

NON-CONSEQUENTIALISM AND UNIVERSALIZABILITY

BY PHILIP PETTIT

If non-consequentialists are to embrace the requirement of universalizability, then they will have to adopt a surprisingly relativistic stance. Not only will they say, in familiar vein, that the premises adduced in moral argument may be only agent-relative in force, that is, may involve the use of an indexical – as in the consideration that this or that option would advance *my* commitments, discharge *my* duty, or benefit *my* children – and may provide reasons only for the indexically relevant agent, in this case me. They will also have to construe the consideration adduced in typical moral conclusions to the effect that this or that option is right or ought to be chosen, or whatever, as itself only agent-relative in force. So at any rate I argue.

To expand a little on the charge: saying that an option is right, like saying that an option is prudent, will refer to one property when my doings are in question, to another when yours are in question; and the rightness of the option will provide a reason to prefer that the option be chosen only for the person indicated. I shall have a reason for preferring that those choices be made that are right relative to me, as I may have reason for preferring that those choices be made that are prudent relative to me. But while you can agree that the choices are right or prudent under that interpretation – that is, relative to me – you will not be given any reason by either consideration for yourself preferring that they be made. The consideration of rightness, like the consideration of prudence, will have reason-giving force only for the agent relative to whom it is interpreted.

This paper tries to establish that non-consequentialists can embrace universalizability only at the cost of casting the predicate ‘right’ and its cognates as agent-relative in this sense, or only at the cost of making an effectively equivalent move. The cost, in an alternative but potentially

misleading vocabulary, is that non-consequentialists will have to live not just with the agent-relativity of moral reasons, but also with the evaluator-relativity of moral judgements.¹ I do not say that that cost is absolutely prohibitive; and I do not suggest that my argument for it is entirely original. But I do think that the cost is substantial, inescapable and insufficiently appreciated.

My paper is in five sections. First, I give a brief account of consequentialism and non-consequentialism.² Next, I characterize and argue for the requirement of universalizability. In the third section I describe a problem raised for non-consequentialism by the requirement; and in the fourth I outline the solution to the problem which I think non-consequentialists should endorse, emphasizing the relativism that it implies. Then, in the final section, I show that two other solutions to which non-consequentialists might look have essentially the same relativistic effect.

I. CONSEQUENTIALISM AND NON-CONSEQUENTIALISM

Non-consequentialism comes in many forms. Kantians say that agents ought to act on the categorical imperative, or ought to treat other people always as ends. Virtue ethicists say that they ought to manifest certain virtues in their behaviour. Deontologists hold that they ought to discharge certain duties: they ought to tell the truth, keep their promises, be non-violent, etc. Rights theorists maintain that they ought to respect certain rights that others have against them. Contractualists assert that they ought to conform to principles that no one could reasonably object to as the bases of social life. Theorists of special obligation say that they ought to deal in a certain way with those who are bound to them. And so on.³

¹ I do not like the language of evaluator-relativity in this context, for the following reason. The relativization of 'right' discussed in the paper relativizes the truth-condition for rightness claims, and their reason-giving force, but it does not relativize their truth-value. Thus, while the truth-condition of 'That is right' in my mouth may be different from its truth-condition in yours, and while it may provide only me with a reason for preference, there is still a non-relative fact of the matter as to whether the sentence in my mouth is true, and that is a fact on which you and I can agree. Talk of evaluator-relativity may suggest a relativity of truth-value – something of dubious coherence – and not just a relativity of truth-condition.

² The account given here of consequentialism and non-consequentialism – strictly, consequentialism and non-consequentialism about acts – draws on earlier work, most recently on my contribution to M. Baron, P. Pettit and M. Slote, *Three Methods of Ethics: a Debate* (Oxford: Blackwell, 1997). See also P. Pettit and M. Smith, 'Global Consequentialism', in B. Hooker *et al.* (eds), *Morality, Rules and Consequences: a Critical Reader* (Edinburgh UP, 2000).

³ I focus here only on the prescriptions supported by non-consequentialists. I say nothing on the fact that because they do not cover all aspects of action and life, the prescriptions supported often entail corresponding permissions.

What is the common thread in these positions? All non-consequentialists speak, at whatever level of abstraction, about what any or every agent ought to do or be; in that sense they are universalists. All non-consequentialists prescribe the same neutral patterns of behaviour or psychology or relationships: that people ought to act on the categorical imperative, manifest certain virtues, respect the rights of their dependants, nurture their friendships, and so on; the patterns are neutral in the sense that they can be understood in the same way by everyone. And all non-consequentialists say that the right thing for agents to do is to instantiate those patterns, so far as these are co-instantiatable, in their own behaviour or psychology or relationships. In particular, they say that that is the thing to do even if instantiating a pattern in the agent's own life means, because of the perversity of his circumstances, that the pattern will be less fully realized than otherwise in the world as a whole. I shall have an agent-relative reason to instantiate the pattern in my life, so it is said, and you will in yours, but neither of us need have an agent-neutral reason to want the pattern instantiated by agents generally.⁴

Consequentialism, as I understand the doctrine, makes a dual assertion. First, it asserts that there are certain potentially shared values by means of which possible states of affairs can be ranked, though perhaps not completely; these may be neutral patterns in human behaviour or psychology or relationships, of the kind that the non-consequentialist prescribes, or they may be neutral outcome-patterns that are more detached from agency, patterns such as the maximization of human happiness or the saving of uninhabited wildernesses.⁵ And secondly, consequentialism maintains that the right choice for an agent to take in any decision is one of those choices, assuming there is at least one and assuming incomplete ranking is not a problem, that promote the overall realization of such values or patterns: we need not rule on what it is to promote values, though the obvious gloss would equate it with maximizing expected value.

⁴ Three points: first, the representation I offer here is not the only one possible, but nothing in my argument is affected by that fact, for the alternative representations that I envisage are only notationally distinct: see my contribution to *Three Methods of Ethics*. Secondly, where I speak here of people's instantiating a pattern in their behaviour or relationships, I have spoken in earlier work of honouring the pattern: see my 'Consequentialism', in P. Singer (ed.), *A Companion to Ethics* (Oxford: Blackwell, 1991), pp. 230–40. And thirdly, I consider here only the sort of non-consequentialism that requires an agent-over-time to instantiate certain patterns; a stricter version would require an agent-at-a-time always to instantiate them, even if that meant that the pattern was less well realized over the agent's life as a whole.

⁵ While I assume that all patterns of the kind that non-consequentialists think should be instantiated in an agent's life can be taken as neutral values to be promoted, I am not now inclined to think that all patterns of the kind that consequentialists think should be promoted can be taken as patterns to be instantiated in an agent's life. See D. McNaughton and P. Rawling, 'Honoring and Promoting Values', *Ethics*, 102 (1992), pp. 835–43, discussed below.

Consequentialism in this sense is a very open-ended doctrine. It says nothing on what the relevant values are. It need not amount, for example, to utilitarianism; it may even espouse as values many of the patterns recommended by non-consequentialists. And consequentialism in this sense leaves open the question as to what day-to-day habits of decision-making ought to be cultivated by agents.⁶ It does not entail, for example, that the agent ought to calculate every choice by reference to the relevant values. Such calculation might do much less well for the promotion of relevant values than a policy of following certain decision-making routines, less well than a policy, say, of in general taking one's cue from the categorical imperative, or from what the manifestation of virtue requires, or from the apparent rights that others can assert.

But however various are the forms of non-consequentialism, and however open-ended is this characterization of consequentialism, the difference between the two approaches is salient. Where non-consequentialism holds that certain patterns are to be instantiated in the agent's behaviour and relationships, even if this means that their overall realization is not thereby promoted, consequentialism reverses that order. It says that certain patterns are to be promoted, even if this means that agents fail to instantiate corresponding patterns in their own lives.

The difference between the two approaches is particularly salient for simple cases where only one pattern or value is taken to be relevant, and there is no issue about the relative importance of different values. To consider those cases, then, where only the value of non-violence is assumed to be relevant, the non-consequentialist pacifist says that non-violence is to be instantiated by agents, even if this means that there is less non-violence overall. The consequentialist counterpart says that non-violence is to be promoted overall, even if this means instantiating violence – say, waging the war to end all wars. In practice the two sorts of pacifist will often converge in their recommendations, as non-consequentialist Quakers converged with Bertrand Russell in their opposition to World War I. But there is always a possibility of their coming apart, as of course the Quakers and Russell came apart in their attitudes to World War II.

II. THE REQUIREMENT OF UNIVERSALIZABILITY

Every prescription as to what an agent ought to do should be capable of being universalized, so that it applies not just to that particular agent, and

⁶ See P. Pettit and G. Brennan, 'Restrictive Consequentialism', *Australasian Journal of Philosophy*, 64 (1986), pp. 438–55; P. Railton, 'Alienation, Consequentialism, and the Demands of Morality', *Philosophy and Public Affairs*, 13 (1984), pp. 134–71.

not just to that particular place or time or context, or whatever.⁷ So at any rate we generally assume in our moral reasoning. If we think that it is right for one agent in one circumstance to act in a certain way, but wrong for another, then we commit ourselves to there being some further descriptive difference between the two cases, in particular a difference of a non-particular or universal kind.

Thus if we say that an agent *A* ought to choose option *O* in circumstances *C* – these may include the character of the agent, the behaviour of others, the sorts of consequences on offer, and the like – then we assume that something similar would hold for any similarly placed agent. We do not think that the particular identity of agent *A* is relevant to what *A* ought to do, any more than we think that the particular location or date is relevant to that issue. In making an assumption about what holds for any agent in *C*-type circumstances, of course, we may not be committing ourselves to anything of very general import. It may be, for all the universalizability constraint requires, that *C*-type circumstances are highly specific, so specific, indeed, that no other agent is ever likely to confront them.

There is no difficulty in seeing how the universalizability challenge is supposed to be met under consequentialist doctrine. Suppose that I accept consequentialist doctrine and believe of an agent *A* that in *A*'s particular circumstances *C*, *A* ought to choose an option *O*. For simplicity, suppose that I am myself that agent and that as a believer in consequentialism I think of myself that I ought to do *O* in *C*. If that option really is right by my consequentialist lights, then that will be because of the neutral values that it promotes. But if those neutral values make *O* the right option for me in those circumstances, so they will make it the right option for any other agent in such circumstances. Thus I can readily square the prescription to which my belief in consequentialism leads with my belief in universalizability. I can happily universalize my self-prescription to a prescription for any arbitrary agent in similar circumstances.

In passing, a comment on the form of the prescription that the universalizability challenge will force me to endorse. I need not think that it is right that in the relevant circumstances every agent do *O*; that suggests a commitment to a collective pattern of behaviour. I shall only be forced to think, in a person-by-person or distributive way, that for every agent it is right that in those circumstances he do *O*. Let doing *O* in *C* amount to swimming to the help of a child in trouble in the water. Universalizability would not force me to think that it is right that everyone swim to the help of a child in such a situation; there might be many people around, and, were they all to swim,

⁷ See W. Rabinowicz, *Universalizability: a Study in Morals and Metaphysics* (Dordrecht: Reidel, 1979).

then they would frustrate one another's efforts. It only requires me to think, as we colloquially put it, that it is right that anyone swim to the help of the child: no one is exempt from this person-by-person non-collective prescription (even if all do face a collective requirement to decide who in particular is going to do the swimming).⁸

So much for the straightforward way in which consequentialism can make room for universalizability. But how is the universalizability challenge supposed to be met under non-consequentialist theories? According to non-consequentialist theory, the right choice for any agent is to instantiate a certain pattern *P*: this may be the pattern of conforming to the categorical imperative, manifesting virtue, respecting rights, honouring special obligations, or whatever. Suppose that I accept such a theory and that it leads me to say of an agent – again, let us suppose, myself – that I ought to choose *O* in these circumstances *C*, or that *O* is the right choice for me in these circumstances. Can I straightforwardly say, as I could under consequentialist doctrine, that just for the reasons that *O* is the right choice for me – in this case, that it involves instantiating pattern *P* – so it will be the right choice for any agent in *C*-type circumstances? I shall argue that there are difficulties in the path of such a straightforward response and that these raise a problem for non-consequentialism.

III. A PROBLEM FOR NON-CONSEQUENTIALIST UNIVERSALIZATION

Suppose I do say, in the straightforward way, that pattern *P* requires not just that I do *O* in *C*, but also, for any agent whatsoever, that that agent should do *O* in *C* as well. Suppose I say, in effect, that it is right for me to do *O* in *C* only if it would be right for any agent *X* to do *O* in *C*. Whatever makes it right that I do *O* in *C* makes it right, so the response goes, that any agent do *O* in *C*. This response, so I now want to argue, is going to lead me, as a non-consequentialist thinker, into trouble.

⁸ The difference is one of scope. I need not say for every *X*, that *X* ought to do *O* in *C*-type circumstances. Rather, for every *X*, I say that *X* ought to do *O* in *C*-type circumstances. This helps to explain why Harean universalization is different from Kantian generalization: see R.M. Hare, *Moral Thinking: its Levels, Method and Point* (Oxford UP, 1981). However, if *C*-circumstances are always guaranteed to be specified in a comprehensive way, as they rarely, in practice, will be specified, then both universal prescriptions can be maintained without the suggestion that a collective pattern is prescribed. We could happily defend the universalization that gives 'it is right' wide scope, for example, though ordinary usage might be strained in our doing so, if the specification stipulated not just that there is a child in trouble in the water, but that there is only one potential life-saver around. I am grateful for comments from Earl Conee on this point.

Judging that an action is right involves approving of the deed and gives one a normative reason to prefer it. Imagine someone who said that he thought his doing something or other, or indeed another person's doing something or other, was the right choice and who thereby communicated that he approved of it. Would it not raise a question as to whether he knew what he was saying if he went on to add that he did not think that there was any good reason for him to prefer that the action should take place rather than not? If the judgement of rightness is to play its distinctive role in adjudicating or ranking actions – if it is to connect with approval in the standard way – then, whether or not it actually motivates the person judging, it must be taken to provide him with a normative reason to prefer that the action should take place.

When I think that it is right that I do *O* in *C*, therefore, I commit myself to there being a normative reason for me to prefer that I do *O*. And when I assert that it is right that anyone should do *O* in *C*-type circumstances, I commit myself – again because of the reason-giving force of the notion of rightness – to there being a normative reason for holding a broader preference. I commit myself to there being a normative reason for me to prefer, with any agent whatsoever, that in *C*-type circumstances that agent do *O*.

The problem with these reasons and these commitments, however, is that they may come apart. For it is often going to be possible that, perversely, the best way for me to satisfy the preference that, for any arbitrary agent *X*, that agent do *O* in *C*-type circumstances, is to choose non-*O* myself in those circumstances.⁹ Choosing non-*O* myself means that there is one person – me – in respect of whom the general preference is not satisfied, but in the perverse circumstances it will mean that there are more agents or actions in respect of whom it is satisfied than there would be did I choose *O*. Perverse circumstances of this kind are not just abstract possibilities, for what an agent does can easily affect the incentives or opportunities of others in a way that generates perversity. The best way to get people to renounce violence may be to take it up oneself and threaten resistance to their violence; the best way to get people to help their children may be to proselytize and not pay due attention to one's own. More generally, the best way to promote the instantiation of pattern *P*, where this is the basic pattern to which one swears non-consequentialist allegiance, may be to flout that pattern oneself.

How can I avoid the conclusion that in such a perverse situation I ought to promote the overall instantiation of my cherished pattern, even at the cost

⁹ Why do I say that this holds often, not always? Because, as effectively noted in the previous footnote, some options may be characterized in a way that does not leave open the perverse possibility. I am thinking of an option such as that of swimming to save a drowning child when no one else is doing so.

of not instantiating it myself? How, in other words, am I to keep faith with the non-consequentialist commitment to the rightness of instantiating *P*, even where this means that the overall realization of the pattern falls short of what it might have been?

It is hardly going to be plausible for me to say that normative reasons bearing on preferences over my own choices trump normative reasons bearing on preferences over how people in general behave. Both sorts of reasons are supported in the common language of what is the right choice or of what ought to be done.¹⁰ And it would surely run against the spirit of universalizability, the spirit in which I deny that my own particular identity is important to the prescription defended, to say that a reasoned preference as to what I do myself should not be responsive to a similarly reasoned preference as to what people in general do – what arbitrary agent *X* does – in the sorts of circumstances in question. This consideration becomes particularly telling when we remember that the satisfaction of the self-related preference may mean that just one person performs the approved action and its non-satisfaction may mean that very large numbers do so. How to justify satisfying the self-related preference, short of privileging the self and rejecting the spirit of universalizability?

Of course acting so as to satisfy a preference for how others behave often involves being a nosy parker and offending against independent values that bear on interpersonal relations, the values associated with being a respectful fellow-citizen or colleague or friend. In these cases the relevance of those values would certainly make a case – it may be cast in consequentialist or non-consequentialist terms – against acting, or at least against acting in certain ways, on a preference for how others behave. But the argument here abstracts from such considerations. It raises the question of how non-consequentialist agents should respond to the recognition, first, that there is a good reason to prefer both that they do *O* in *C* and that arbitrary agent *X* do *O* in *C*; secondly, that these preferences cannot be simultaneously satisfied; and thirdly, as we can now add, that there are no other relevant values

¹⁰ A non-consequentialist might think of arguing at this point that the notion of rightness has reason-giving force for someone who makes a judgement of rightness only where the judgement bears on his own action. But it is hard to make good sense of this move, short of the indexicalization, or other equivalent moves, introduced below. If self-prescription gives me reason to prefer that which I prescribe for myself, so prescription for agents in general would seem to give me reason to prefer that which I prescribe for such agents. Non-consequentialists who are attracted by the move envisaged here need a theory of responsibility under which a person is given a reason for action by a reason for preferring that *X* do something only when he is *X*. But it is hard to see how to design a plausible theory of responsibility – itself non-consequentialist in character – that would do the job. And in any case such a theory of responsibility would have a relativizing effect similar to that which attends the indexicalizing–decentring strategy considered later.

that would argue against satisfying either the self-directed preference or the more general preference. That question still remains in place, and the argument still suggests that the non-consequentialist may be forced by the spirit of universalizability to give it a distinctively consequentialist answer.

The upshot is that if as a non-consequentialist theorist I straightforwardly universalize the prescription that in a certain situation I should instantiate a favoured pattern *P*, then the prescription to which I thereby commit myself, that in that situation any *X* ought to instantiate pattern *P*, may force me to revise my original self-prescription. I have equal reason to prefer both that I instantiate *P* and that any agent instantiate *P* – this reason is expressed by the use of the word ‘right’ or ‘ought’ in each case – and the spirit of universalizability blocks me from treating myself as in any way special. Thus if the preferences are inconsistent in a certain situation – if the choice is between my instantiating *P* alone, for example, or my acting so that many others instantiate *P* instead – then I shall have reason not to instantiate *P* myself.

As a would-be non-consequentialist thinker, my initial claim must have been that the point is to instantiate *P* in my own life, not promote it generally. But I countenance the general claims of the *P*-pattern when I universalize in the straightforward way: I prescribe general conformity to that pattern, not just conformity in my own case. Thus it now seems that what I must think is that this general conformity is to be promoted, even if that means not myself instantiating the pattern in my own behaviour or psychology or relationships. It seems that what I must embrace, in effect, is a consequentialism in which conformity to pattern *P* is the ultimate value to be promoted.

The sort of consequentialism envisaged here is not an outlandish doctrine to will on some self-described non-consequentialists. It is a consequentialism in which the goal is maximal conformity with a reflexive pattern, with a pattern whereby, for any *X*, *X* instantiates a certain pattern in *X*'s behaviour, or psychology, or relationships. *X* does not treat other people only as means; *X* displays a certain pattern of virtue; *X* keeps *X*'s promises; *X* respects the rights of those with whom *X* comes into contact; or whatever. It may well be that many self-described non-consequentialists are really reflexive consequentialists in this mould. They may be reflexive consequentialists who think that it is extremely unlikely that promoting the reflexive pattern they celebrate will ever require flouting it and who have little incentive, therefore, to emphasize the consequentialist aspect of their commitment.

One example of a reflexive consequentialist is the person described by Robert Nozick as a utilitarian of rights, someone who thinks that an agent

like the state ought to act so that there is more rather than less rights-satisfaction, even if that means flouting some rights itself.¹¹ And another is the just-deserts theorist who thinks that the courts ought to pass occasional exemplary sentences in order to ensure that a regime of just deserts is promoted, that is, in order to minimize the number of cases where offenders go unpunished.¹² Such thinkers often describe themselves as non-consequentialists, but clearly what they really espouse is a consequentialism in which the overall aim is to promote a certain regime of rights or deserts.¹³

IV. A WAY OUT OF THE PROBLEM

But many non-consequentialists will shrink from embracing reflexive consequentialism. So is there any other recourse available? Is there any way in which a non-consequentialist can consistently endorse the universalizability constraint?

The reason why the straightforward response to the universalizability challenge raises a problem for non-consequentialism is this. It would force me to think that just as it is right that I instantiate *P* in a certain situation, so in the very same sense of 'right' it is right that anyone instantiate *P* in that sort of situation. It would force me to employ the same currency of evaluation – the same language of 'right' or 'ought', or whatever – in expressing the value I attach to my instantiating *P* and to anyone's instantiating *P*. The only escape from the problem, then, will be for the non-consequentialist to argue that this is mistaken, and that in universalizing the particular judgement as to what it is right for me to do – in moving to a judgement as to what it is right for anyone to do – I introduce an evaluative vocabulary that does not commit me in the same way.

As it happens, there is a simple strategy by which non-consequentialists can effect this escape. In the straightforward mode of universalization, 'It is right for me to do *O* in *C*' goes over to 'It is right for arbitrary *X* to do *O* in *C*', where 'right' in each case is used to express the same sort of prescription. But non-consequentialists can escape the problem raised by arguing that 'right', 'ought', or whatever, does not express the same sort of prescription in the two judgements.

¹¹ R. Nozick, *Anarchy, State, and Utopia* (Oxford: Blackwell, 1974), p. 28.

¹² See J. Braithwaite and P. Pettit, *Not Just Deserts: a Republican Theory of Criminal Justice* (Oxford UP, 1990).

¹³ Another example of a reflexive consequentialism would be the consequentialism that David Gauthier derives from Kantian principles in *Kantian Consequentialism* (Oxford UP, 1996), a consequentialism under which the maximand is the treatment of rational beings as ends in themselves and not as means only.

There are two steps whereby non-consequentialists can achieve this result. First, they can argue that when a person A says that something is right, or at least right for himself, then ‘right’ is indexical and is best represented as ‘right _{A} ’; this indexicalized predicate serves to identify the property of meeting A ’s standards, as the corresponding predicate ‘right _{B} ’ – ‘right’ as used in B ’s mouth – would serve to identify the property of meeting B ’s standards.¹⁴ And then, secondly, they can say that the way to universalize A ’s judgement ‘It is right _{A} that I, A , do O in C ’ is not

(i) $(\forall X)(\text{It is right}_{A} \text{ that } X \text{ do } O \text{ in } C)$

but rather

(ii) $(\forall X)(\text{It is right}_{X} \text{ that } X \text{ do } O \text{ in } C).$

Whereas indexicalizing the rightness predicate leaves open both of these options, endorsing the second means taking the predicate to be not only indexical but also, as we can say, decentred.

If non-consequentialists take this line, then they can escape the pressure towards reflexive consequentialism. For by taking ‘right’, ‘ought’, ‘should’ and other such words to be indexical they can ensure that my commitment to what I do myself has a hold on me that my commitment to what any other agent does in similar circumstances need not have. Universalizing no longer has to mean countenancing a universal prescription implicit in my prescription for what I should do myself. It need only mean recognizing that as I prescribe for myself in my own position, so others must be allowed to prescribe for themselves in theirs.

Under this approach, the question raised when people ask if it is right for them to take a certain choice, like the question raised when they ask if it is prudent, varies from person to person. A asks if it is right _{A} , B asks if it is right _{B} , and so on. The non-consequentialist claim is that it is right _{A} for A to instantiate P in A ’s life and similarly that it is right _{B} for B to instantiate P in B ’s life. And so universalization does not commit a non-consequentialist A to thinking that any general pattern is right _{A} ; that commitment would raise the same problem as before. It commits A to thinking only that as A ’s instantiating P is right _{A} , so B ’s instantiating P is right _{B} , and so on.¹⁵

¹⁴ See James Dreier, ‘Internalism and Speaker Relativism’, *Ethics*, 101 (1990), pp. 6–26. On the recourse to indexicalization, and other reasons for invoking it, see I.L. Humberstone, ‘Two Kinds of Agent-Relativity’, *The Philosophical Quarterly*, 41 (1991), pp. 144–66, esp. pp. 148–9, where Humberstone envisages a formulation of ethical egoism that parallels the formulation which I envisage here for non-consequentialism.

¹⁵ Such non-consequentialists may hold that for any X and Y , if it is right _{X} that X do something, then it is right _{Y} that Y allow X to do it; the important point is that they do not hold that in such a case it has to be right _{Y} that X do it.

It may be implausible to take the two steps described – to indexicalize and decentre ‘right’ – but it is certainly not an absurd strategy for non-consequentialists to adopt. When we use the word ‘prudent’ we do often change the indexical reference of the term – here I mean prudent for me, there I mean prudent for you – and so there is no incoherence in saying that when I speak of rightness, now in relation to my actions, and now in relation to yours, the index of the predicate shifts.¹⁶

If non-consequentialists are forced to adopt this strategy, however, then that is very revealing. It shows that non-consequentialism involves a counter-intuitive relativization of rightness. Where consequentialists hold that the successful justification of a moral judgement must always take people back to common values – it must show that the judgement is required by those shared values – non-consequentialists lower the standards that justification is required to satisfy. On their view, all that justification need do is to show each person addressed that the judgement made, and the action taken, was reasonable from the agent’s point of view; in particular, it was a counterpart of the judgement and the action that each might himself have endorsed in his own case.

On this approach, speaking of what is right in one’s own case never amounts to speaking about a common property. It amounts only to employing a common schema whereby each participant can refer to the property that matters to himself, the property of being right as that is indexed to his position. My trying to argue that something I did was right, therefore, is not an exercise pursued with reference to any common lights. It is an attempt to show that the lights whereby the action looked acceptable to me, while they are not necessarily shared with any others, are at least lights that others should understand. It is an attempt to display not an identity in evaluative commitments, but rather an isomorphism of evaluative standpoints.

Here is an analogy. Patriots all care about their own countries, and to that extent have something in common. But the property that moves the patriot in each country is different from the properties that move patriots elsewhere: it is the property of serving his particular country that each finds moving, not the property of serving whatever one’s country happens to be. Similarly, on the formulation of non-consequentialism emerging here, the

¹⁶ But the sort of move envisaged here is not consistent with an expressivist, as distinct from a cognitivist, viewpoint. For according to expressivism ‘right’ in the speaker’s mouth will always express the speaker’s attitudes, never the attitudes of an arbitrary agent *X*; it will always serve in parallel to the way ‘right_A’ serves for the story told here, never in parallel to the way ‘right_X’ serves. Dreier, in ‘Accepting Agent-Centered Norms: a Problem for Non-cognitivists and a Suggestion for Solving It’, *Australasian Journal of Philosophy*, 74 (1996), pp. 409–21, offers the non-cognitivist a way out of this difficulty, using the approach I describe in the next section. I do not think much of that escape myself, but this is not the place to argue the point.

property invoked to justify any particular agent will be a different property from that which is invoked in justification of others. What such agents are required to have in common need be nothing more substantive, then, than the common structure of attachment that characterizes patriots.

V. OTHER WAYS OUT

I have described a way in which non-consequentialists can escape the problem posed for them by the requirement of universalizability – the simplest mode of escape, as it seems to me – and I have argued that this commits non-consequentialism to a counter-intuitive relativization of rightness. But there are other accounts of non-consequentialism that might also claim to offer an escape from this problem, and my last question is whether they involve the same sort of relativization. I believe that, one way or another, they all do have this relativizing effect – in this respect they do not differ from the indexicalizing–decentering mode of escape – and I would like to support that claim by consideration of two of the most sophisticated of recent suggestions.

Dreier ('Accepting Agent-Centered Norms' p. 419) offers a first example of how non-consequentialism might escape the problem posed, without indexicalizing or decentering the rightness predicate. According to his approach, one he develops for other purposes, we should acknowledge three claims:

1. Saying that an action like promise-keeping or promoting happiness is right involves recognizing a normative reason for wanting the property of promise-keeping or promoting happiness, the target-property of the action
2. Wanting such a property means wanting to instantiate it oneself, wanting to be a promise-keeper or a promoter of happiness
3. Therefore judging an action right, and approving of its target-property, is consistent with not recognizing a normative reason to want others to instantiate that property and perform that action.

One will certainly not want others to instantiate the property, Dreier suggests, in those perverse cases where their instantiating it is inconsistent with one's instantiating it oneself. That others promote happiness may always be consistent with promoting it oneself, but that others keep their promises need not be consistent with one's keeping one's own. And, other things being equal, one will not want others to keep their promises – one will want them not to keep their promises – if that means that one cannot keep one's own. One will describe their promise-keeping as right – one will approve

of that target-property – but still not want them to keep their promises. And that will be fine; for the judgement that their promise-keeping is right, on this approach, does not provide a normative reason for wanting them to keep their promises.

On this approach, the judgement that an action is right will only have reason-giving significance for what one should desire and do oneself. It will amount to a judgement that the target-property in question is one that it is right to instantiate and reasonable to prefer, where that means reasonable to prefer that one instantiate it oneself. And so the judgement that an action is right will only provide a ground for preferring that one perform it oneself, where that is relevant; it will have no bearing on how one should form preferences in relation to the actions of others.

But this approach offers no consolation for non-consequentialists, since it supports the same sort of relativization as the indexicalizing–decentering strategy described in the last section. The indexicalizing–decentering strategy leaves in place the rule of inference whereby the rightness of two or more actions, their rightness in the same sense, would argue equally for there being a good reason to prefer that each be performed regardless of the identity of the agent. Thus, consistently with leaving that rule of inference in place, it argues that the non-consequentialist must find a different sense of rightness at work in claims about what it is right that an agent himself should do and in claims about what it is right that an arbitrary agent should do. Indexicalization and decentering ensure that such a different sense is available.

What Dreier's approach does instead is revise the rule of inference mentioned. It construes things so that the observation that an action is right – that it realizes a target-property of which the observer approves – only ever offers a reason for why that person should prefer that he should perform the action; it never gives him reason, just on its own, to prefer that anybody else should do so. This construction involves relativizing the rule of practical inference whereby judgements of rightness deliver reasons for preference in a way which tracks the strategy that I have described. Where that strategy would relativize the judgement of rightness itself and keep a standard rule of inference, this construction achieves the same end by keeping the judgement of rightness standard and relativizing the rule of inference. The two approaches invite non-consequentialists to relativize the significance of rightness; they differ only in where the recommended relativization is introduced.

A second account of non-consequentialism that might seem to offer a way out of the universalizability problem is presented in a series of papers by David McNaughton and Piers Rawling. What they do is suggest that

non-consequentialists assign a canonical form to judgements of rightness such that under this canonical form universalization proves no problem: it does not generate a slide into consequentialism.¹⁷

The canonical form that they recommend for relatively simple judgements bearing on what an agent is obliged to do is this: I should ensure (for short, IS) that I do O in C , or, schematically, IS(I do O in C). They claim that non-consequentialists will universalize that judgement in the spirit of a corresponding formulation: for any X , X ought to ensure that X does O in C , or, again schematically, $(\forall X)[XS(X \text{ does } O \text{ in } C)]$. And, crucially, they read that universal prescription in turn in such a way that I, the universalizing subject – I, the agent who is given normative reason by the original judgement to want that I ensure that I do O in C – am not given the same reason, in the act of universalization, to want that arbitrary X ensures that X does O in C .

What McNaughton and Rawling do in imposing their canonical formulations and interpretations is regiment language in such a way that the problem I posed for non-consequentialism cannot be raised. That problem arises only if we are allowed access to a deontic operator or predicate – ‘right’ or ‘ought’ or ‘should’, or whatever – that can serve to articulate what there is normative reason for me to want, even when the matter in question involves the actions of others rather than my own. McNaughton and Rawling allow me to say that for arbitrary X , X ought to ensure that X does O in C . But they insist that this is to be read in such a way that the ‘ought to ensure’ does not have the same force as the ‘ought to ensure’ in ‘I ought to ensure that I do O in C ’. It does not articulate a normative reason for preferring that arbitrary X ensures that X does O in C , in the way it can articulate a normative reason for preferring that I ensure that I do O in C .

This way of regimenting deontic language, however, is effectively equivalent to prescribing the indexicalizing–decentring move. If we were forced to indexicalize ‘right’, and to decentre the index in universal prescriptions, then we would be forced into precisely the groove in which McNaughton and Rawling would pin us. A could say in his own case that it is right_A that he do O in C , or ensure that he do O in C . And the only universalization available to A would be to say that for arbitrary X , it is right_X that X do O in C , or that X ensure that he does O in C . Thus it could be held that while the self-prescription gives A normative reason to prefer his doing O in C , the universal prescription does not give A normative reason to prefer that arbitrary X does O in C .

¹⁷ McNaughton and Rawling, ‘Agent-Relativity and the Doing–Happening Distinction’, *Philosophical Studies*, 63 (1991), pp. 167–85; ‘Honoring and Promoting Values’, *Ethics*, 102 (1992), pp. 835–43; ‘Value and Agent-Relative Reasons’, *Utilitas*, 7 (1995), pp. 31–47.

Dreier's interpretation of non-consequentialism would impose the same sort of relativization as the indexicalizing–decentering strategy, locating the relativization in a different place, i.e., in the rule of inference from the judgement of rightness, rather than in the content of the judgement itself. The interpretation offered by McNaughton and Rawling would have exactly the same relativizing effect as those strategies, but would dress up that effect as an artefact of what, allegedly, is the one and only correct way to formulate and interpret self-prescriptions and universal prescriptions. I think that the indexicalizing–decentering strategy ought to be preferred by non-consequentialists, on the grounds that it makes clearer what is going on and how exactly they manage to avoid the slide from universalizability into consequentialism. But that is not a point that I wish to establish now.

The lesson I want to emphasize, rather, is that no matter how non-consequentialists go in formulating the import of their stance, they do at some point have to embrace a distinctive sort of relativization. They have to think that moral justification cannot be required to resort, even at the most ultimate level, to values that all can countenance in common. They have to think that the best and only sort of justification that someone can offer to others for the action he took is this, that it is a sort of action that they would have found justifiable were they in the agent's position, not that it is a sort of action that advances anything that they should care about in common. Non-consequentialists have to think that the best that moral agents can achieve in the way of mutual understanding is the kind of shared comprehension that patriots from different countries achieve, when they recognize that while their concerns are different, and even perhaps in competition, they are at least isomorphic. Non-consequentialists have to think, in an even sharper parallel, that it is the kind of mutual understanding that egoists attain when they recognize that while their hearts are set on quite different ends, they do pulsate to the same egocentric rhythm.¹⁸

Australian National University

¹⁸ My warm thanks to Don Regan, who challenged me at every step in the argument. My thanks too to the audiences at a number of seminars where I presented the paper: at the Colloque de Cérisy, La Trobe University, Monash University, the University of Canterbury, Christchurch, L'Université de Montréal, Brown University and St Louis University. I was particularly helped by comments from Diane Proudfoot and Jack Copeland, and by discussion with Lloyd Humberstone, Peter Railton, Piers Rawling, Michael Ridge and Michael Smith.