective normative
ethic. It is surely right to remark, as Daniels has, that it is both "commonplace,
and true, to note that there is variation and disagreement about considered
moral judgements among persons and cultures" and that it is "also common-
place, and true, to note that there is much uniformity and agreement on
considered moral judgements among persons and cultures".[33] Daniels sugges-
tion is that "an underlying agreement on features of the component background
theories" may account for the extensive agreement we do find.[34] But there is

[31] See my "Testing Ethical Theories: The Right and the Good Again", *Critica*, Vol. XI
No. 32 (August, 1979), pp. 15–26.
[32] In the articles cited in footnote 12, I have said something about what principles of
rational action and rational belief we have at hand.
[33] Daniels, "Wide Reflective Equilibrium and Theory Acceptance in Ethics", p. 272.
[34] *Ibid*.

in reality wide disagreement over philosophical accounts about the nature of persons, human nature, the role of morality in society and the like and it is a further commonplace that social science has not yet found its Newton and that there are wide disagreements about which social theories, if any, are correct, what their scope can be and about what level and what type social theories we can appeal to as something which might be reasonably well founded. We are not even clear about how non-ideological social science can be. If, on the one hand, we stick with the social science quantifiers, we are very unlikely to get anything very helpful here, but, if, on the other, we move to the great synthesizers (Marx, Weber, Pareto, Durkheim, Freud, Dewey and contemporary developments of that genre) we a) get theories whose very bases are problematical, b) theories which are surely not *wertfrei* and c) theories at least some of which may be inextricably ideological.

I do not believe there are any *a priori* barriers to developing such theories along lines which would lead to a wide intersubjective rational consensus. In my optimistic moments, I have hopes that some rationally reconstructed social theory emerging from the Marxist tradition will achieve just that. But surely there can be no claim that such agreement about background social theories has been achieved. And here the proof of the pudding is in the eating. We only should avoid being incapacitated by historicist or positivist myths. We can be hopeful here, but until the work is actually done, if indeed it is ever done, we are hardly in a position to make objectivist claims. Hopes and expectations do not add up to a defendable theory. The recognition of the importance of the method of wide reflective equilibrium is I believe, a start but until moral theory is much more deeply integrated with critical social theory than it is at present, I do not think there is much hope of showing what it would be like to have a sound normative ethical theory which would provide us with an objective Archimedian point for assessing social institutions and for guiding our lives as human beings. Such critical social theorizing – a moral-cum-social theorizing – needs to be done before systematic work in the foundations of ethics or meta-ethics will be likely to have much point. *Perhaps*, after that critical social theorizing is done, the foundational and meta-ethical work will be seen to be unnecessary or at best a pleasant little game for those who like tidying such things up?

THE HASTINGS CENTER
& THE UNIVERSITY OF CALGARY