

GRAHAM ODDIE

MORAL REALISM, MORAL RELATIVISM AND MORAL RULES
(A COMPATIBILITY ARGUMENT)

Relativism and realism are often held to be in conflict. It is not obvious what the exact nature of the conflict is supposed to be, because it is not obvious what the two doctrines amount to. But that there is some kind of conflict between moral realism and moral relativism is evidenced by the fact that some take moral relativism to be the straightforward denial of moral realism. David Brink, for example, baldly states that 'Moral relativism is usually understood as a denial of realism or objectivity about ethics' (Brink 1989, 43). This definitional thesis is wrong. Even so, in the case of morality at least the tension between the two seems clear enough. Why?

The moral realist holds *inter alia* that there are objective truths in the moral domain, and further that these objective truths are not mere human artefacts, but are backed up by special kinds of facts. The moral relativist, on the other hand, claims that judgements of right and wrong, permissibility and obligation, and so on, are in an important sense socially constructed; that they are human artefacts, perhaps of a very sophisticated sort; that they depend on the existence of human institutions, agreements and conventions, tacit or explicit, and vary from one set of such institutions and agreements to another. For example, Richard Boyd writes that 'Moral realism . . . contrasts with views according to which moral principles are largely a reflection of social constructs or conventions' (Boyd 1988, 182). Mark Platts (characterizing relativism): 'Moral judgments are partly the result of conventions which could have been and in some places are, otherwise' (Platts 1988, 286). And finally, Gilbert Harman: 'Moral relativism denies that there are universal basic moral demands and says that different people are subject to different basic moral demands depending on the social customs, practices, conventions, values and principles that they accept' (Harman 1989, 363).

I will show that these apparently incompatible theses could both be true: that is, that realism and relativism are logically compatible in the moral domain. The compatibility argument I offer does not exploit the incommensurability of rival conceptual frameworks. Nor does it turn on our limited epistemic access to, and hence disagreement over, the alleged



moral facts. Rather, the compatibility turns on a particular analysis of moral rules, of their nature, provenance and function. The basic ingredients of the analysis are not novel. In fact I think they have a ring of obviousness to them. However the consequences of the analysis have not always been understood. The conception of moral rules I offer, in addition to securing the compatibility thesis, explains and unifies a range of apparently diverse moral phenomena, including: the centrality of rules in morality, universalizability, the supererogatory, the limits of obligation, and moral reform. Since, other things being equal, a theory which unifies and explains a wide range of phenomena is to be preferred, the compatibility proof is more than that. It is also a strong plausibility argument for this conception of the moral.

1. REALISM

Robust realism about a domain is committed both to a truth doctrine and an irreducibility doctrine. The truth doctrine states that discourse about the domain is truth-bearing, and that there are truth-makers (properties, relations, magnitudes and states of affairs) to undergird the true, or approximately true, theories in that domain. The irreducibility doctrine states that the entities (properties, relations and magnitudes) correctly so described are not reducible to anything more fundamental. In particular, irreducibility entails mind-independence: the properties and relations in question are not reducible to, or constituted by, congeries of mental states.

Doubtless such a simple statement contains ambiguities and omits subtleties which it would be fruitful to explore in depth. The large literature on different kinds of scientific realism testifies to the fruitfulness, or at least to the seductiveness, of that enterprise. But simplicity has some virtues, and this dual doctrine – truth and irreducibility – captures something important at the core of the realist's outlook.

Moral realism is thus committed to irreducible, mind-independent truth-makers for the claims made in at least some moral judgements. This characterization is compatible with there being vastly different kinds of moral realists. Morality is usually thought of as embracing both the axiological realm (*good, evil, better than* and so on) and the deontic realm (*permissible, obligatory, supererogatory, forbidden* and so on). For the purposes of this compatibility proof I will concentrate on one particular kind of realist – one who takes the axiological to be metaphysically basic and irreducible, and regards the deontic as in some way derivative. Call this position *axiological realism*. In order to make the compatibility argument interesting I will try to make this version of realism as *robust* as possible.

According to the axiological realist then, axiological features are basic, and deontic features are derivative.¹ Briefly, deontic features are determined by, or supervene on, axiological features.

Axiologies can be subdivided into those which are *agent-neutral* and those which are *agent-relative*. Utilitarianism is a paradigm of agent-neutral axiology, while egoism is a paradigm agent-relative axiology. If we take the *relativity* in agent-relative value quite literally then this distinction is straightforward. The agent-neutralist takes value to be borne by, say, states of affairs *simpliciter*. If value admits of degrees then agent-neutral value can be represented as a *one-argument function* taking possible states of affairs to degrees of value. Agent-relative value, on the other hand, is a *two-argument function*, one which assigns degrees to *pairs* consisting of a possible state of affairs and an individual. In order to make the compatibility thesis interesting I will assume our realist to be committed to the agent-neutrality of value, or at least to the agent-neutrality of the values which ground obligations and permissions. (Maybe the thoroughgoing agent-relativity of value is compatible with axiological realism, but if so then the compatibility is there from the start.)

The realist about value is essentially distinguished from the anti-realist by commitment to the idea that there is some irreducible fact of the matter which determines the shape of value. Further, our robust realist takes actual value to be both determinate and non-trivial. That it is determinate means that it is representable by a real-valued function. That this is non-trivial means that it is neither everywhere undefined, nor does it assign the same value to all states of affairs.² This may sound a bit queer, but only because realism about value is a bit queer (Sylvan 1988).³

While the *determinateness* of value may seem rather a strong condition to place on realism (even a realist about value could countenance a degree of value fuzziness) it is in keeping with the aim of making the compatibility thesis both robust and interesting. That is, for the purposes of the compatibility argument we take the *strongest* version of realism on offer. Obviously the compatibility of realism and relativism would be less surprising if there were a number of quite distinct and incompatible actual value functions. A realist who was also a pluralist about value would be able to accommodate some quite robust relativist theses.⁴

The axiologist maintains that value is what undergirds obligation. Valuable states of affairs make a claim to be realised. The degree of value is a measure of the weight of that claim. The greater the value, the greater the claim. One interesting, but metaphysically extreme, position (recently espoused by John Leslie (1979)) is that value is causally productive of those valuable states of affairs. This thesis is fraught with difficulties. If we

eschew it, then for value to be realised, value which would not otherwise have been realised, it must either await chance, or the deliberate actions of agents capable of choosing between options which differ in value. According to axiological realism valuable states of affairs make a claim to be realised on those agents who can so realise them. Thus in some way or other value generates obligation. The core of the axiologist's thesis about the deontic is just this: that the deontic features of an option, or series of options, or of total states of affairs, are determined by axiological features.

We have assumed that possible states of affairs possess different degrees of value. Exactly how is the value of an action, or of any state of affairs or event, related to the value of the various more determinate states of affairs which realize it? This is a more general version of the problem of the relation of the value of an act to the value of the possible outcomes of the act. From extremely weak and attractive assumptions it is possible to show that the objective value of a state of affairs (if there is such a thing) is its objective *expected* value, where this expectation is determined by the objective chances and the objective values of the various possible realizations or outcomes (Oddie and Milne 1991). This is what I will assume, but for the purposes of the compatibility proof we need not be detained by arguing directly for it. For the assumption that the objective value of an act or state is its objective expected value (where both probability *and* value functions are objective) is undeniably *consistent* with a robust realism about value.

There are other plausible candidates for the deontically relevant value of an act, but each of them involves an element of subjectivism. The chance and value functions could each be replaced by subjective counterparts – chance by credence, and value by desiredness – to yield three more possibilities. Both credence and desiredness are notoriously agent-relative, so each of these three possibilities would build in an element of relativity. Suppose, for example, we replace the chance function by a credence function, and base deontic categories on the resulting mixture of subjective probability and objective value. Given that beliefs about the natural world vary greatly, and that many of these differences are culturally conditioned, different deontic demands could well apply to different people.⁵ But the compatibility I will argue for does not turn on such epistemic relativity. It is with a completely objective, agent-neutral, axiological realism that I claim moral relativism is compatible.

Given that individual acts possess differing degrees of objective value the axiologist may wish to sort and rank them deontically according to those values. One determination thesis which apparently recommends itself is this: that only the best is morally permissible. Anything less than the

best is morally forbidden. This kind of act-maximizing theory faces a host of familiar objections, all of which involve the problem of accounting for the role which rules and maxims clearly play in the moral life. Indeed, a morality without some system of rules and maxims is scarcely imaginable. At the risk of boring the reader I will rehearse some of the main objections to act-maximization.

Firstly then, there is F. H. Bradley's oft quoted objection to act-utilitarianism, that 'this is to make possible, to justify, and even to encourage an incessant practical casuistry; and that, it need scarcely be added, is the death of morality' (Bradley 1962, 109). Generally, if maximization is true then the appropriate method for selecting an option would involve calculating, over each and every set of options, the objective expected value of each. Such an approach would not only be the death of morality, but the death of all spontaneity, perhaps the death of action. Actual moral reasoning, of course, proceeds not so much in accord with calculations over all possible consequences, but rather in terms of the *kinds* of act that the agent faces. Typically the agent locates the choices he is facing within various general kinds, kinds about which he typically accepts certain normative principles or rules.

Secondly, there is the supererogatory. According to the act-maximizer options that fall short of the best, even by some small amount of value, are morally wrong. But this excludes the supererogatory. A simplistic lowering of the threshold would not only face the charge of arbitrariness, but would fail to accommodate other intuitions about morality. There are areas of value promotion and enhancement which exceed what is morally obligatory, but which are not supererogatory. Suppose I have two bottles of whisky in the cupboard: one a rather ordinary blend, the other a good malt. I reach into the cupboard to pour myself a drink and without much thought I take the blend. The malt would have given me more pleasure *now*, and given the small chance that this may be the last whisky I will ever enjoy together with the fact that none of my heirs appreciate whisky, a swift calculation would have revealed that having the malt now has higher objective expected value than having the blend. It sounds strange to say that in choosing the blend I did something morally wrong. It also sounds odd to say that pouring the blend was a moral peccadillo, something which was morally wrong, but only in a minor and trivial way, a moral wrong not worth worrying about. (Anyone who suggested that should be advised to lighten up a bit.)

Thirdly, there are tragic moral dilemmas, when it is impossible for an agent to do anything without thereby doing something that is morally impermissible. According to the act-maximizer there must always be an

optimal option and that will always be morally permissible. Moral rules might help to explain the existence of such dilemmas, for rules might come into conflict in odd situations.

Fourthly, there is universalizability. Some versions of universalizability the axiologist may accept without qualms. For example, she may be happy to endorse the principle that no two situations differ in value without differing in some non-evaluative feature; or that two agents in exactly similar situations with exactly similar options, are morally obliged to choose exactly similar acts. Of more interest is the substantive Kantian thesis which enjoins us to act in such a way that the maxim under which the act falls is one which we can will as a universal rule. That is, in acting morally it is sometimes pertinent to ask what would happen if everyone were to follow suit. The significance and scope of this idea is hotly debated, but it does seem to play an important role in moral reasoning.⁶ It would be nice to be able to accommodate those applications of the principle where it is indeed attractive, as well as to be able to explain, in those cases where it is not, why it fails.

Considerations like these have moved some to adopt some version of rule consequentialism (or more broadly, rule-axiology) a doctrine which is now enjoying a little renaissance after a long spell in the shadows. Still, it seems that it would be crazy for an axiologist to demand strict rule-following in circumstances in which breaking the rule would, all things considered, promote greater value – ‘rule worship’ as Jack Smart has so aptly described it (see Smart and Williams (1973)). A solution to the above problems would have to be able to cope with that kind of objection to rule-consequentialism.

2. RELATIVISM

Just as there are a variety of realisms there are also a variety of relativisms. But the core of moral relativism is, of course, that deontic judgement – of moral right and wrong, moral permissibility and impermissibility, the wicked and the supererogatory – are *relative*. To *what* they are relative will differ from one account to another. The version I want to consider is a form of cultural relativism often called *conventionalism*: that moral obligations and permissions are relative to cultural or social conventions or agreements. Conventions are socially constructed entities. Conventions can and do differ from one society or group to another. And so if moral obligations depend on, or make sense only in relation to, such conventions then moral relativism in one clear sense is true.

There are two doctrines of relativism possible here, a weak doctrine and a strong doctrine. The first, weaker, view is that moral obligations are generated by social conventions, and that different social conventions generate possibly incompatible moral obligations. This weak doctrine can be consistently combined with the view that moral judgements make sense outside a context of social conventions. The second, stronger view is that moral judgements only make sense in relation to social conventions and agreements. I will consider both views, and their compatibility with realism.

Gilbert Harman apparently defends the stronger of these two versions of conventionalist relativism:

My thesis is that morality arises when a group of people reach an implicit agreement or come to a tacit understanding about their relations with one another. Part of what I mean by this is that moral judgments – or, rather, an important class of them, make sense only in relation to and with reference to one or another such agreement or understanding. . . . It should be clear that I intend to argue for a version of what has been called moral relativism. (Harman 1989 33)

And:

. . . the judgement that it is wrong of someone to do something makes sense only in relation to an agreement or understanding . . . an action may be wrong in relation to one agreement but not in relation another. . . . it makes no sense to ask whether an action is wrong, period, apart from any relation to an agreement. (Harman 1989a, 33ff)⁷

This agreement account of morality is appealing to those who are naturalistically inclined. Further, since the content of the agreements will presumably be various rules and regulations governing certain kinds of behaviour, this kind of theory has the resources to solve the kinds problems raised in the last section against act-maximization. However, conventionalism does face a number of well-known problems of its own, just three of which I will raise here.

Firstly, if agreements generate moral obligations then there must be some connection between such agreements and the obligations they generate. This connection cannot be a purely naturalistic phenomenon, for one term of the relation is moral. It connects a natural phenomenon (agreement) with a moral one (obligation). This connection (that conventions are morally binding) is a moral feature of the universe. We now face a dilemma. Either this connection is a basic feature of the world, or else the connection is reducible to some more basic feature. The former horn of the dilemma undercuts naturalism – the very feature which makes conventionalism appealing to advocates like Harman – while the second horn of the dilemma generates a regress. Moral features have to be grounded in an agreement, but on pain of an infinite regress at least one kind of agreement

must generate moral obligations without there being some prior agreement to undergird that. That casts us back on the first horn of the dilemma.

Secondly, while it seems attractive and natural to regard moral systems as generated by social conventions and agreements, and to see moral obligations as generated by participation in such systems, it is not so attractive to accept what seems to be an obvious corollary: namely that there is nothing external to the social conventions against which they can be judged – except, of course, other social institutions, other conventions, other agreements. But we do want to be able to criticize some agreements, conventions, and whole moralities, as morally inadequate, without that criticism reducing simply to an appeal to another set of conventions in which we happen to be embedded.

Thirdly, agreement theories of morality seem to leave out of their purview a range of morally relevant states and morally relevant beings. For at least those within the Hobbesian social contract tradition tend to give only the social contractors themselves full moral status. Morally significant beings who are unable to enter into the contract (babies, animals, mental defectives, perhaps trees and ecosystems) seem either not to have been brought within the purview of morality at all, or are treated of only insofar as the contractors' interests happen to coincide with theirs.

Finally, conventionalist moral theories raise rather sharply the problem of justifying morality. *Why be moral?* becomes *Why obey the social conventions practised around here* when doing so clashes with prudence? While it may well be mostly prudent to abide by Roman conventions when hanging out in Rome, when prudential push comes to shove, that justification won't wash.

It should be clear from the second and third objections that I am not interested in extreme versions of moral relativism, according to which absolutely any system of conventions goes. We can consider a spectrum of possible positions ranging from moral absolutism (that there is just one true morality), to an extreme relativism (that any moral system at all is on a par with any other). I assume that neither of these extremes is plausible, and that what is attractive is a moderate version of moral relativism: one which holds that there is a possibly rich variety of legitimate moral systems which generate incompatible moral obligations for their respective adherents. A theory which could deliver this rational kernel of moral relativism, while avoiding the supposed pitfalls, would be worth a look.

3. THE SOCIAL CONSTRUCTION OF MORALITY

Let us begin by assuming, with the axiological realist, that some possible situations are good and some are bad *simpliciter*. This does not preclude the possibility that agent-neutral goodness or badness of situations is determined by the welfare and interests of humans, or of sentient creatures generally, or of some broader range of entities that might be considered to have interests. Further, these interests in themselves may be essentially agent-relative. As a general principle there is nothing to stop us assuming that the totality of agent-relative interests collectively settle overall agent-neutral value. That is, agent-neutral value may well globally supervene on agent-relative value: no difference in agent-neutral value without some difference in agent-relative value. One particularly attractive principle is Pareto optimality: that agent-neutral value is raised whenever the agent-relative value for some being is raised and the agent-relative value for no being is diminished.

There are two aspects of the focus of morality which are of interest. Firstly, much of what is of interest in a moral system involves *interactions* between agents. Exchanging information, exchanging goods, making love, handling conflicts, raising a family, caring for the sick and injured, eating – these activities are typical of the interactions on which moral systems like to comment. Typically moralities will lay down obligations and permissions – in brief *norms* – for such situations. For example: don't lie, don't steal, don't rape, take an eye for an eye, honour your father and mother, feed your children, you may leave old and weak grandparents to perish in the snow, you must not kill innocent human beings, you may kill in self-defence, you may let innocent human beings die, you must not eat your parents, you must eat your deadly enemies, and so on. Secondly, what follows immediately from this, but is nevertheless worth noting separately, moralities involve both *kinds* of actions and kinds of *interactions*. This generates two observations:

Firstly, the fact that it is recurring *kinds* of interaction which feature here, interactions in which certain kinds of behaviour produce good outcomes, and certain other kinds produce bad outcomes, enables even the extreme act-axiologist to begin accommodating moral rules, rules subordinate to the overall principle of value promotion on each individual act.

Secondly, it is important for the act-axiologist that we are dealing here with kinds of *interaction*. In an interaction the goodness or badness of an outcome typically depends not only on what an isolated agent does, but also on what the other participants do. The outcome is a product of their

collective action. It is this feature, even more so than that concerning kinds, which necessitates a system of moral rules for the act-axiologist.

Let us consider the simplest possible kind of interaction imaginable: one involving two agents, each of whom has just two options: A and B. Let us suppose that the value of the outcome depends on what both do. So we have four possible outcomes and four possible (agent-neutral) values attached to those.⁸ Many patterns of value distribution are imaginable. Suppose that two of the outcomes are good, say those in which both do the same thing, and the other two are bad. So the scheme looks like this:

		Agent 1	
		A	B
Agent 2	A	good	bad
	B	bad	good
		Recurrent situation kind S	

Even such a simple model as this at least approximates some of the kinds of interactions in which we can find ourselves. The most commonly cited example is that of driving on the roads. Here action A is driving on the left, and B is driving on the right. If we both drive on the left then even if we approach each other from different directions the outcome will be good. The same is true if we both drive on the right. But if one drives on the right and the other on the left – disaster.

What should our agents in a situation of kind S do? The tempting answer for the axiologist is that *jointly* they should produce a good outcome, rather than a bad one. They should coordinate their behaviour, so that either both do A, or both do B. But while such advice is true, it is not directly of assistance to our two agents. The act-maximizer also tells each *individual* agent to perform that option which, in the circumstances, has greatest value, where value is just objective expected value. What implications does that have for such situations? Which action of agent 1 has greatest value here?

We cannot answer that question, for we are simply considering a situation *kind* and have not described it in a sufficiently detailed manner to yield values – that is, objective expected values. There are no expected values here until we also attach objective probabilities to the agents acting in various ways.

However, what we can say is that, in any given occurrence of kind S of interaction, if agent 2 is objectively more likely to do A than B then agent 1's doing A has greater value than his doing B. But if agent 2 is more likely to do B than A, then agent 1's doing B has greater value. The same is true

of the relative merits of agent 2's doing A and B. Thus for each agent, which of his options has greatest value depends on what the other agent is likely to do.

Such kinds of interactions are almost identical to what are known in Game Theory as coordination problems with more than one equilibrium. Coordination problems arise when agent-relative values are not in conflict, but coincide (either partially or totally), and an equilibrium point (respectively: *proper* equilibrium point) is one such that had any one participant acted differently from the way he acted in fact, given the ways the others acted in fact, he would have been no better off (respectively: he would have been worse off).

The theory of conventional solutions to recurrent coordination problems in the case where the values are agent-relative and the probabilities are subjective has been thoroughly worked out.⁹ Two differences between that subjectivist theory and the way I apply that framework here are worth noting.

Firstly, I have been assuming that coordination problems can be generated by the objectivist's agent-neutral value. That is unproblematic since an agent-neutral value distribution is formally tantamount to the coincidence of agent-relative values.

Secondly, for the purposes of meshing with our realist theory of value I assume that the relevant probabilities are objective: the objective dispositions or propensities of agents to act. The standard theory is subjectivist in that the relevant probabilities involved are degrees of belief, not objective propensities. But these differences do not significantly impugn the substantial overlap between these objective coordination problems and their subjective counterparts.

Given that our two agents face recurring interactions of this kind it would clearly be good if they could settle into an easily discernible and regular pattern of behaviour. That is, if they could institute a *convention* which would ensure the coordination of behaviour and the attainment of an equilibrium point.¹⁰ What I will give here is a necessary condition for the concept of a *moral convention*. A moral convention is, like its subjectivist counterpart, a regularity in behaviour by members of a community which constitutes a standing solution to a recurrent coordination problem. The crucial differences are generated by the use of objective agent-neutral value, and objective propensities to action.

DEFINITION. A regularity *R* in the behaviour of members of a population *P* when they are agents in a recurrent interactive situation *S* is a (proper) *moral convention* only if:

- (i) uniform conformity to *R* in situations of kind *S* is normally a (proper) agent-neutral coordination equilibrium;
- (ii) members of *P* conform to *R* in instances of kind *S* with a high degree of objective probability.

The concept of a *moral* convention thus builds in objective success. To count as a possible moral convention a regularity in collective behaviour must actually succeed in ensuring the attainment of an agent-neutral equilibrium point in normal cases. Moral conventions can be contrasted, on the one hand, with social conventions which do not necessarily promote overall objective agent-neutral value (say those which promote merely the interests of the group which institutes them but are detrimental to overall value promotion); and on the other with regularities in collective behaviour commonly called conventions but which promote neither agent-relative nor agent-neutral value.

Why call such a regularity a convention? The term “conventional” is, as Goodman points out, intricately ambiguous.

On the one hand the conventional is the ordinary, the usual, the traditional, the orthodox as against the novel, the deviant, the unexpected, the heterodox. On the other hand, the conventional is the artificial, the invented, the optional as against the natural, the fundamental, the mandatory. (Goodman 1989)

It is primarily in this latter sense that solutions to coordination problems are conventional, although any such convention may in time also become conventional in the former sense. The convention *that agents in recurrent circumstance S do A* is certainly optional, because its rival – *that agents in recurrent circumstance S do B* – is just as good a solution to the coordination problem. Nothing here makes one of these two options more natural, fundamental or mandatory than the other. Either convention would do as a means for collectively directing groups of agents along paths of reliably high value, rather than along paths of uncoordinated low value. There is nothing to recommend one convention over the other, and which one, if either, is instituted to solve this recurrent coordination problem is thus, in a perfectly clear sense, *arbitrary*, although it would not be arbitrary to institute some moral convention or other.

I have said nothing in this brief account about the subjective side of a convention. What undergirds the effective operation and maintenance of a convention are shared beliefs and desires. What beliefs and desires? Often it is the common belief that if all stick to the convention in situations of type *S* a good outcome will be secured and a bad outcome avoided, together with the common expectation that others will indeed stick to the convention. Each person judges it to be more valuable to abide by the convention on the condition that others will do so, and expects others to do so. In other

words, each recognises, clearly and distinctly, that they are all operating an explicit agreement in order to ensure good outcomes in situations which present a coordination problem. However, I have not added this to the basic account of a moral convention, since a moral convention could become operational and actually achieve its desirable effects, of guiding the members of a group along a path of high value, without the individual members actually believing that they are operating a convention. In the case of moral conventions the members of the community will typically consider that they are attempting to abide by the one and only morally correct way to behave in such situations. And they continue, on the whole, to abide by the convention without knowing that it is just one conventional solution amongst others to a recurrent coordination problem. Coming to an awareness of the conventionality of the rule might well undermine its effective operation. Indeed, the belief in absolutes might even be psychologically crucial to the effective operation of morality.

How plausible is it to suppose that moral norms can be generated in this way by coordination problems? This is a heavily theory-laden question. It presupposes the axiological realist's fundamental thesis that there is objective agent-neutral value, but without giving us any of the details of the nature agent-neutral value. Still, the claim does have a certain amount of plausibility given quite modest theses. Suppose, quite plausibly, that agent-neutral value is either total or average agent-relative welfare, or some other such function of overall agent-relative welfare. Then there are a host of coordination problems, or approximate coordination problems which moral norms do seem to address. Two which immediately spring to mind are coordination problems generated by the use of land and material goods, and those generated by sexual relations. To these recurrent coordination problems the norms associated with the institution of private property, and with lifelong monogamy, may be candidate solutions. But it is implausible that they are the only solutions, or that they do not admit of a number of important variations. Further, it is not necessary for the purposes of the compatibility argument that *every* moral norm is a conventional solution to an agent-neutral coordination. It is sufficient that some central and important moral norms are so generated.

4. MORAL OBLIGATIONS AND MORAL CONVENTIONS

The conjecture about the function and nature of moral conventions permits two quite different developments of a theory of moral obligation, and these yield interestingly different conceptions and results. I will call them the narrow view, and the global view, for they arise respectively from narrow

and global versions of the thesis of the supervenience of the deontic on the axiological.

The narrow view of the supervenience of the deontic on the axiological is this: that one's moral obligation is (always and everywhere) to maximize objective, agent-neutral value, on an act-by-act basis. Two individual acts with the same value profile must have the same deontic profile.

This narrow view would facilitate an interesting version of the compatibility argument. Suppose that there are recurrent coordination problems generated by objective agent-neutral value, the only effective and systematic solutions to which are moral conventions. Then it will be true that a moral convention will generate objective moral obligations, and that quite different moral conventions would have generated quite different objective obligations. In other words, the narrow view of obligation yields the compatibility of what I have called weak relativism with robust realism.

To see this, note that each moral convention will be associated with a norm: a norm which spells out how agents in such situations ought to act. The norm in the example at hand will either be *in recurrent situation kind S one ought to do A*; or else *in recurrent situation kind S one ought to do B*. Suppose one of these norms, say the former, is one to which members of a group in fact conform with a high degree of (objective) probability. Then typically when *S* occurs it will be objectively probable that each participant will abide by the norm and do *A*. This by itself makes it *objectively more valuable* for each of the parties to the convention to abide by the associated norm. By the same token, the existence of some completely different convention would have made it objectively more valuable to abide by a completely different norm. Thus if it is one's moral obligation to maximize objective value, different socially constructed conventions will generate quite distinct and incompatible types of objective obligations.

If we descend somewhat from the abstract plane to our homely traffic analogy then the compatibility is quite obvious. In a state in which the traffic convention is that of driving on the left it is objectively more valuable to drive on the left than on the right. In the absence of any existing convention there is no mandatory rule about which side of the road to drive on, and the rules of the road are clearly social constructions, but it is nevertheless *objectively* more valuable to abide by the local convention *whatever it happens to be*.

But this is not the only way or even the most attractive way in which a theory of moral obligation could be developed from these assumptions. Moreover, the narrow view still suffers from some of the faults of realism mentioned in Section 2, precisely because it does not give sufficient weight to the *constitutive* role of norms in a deontic system. On the narrow view

the *moral permissibility* of a particular choice would not be determined by whether or not it conforms to the moral norms, but rather, by whether or not it individually maximizes value. As far as *deontic judgements* go, the norms themselves would just be rules of thumb. Further, it would be difficult to accommodate both the supererogatory, as well as those acts of value enhancement which seem not to involve moral obligations at all.

This suggests a quite different way of defining moral obligation, via the norms themselves and *their* value profiles, rather than via the individual value profiles of acts. Let us suppose that interactions between agents do fall into recurrent kinds which typically (although obviously not *exceptionlessly*) exhibit the features of coordination problems with more than one equilibrium. Then the only way of effectively ensuring that agents reliably produce good outcomes in such situations would be via a system of conventions. A moral system, or morality, could then be regarded as a collection of norms which, if followed with a high degree of probability within a given society, would constitute moral conventions within that society.

According to this view the global axiological features of recurrent coordination problems, together with a set of collective responses to those, determine which norms are in force in a population, and those norms in turn determine the deontic properties of individual acts. On this broader view of the supervenience of the deontic on the axiological, an act would be *morally* permissible or obligatory, or forbidden, only relative to a moral system. That is, this view would be consistent with the *strong* view of moral relativism.

Here is one attractive way of implementing this idea. An act *A* is morally permissible relative to moral system *M* (or *M-permissible*) just in case it is of a kind which does not violate any norm in *M*; it is morally impermissible relative to *M* (*M-impermissible*) just in case it is not *M-permissible*; and it is morally obligatory if it is *M-permissible* and every way of refraining from the act is *M-impermissible*.

We thus have a conception of morality – let us call it *moral conventionalism* – which obeys the axiologist's requirement that the deontic broadly supervene on matters of value, and is thus compatible with a robust realism about value. Ultimately it is actual objective facts about value promotion, together with actual practices in a society, which undergird the deontic facts. Further, moral conventionalism is relativistic in the strong sense. Moral obligations, moral permissions, and so on, are relative to systems of norms, socially constructed systems of moral conventions. These conventions could have been entirely different, and there is a clear sense in which some rival conventions would have been just as good.

While all this seems rather close to the contractarian tradition in ethics one crucial difference in both intent and consequence should here be noted. One way of reconstructing the problem the contractarian is addressing is this: starting with the assumption that value resides only in valuers, and that individuals have differing and conflicting values, the existence of a single commonly accepted morality seems puzzling. It gives the appearance of both objectivity and agent-neutrality since a moral system typically restricts unfettered pursuit of an individual's own interests. Or to put it slightly differently: how is it that a single agent-neutral morality can be legitimately undergirded by a collection of conflicting agent-relative value functions? The contractarian tries to explain how it is that rational individuals pursuing their own agent-relative values would rationally converge on a common agent-neutral system of moral norms. The contractarian thereby explains away the appearance of objectivity which is generated by a common morality.

The question raised and answered here is in fact a mirror image of this. Starting from the realist's assumption that there is a single, objective, agent-neutral value distribution over states of affairs, how is it possible for there to be more than one legitimate system of morality? The moral realist thus has the task of explaining away the apparent subjectivity which is suggested by the existence of conflicting but equally viable and legitimate moral systems. Moral conventionalism achieves just that.

5. MORAL CONVENTIONALISM AS A SOLUTION TO THE PROBLEMS

Having delivered the compatibility argument it is time to make good the promise that some of the standard problems raised for the axiological realist as well those raised for the deontic relativist can be elegantly solved by the position I have called moral conventionalism.

Bradley's objection to act-maximization is immediately defeated. In determining what it is *morally* permissible or obligatory to do one is not required to engage in a massive and detailed consequentialist calculation over all possible futures to determine or estimate the expected value of each option. Rather, on the assumption that one is participating in a genuine system of moral conventions *M*, in any morally pregnant interaction one considers the kinds of options one faces, and whether or not these fall under the moral norms that prevail. This is not to say that some calculation or estimation of relative values will not be important for determining whether, in a given situation, there may not be *reasons* determined by value for acting in one *M*-permissible way rather than another. But these reasons will not necessarily be demands of morality.

The second objection involved the inability of the act-maximizer to countenance supererogatory acts. Suppose that an agent is faced with a range of choices, some of which are and some of which are not *M*-permissible. Provided that the norms are genuine moral conventions and that the situation is not atypical, the *M*-permissible options will all be good ones. But some of the *M*-permissible options may well realize more value than others. If our agent chooses one of those *M*-permissible options with more than minimal value she realizes more than the minimum value demanded by the moral norms. As far as value promotion is concerned she goes beyond what is morally obligatory.

It might be tempting, at first, to call *all* actions which fall into this category *supererogatory*. But this is undermined by the third objection: that some actions of value promotion fall outside the scope of the deontic altogether. It is not required by morality that I have the malt rather than the blend – both are *M*-permissible – and the malt realizes more value than the blend. However, it would not normally be considered supererogatory of me to pour myself the malt. Why is this?

There are two components of a supererogatory act. One is that it promotes more value than what morality requires. The second is that it would not be reasonable to *expect* the person to go beyond morality in that particular way. Some, but only some, of our reasonable expectations about what people will do are governed by prevailing conventions. It is reasonable to expect that people will abide by the moral conventions, since a regularity in behaviour is a convention only if there is a high degree of conformity. Other expectations that we have about behaviour are just part of a stock of common knowledge about human beings and what they are likely to do. For example, we don't expect people, being what they are, to go out of their way to promote value when it is at a considerable cost to themselves. Only those acts which both exceed what morality demands, and which it would not be reasonable to expect people to engage in, count as supererogatory. Returning to the whisky, it is clear that while the choice of malt satisfies the first condition (it creates more value than morality demands) it does not satisfy the second (it would clearly not be unreasonable to expect a whisky drinker to take the malt rather than the blend).

What this highlights is the fact that not all reasons for acting will be moral reasons. And, more interestingly, not all reasons connected with agent-neutral value promotion will be moral reasons for acting. There is thus more to value and reason than is countenanced in morality.

Turning to the fourth problem, it is not difficult to see how tragic moral dilemmas can arise within moral conventionalism. Norms govern recurrent interactions. Two kinds of interaction which are usually distinct may well

coincide on a particular occasion. Further, in such a case the act which satisfies one of the norms could be the kind of act which fails to satisfy the other, making it impossible to satisfy both norms at once. Hence one can be left with no *M*-permissible option. Of course, although all of the options will be *M* impermissible in such a case, at least one will be of optimum value. And so there may be an objective *reason* to choose one option rather than another, but whatever is chosen will still conflict with the demands of morality. Again we see that not all reasons for acting are strictly moral reasons.

The final objection to our version of moral realism involved the scope and limits of universalizability. It is obvious that moral conventions are connected with something analogous to universalizability. A moral convention serves the purpose of directing interacting individuals along paths of high value on each *particular* occasion only if the convention is adhered to by all the parties on that occasion. It thus makes sense for one person to conform to the relevant norms only given the supposition that everybody else also conforms. Those who depart from the moral norms thus have a case to answer. General conformity with the norm, by virtue of its connection with the objective propensity of individuals to conform with the norm, undergirds the value of each and every individual act of conformity with the norm. Thus departures from the norm threaten, firstly, the attainment of value in that particular instance; and secondly, the dispositions which make conformity with the norm valuable. Too many departures will undermine the existence of the convention, and with the disappearance of the convention goes the value of general conformity to the norm.

While moral conventionalism has something in common with rule-consequentialism it is this point which sets its justification apart. The rule consequentialist, like the Kantian subscriber to the categorical imperative, casts about for rules such that if everybody followed them a good outcome would obtain. But if we are dealing with recurrent coordination problems with multiple equilibria there are just too many such rules, rules which are incompatible with one another. *Drive on the right* is one rule which passes the Kantian test: *drive on the left* is another. But it does not follow that, both rules having passed the test, I can thereby please myself as to which rule to obey. What matters is what everybody else *is* doing or *is likely* to do. The mere fact that it would be good if everybody were to follow rule R does not, by itself, justify following rule R. It matters how probable it is that others follow suit. It matters, that is, what the prevailing convention happens to be. This is perhaps the most interesting and important feature of moral norms: *their value is generated by conformity to them*. Absent actual

conformity within a community, adherence to a possible moral convention does not in itself guarantee that such adherence has any value.

Not only does moral conventionalism thus help to explain the scope of one version of universalizability it avoids one criticism often levelled at rule-consequentialism, Smart's 'rule worship' objection. For suppose that you are in an *atypical* situation, one in which violating a moral norm will produce more expected agent-neutral value than conformity, even taking into account the effect this violation will have on the objective strength of the convention, and hence of the value of future conformity. What, in such circumstances, should you do? According to moral conventionalism, it is *morally* impermissible to violate the norm. Still, would it not be *better* to violate the moral convention, no matter how entrenched the convention is? It is obvious that this result *is* precisely that delivered by moral conventionalism. It would be a breach of morality to violate a prevailing moral norm, and so it will always be *morally* impermissible for an agent to do so. But sometimes it will, nevertheless, be more *valuable* to violate the norm than to conform to it. Rule worship does not follow *once we surrender the thesis that the only reasons for acting are moral reasons*. What does follow is that there can be a good (non-moral) reason for an agent to do what, *morally*, he ought not to do.

A legal analogy may help here. We can imagine a traffic code in which it is illegal to drive on the wrong side of the road except when overtaking. But suppose an opposing vehicle is out of control and careering towards you on the wrong side. To avoid a collision you will also have to drive on the wrong side. Since you are not overtaking you thereby break the law. But you nevertheless had a good reason, albeit a non-legal one, to do what legally you ought not to do. Not all reasons for actions are legal reasons.

The attraction of this solution is that, like rule-consequentialism, it maintains a high view of the role of moral rules in generating *moral* obligations. But, unlike rule-consequentialism, it allows that considerations other than moral ones can provide agents with reasons, and good ones at that, for overriding moral obligations.

Why is this escape, or an analogous one, not available also to the standard rule-consequentialist? Could she not say that sometimes it is more valuable to break the rule, and hence that moral rules can be reasonably overridden? She could, but she would thereby lose the chief justification for supplanting act-maximization. The extra ingredient which makes moral conventionalism preferable is the insight that moral conventions are *necessary* for the solution of recurrent coordination problems, and that the high objective probability of general conformity with the convention is what *makes* individual acts of conformity objectively valuable.

Let us now turn to the three criticisms of relativism. It turns out that each of them is answered within our version of moral conventionalism, simply because the theory does not repudiate an underlying robust realism about value.

The first criticism, that an agreement theory of morality cannot explain the connection between agreement and moral obligation, is met by the combination of realism about value together with the thesis of the supervenience of the deontic features on value features. It is the realization of value which generates moral obligations, via the norms which solve recurrent coordination problems. Since my version of moral conventionalism is not motivated by an extreme naturalism, the idea that there are metaphysical connections between natural features and evaluative features, and between evaluative features and deontic features is not itself particularly worrying. Since naturalism is not the point of the theory, its failure to be naturalistic is not an internal worry.

Not any old regularity in collective behaviour constitutes a moral convention – that is, a solution to a recurrent coordination problem. To be a *moral* convention a regularity or agreement must *successfully* direct agents into paths of high objective value. Regularities in group behaviour which fail to achieve this can be criticized as morally inadequate. Thus the second and third criticisms – that relativism does not leave room for a moral critique of social conventions, and that it fails to take into account those morally significant beings that cannot enter the social agreements – are defeated. Not every agreement generates a genuine moral convention. And, to the extent that maltreatment of morally significant beings diminishes overall value, moral conventions must be sensitive to them.

Thus contrary to a popular and widespread criticism, the moral conventionalist can endorse the overthrow of a prevailing morality. But the moral conventionalist may well be justifiably cautious in doing so. Here's why. A prevailing regularity in behaviour might be found wanting and yet it might be very dangerous to overthrow it. Consider the following recurrent coordination problem:

		Agent 1	
		A	B
Agent 2	A	good	bad
	B	bad	excellent
		Recurrent situation	

Suppose that for some reason a population has adopted the following regularity – in these situations *do A* – so that there is a high degree of probability

that an agent will indeed do A on any such occasion. Then unilateral departures from that regularity will generate a worse outcome. Hence the regularity counts as a moral convention. Further, given that the convention is in place, it will be objectively more valuable for individuals to abide by the convention than to violate it.¹¹ And yet there is something unsatisfactory about the convention. There is a better convention available, one which would ensure greater overall value on each such occasion: namely, *do B*. And so even a successful moral convention can be criticized on the grounds, quite independently of any agreement, that it fails to realize as *much* value as some other moral conventions would realize.

While this allows legitimate criticism of conventions it also opens up a dilemma for the moral reformer. The lifelong-monogamy convention may be sub-optimal in an analogous way, and some other convention might achieve higher overall value. But the problem is to achieve a *collective* change to a different convention. Inevitably, it will be difficult to change the ingrained expectations which undergird the convention. The passage to the new convention, when no conventions at all are in force, will be fraught with casualties. And a new moral convention may never become properly established. The moral reformer will thus sometimes face the dilemma of whether or not to run the risks inherent in breaking up an existing but inferior conventional morality in the hope that a new and superior conventional morality can supplant it.

Finally consider the old question: *Why be moral?* Any answer seems unsatisfactory. To appeal to the moral obligatoriness of the moral is clearly circular. But suppose we try to avoid circularity by appealing to some non-moral feature which moral actions possess (for example, promotion of self-interest). If it is the concomitant feature which justifies doing the moral thing then why not go straight for acts with that concomitant feature? In this case the account seems otiose. And if the morally right thing to do can fail to have the concomitant feature then the justification appears to collapse.

Moral conventionalism delivers us from this trilemma. The justification for abiding by moral conventions is simple: *moral conventions promote good by virtue of their being effective solutions to recurrent coordination problems*. Absent the operation of any moral conventions we would regularly be getting ourselves in a terrible mess. And once genuine moral conventions are in place, to unilaterally depart from them will typically undermine value.

This justification is not viciously circular. I am not saying that you ought to obey moral conventions *M* because obeying the moral conventions is the *M*-obligatory thing to do. Promoting good is not in itself *M*-

obligatory. But promoting good is a feature which, by its very nature, any system of genuine moral conventions possesses.

Is the justification otiose? Why not cut to the chase and just say that we have most reason to promote – indeed, to maximize – good? But as we have seen, recurrent coordination problems *demand* a set of conventions. Absent such a set we just won't know how to interact and we will regularly go awry. A system of coordinating rules is essential for the collective promotion of good, and once it is in place it is self-supporting. It is the operation of a moral convention which ensures that obeying that very convention possesses value.

What about vulnerability to contingencies – those atypical cases in which abiding by the conventions fails to promote, or even radically undermines, good? In that case the justification does break down. But it seems it should break down. In such a case you might have most reason to do what is morally forbidden. You are caught in a dilemma, maybe even a tragic one. That in itself does not make the optimal act *morally* permissible or *morally* right, on pain of contradiction. The moral conventions define what is morally right and to flout moral convention is to act immorally.

Moral realism and moral relativism can thus happily coexist, at least in the form of moral conventionalism. More precisely, axiological realism is compatible with widespread deontic relativism. This compatibility entails that the attractions of realism are not necessarily criticisms of relativism, and vice versa. In particular, that hoary chestnut of the first-year student of ethics, the rich diversity of legitimate moral systems, is not, in itself, the basis of a valid argument against robust moral realism.

NOTES

¹ We could consign all other moral theories, all those according to which the deontic concepts cannot be so reduced purely to considerations of value, to the deontologist. See Vallentyne (1987).

² The value function need only be unique up to linear transformations, although a *really* robust realist might assume that there was a determinate origin to the scale of value – that states with positive numerical value are in some absolute sense good, those with negative value, bad. Further, we could allow that the values are multi-dimensional vectors, rather than numbers, to allow for different dimensions of value. (See Oddie and Milne 1991).

³ Sylvan argues cogently that it is the uniqueness of the actual world which is at the core of realism.

⁴ See Sylvan (1988). In any case it appears that the deontic facts to which we have *access* radically underdetermine value. In Oddie and Milne (1991) it is shown that the judgments of absolutely any reasonable moral theory at all can be represented within an agent-neutral axiology.

⁵ This is one way of interpreting Frank Jackson's recent proposal that moral obligation be settled by subjective probability and real value. See Jackson (1986). Jackson shows how the relativity of the credence function can generate relativity in obligations. For an objectivist's reply to Jackson see Oddie and Menzies (1992).

⁶ See the collection of essays in Potter and Timmons (1985).

⁷ See Harman (1989a). See also Harman (1989). However, I have not incorporated the details of Harman's particular version of relativism, which is undergirded by a fundamentally subjectivist conception of moral reasons. In terms of the framework of axiologies introduced, Harman's view seems to be that the kind of value which is pertinent to the deontic status of an act is that which is generated by subjective probability and subjective value.

⁸ The values of these 'outcomes' will in turn be expected values, since these outcomes will all be compatible with many different possible futures.

⁹ See David Lewis's (1986) which builds on the pioneering work of Thomas Schelling (1960).

¹⁰ An equilibrium point (respectively: proper equilibrium point) will be one such that had any agent unilaterally acted differently no more value (respectively: less value) would have been realised.

¹¹ As is easily checked: if the objective probability of each agent conforming is high then each maximizes value by doing A. That such situations are prima facie problematic for act-consequentialism has often been noted. See, for example, Ellis (1981), Vallentyne (1987a).

REFERENCES

- Boyd, R.: 1988, 'How to be a Moral Realist', in G. Sayer-McCord (ed.), *Essays on Moral Realism*, Wadsworth, Belmont CA.
- Bradley, F. H.: 1962, *Ethical Studies*, Oxford University Press, Oxford.
- Brink, D.: 1989, 'Moral Realism Defended', in L. Pojman (ed.), *Ethical Theory: Classical and Contemporary Readings*, Wadsworth, Belmont CA.
- Ellis, B.: 1981, 'Retrospective and Prospective Utilitarianism', *Noûs* **15**, 329–335.
- Goodman, N.: 1989, 'Just the Facts, Ma'am', in M. Krausz (ed.), *Relativism: Interpretation and Confrontation*, University of Notre Dame Press, Notre Dame, IN, pp. 383–386.
- Harman, G.: 1989, 'Is There a Single True Morality', in M. Krausz (ed.), *Relativism: Interpretation and Confrontation*, University of Notre Dame Press, Notre Dame, IN, pp. 383–386.
- Harman, G.: 1989a, 'Moral Relativism Defended', in Louis P. Pojman (ed.), *Ethical Theory: Classical and Contemporary Readings*, Wadsworth, Belmont, CA, pp. 33–41.
- Jackson, F.: 1986, 'A Probabilistic Approach to Moral Responsibility', R. Barcan Marcus (ed.), *Proceedings of the 7th International Congress of Logic, Methodology and Philosophy of Science*, North Holland, Amsterdam.
- Jackson, F.: 1991, 'Decision-Theoretic Consequentialism and the Nearest and Dearest Objection', *Ethics* **101**, 461–82.
- Leslie, J.: 1979, *Value and Existence*, Rowan and Littlefield, Totowa, NJ.
- Lewis, D.: 1986, *Convention: A Philosophical Study*, 2nd ed., Basil Blackwell, Oxford.
- Oddie, G. and P. Menzies: 1992, 'An Objectivists Guide to Subjective Values', *Ethics* **102**, 327–50.

- Oddie, G. and P. Milne: 1991, 'Act and Value: Expectations and Representability of Moral Theories', *Theoria* **48**, 165–82.
- Potter, N. and M. Timmons: 1985, *Morality and Universality*, D. Reidel Publishing, Dordrecht.
- Schelling, T.: 1960, *The Strategy of Conflict*, Harvard University Press, Cambridge.
- Smart, J. J. C. and B. Williams: 1973, *Utilitarianism: For and Against*, Cambridge University Press, Cambridge.
- Sylvan, R.: 1988, 'Radical Pluralism – An Alternative to Realism', in R. Nola (ed.), *Relativism and Realism in Science*, Kluwer Academic Publishers, Dordrecht, pp. 253–92.
- Vallentyne, P.: 1987, 'The Teleological/Deontological Distinction', *The Journal of Value Inquiry* **21**, 21–32.
- Vallentyne, P.: 1987a, 'Utilitarianism and the Outcomes of Actions', *Pacific Philosophical Quarterly* **68**, 57–70.

University of Colorado
Department of Philosophy
Campus Box 232
Boulder, CO 80309-0232
U.S.A.