

UNIVERSALIZABILITY AND THE COMMITMENT TO
IMPARTIALITY

I

Let us assume here that we are willing to reason and act within the bounds of morality, that there is something that we recognize as the moral point of view which we are committed to and which in certain respects circumscribes our actions. Such a conception would also have a conception of justice as a proper part and it is at least plausible to believe that such a conception would have as an integral element a *principle of impartiality: a principle to the effect that all human beings have an equal right to the fulfillment of their interests*. What I want to inquire into is whether there is a statement of the *principle of universalizability* which is both *categorically prescriptive* and rationally undeniable which entails *the principle of impartiality*.¹ In fine, can we get *the principle of impartiality* from *the principle of universalizability*? I want to know whether it is the case that, if you start from *universalizability* and you are consistent, you will be required also to accept *the principle of impartiality*?

The received wisdom in ethical theory now (1982), in the aftermath of the extensive discussion of R. M. Hare's and Marcus Singer's accounts of universalizability and its placement in morality, is that no such acceptance is required.² I want to consider here whether the received or at least dominant view is also the right view or at least the most plausible view.

'Universalizability' is a term of art and there are a number of different readings of it. I want to start with something which is tolerably unproblematic and to move, only as far as necessary in the way of increasing precisification, to something which is so tied to the very concept of morality and moral judgment, that, if there are any conceptual truths at all, it cannot be denied that all moral judgments are in that sense universalizable. That is to say, I want to state a principle of universalizability which is unassailable and holds in virtue of what it is for something to be a moral or evaluative judgment.

II

To begin, then, at the beginning. *To say that a judgment of rightness or wrongness is universalizable is to say that whatever is right or wrong for*

one person is right or wrong for every relevantly similar person in relevantly similar circumstances. More generally, and still hopefully uncontentiously, we might put it as J. L. Mackie puts it in the beginning of his discussion of universalizability: "Anyone who says, meaning it, that a certain action (or person, or state of affairs, etc.) is morally right or wrong, good or bad, ought or ought not to be done (or imitated, or pursued, etc.) is thereby committed to taking the same view about any other relevantly similar action (etc.)."³ To claim universalizability for norms and evaluations is to claim that if *A* is a good pencil (teacher, comrade, institution) then anything exactly or precisely like *A* is also a good pencil (teacher, comrade, institution).⁴ Intelligibly to assert that one pencil, teacher, comrade or institution is good and the other is not, there must be some difference between them in virtue of which this judgment can be made. The same thing holds for judgments of rightness or judgments concerning what ought to be done or what ought to be the case or what someone is morally justified in doing. If Axel is justified in doing *A* then anyone exactly like Axel is justified in doing anything exactly like *A* in exactly the same circumstances. Similarly, if *A* is the right thing to do for Mary or if *A* is something Mary ought to do, then anyone exactly like Mary and in the same circumstances ought to do *A* as well and the doing of *A* in such a circumstance is also the right thing for them to do.

It will, of course, be said that the people, objects, roles, institutions in question are never, and never can be, exactly alike, for they are at least numerically distinct. It is correct to respond to that that being merely numerically different doesn't count, but, that response notwithstanding, it still remains the case that, numerical differences apart, there are also always differences between people, institutions, roles and the like.

To salvage universalizability from uselessness 'relevantly similar' must, in the above formulations, replace 'exactly alike' or 'exactly the same' and the like.

A perhaps more useable *principle of universalizability* will read: if an action or attitude is right, if something ought to be done or to be the case, or if something is good, then, anything relevantly similar to it is also right, ought to be done or ought to be the case or is good in relevantly similar circumstances and (where this is relevant) for relevantly similar persons.

The most obvious problem here is to specify acceptable criteria for what are to count as *relevant similarities*. It appears at least to be the case that nothing formal or definitional, except arbitrarily by way of stipulation, will establish criteria for relevant similarities. It appears at least to be the case that non-formal, and perhaps invariably contestable, evaluative or normative

judgments must be made in coming to an agreement about what the criteria for relevant similarity are.

It does not appear at least to be the case that we can ascertain what the criteria are from simply becoming clear about the concept of morality or about what moral judgments are or what it is to take the moral point of view such that all informed people with such an understanding must agree on what the criteria for relevant similarities are.

It is also important to notice that the *principle of universalizability* does not tell us *what* is right or wrong, good or bad, or *what* ought or ought not to be done. It rather says that *if* one thing is right, good or ought to be done, then another thing relevantly similar to it is too.

The *principle of universalizability* so construed does not, note, take us to the *principle of impartiality*; namely, to a principle which asserts that *all human beings have an equal right to the fulfillment of their interests*. Only if we judge that there are no relevant differences in this respect between people could we conclude that all human beings have an equal right to the fulfillment of their interests on the ground that if *A* has a right to the fulfillment of his interests, then anyone else does too, because there are no relevant differences between *A* and other human beings. But that there are no relevant differences between *A* and any other human being will be challenged, e.g. the Son of Sam is not Malcolm X, Ronald Reagan is not Salvadore Allende. Moreover, the *principle of universalizability* will not itself determine that anyone, let alone everyone, has a right to the fulfillment of her/his interests. What it will tell us is that *if* Fran has a right to the fulfillment of her interests, anyone relevantly similar to Fran and in a relevantly similar situation will have an equal right to the fulfillment of her or his interests as well. It will not be right, in this respect, to treat Hans differently from Fran, unless there is a relevant difference between them or a relevant difference in their situations which justifies a difference in treatment. But the *principle of universalizability* or anything entailed by it will not tell us whether there is a relevant difference between them. It will not tell us that Fran has a right to have her interests fulfilled, let alone tell us – or legitimately give us to understand – that all of us have an equal right to the fulfillment of our interests.

Only on the assumption that we are sufficiently similar to be entitled to have our interests equally protected, to be equally free from the deliberate frustration of our interests, can we correctly claim that this is something the *principle of universalizability* supports. But without that assumption about human nature and entitlement, such an equal concern for the interests of

all human beings is not established by the consistency requirements of *universalizability* alone. *Universalizability* by itself does not give us the rational foundations of justice.

These normatively substantive things the *principle of universalizability* cannot do. Yet, what seems at least unassailable about the *universalizability thesis* is this: on the one hand, if something is right, or if something ought to be done or if something is good, and, on the other hand, if another thing is not right, ought not to be done or is not good, then there must, simply as a matter of logic, be some difference between them which explains why one thing is right and the other not, why one thing ought to be done and the other not, why one thing is good and the other not. In spite of Don Locke's sharp criticism of the appeal to *universalizability* in ethics, this is something that he and R. M. Hare and Paul W. Taylor are all agreed upon.⁵ This means that if something is the morally inappropriate thing to do in a certain circumstance, then there must be some reason why it is morally inappropriate; to say 'This is wrong' is to presuppose a moral principle — 'Whatever is like this in certain respects is also wrong.' If you want to talk that way, you can say that this means that the Principle of Sufficient Reason applies in ethics.

What is very often claimed by moral philosophers — philosophers I like to call ethical rationalists — who try to give some strategic deployment of *universalizability* in moral philosophy is the cluster of claims I shall make in the rest of this paragraph. When we reflect on what moral principles really are — what essentially their nature is — we will come to see that moral principles do not make exceptions of particular individuals just because they happen to be particular individuals. If a moral principle applies or does not apply to a particular person, it is not because he or she is that particular person, but because he or she possesses some feature or combination of features which could, at least in principle, be possessed by others.⁶ It does not matter *who* you are — that is morally irrelevant; what counts is only general characteristics of you and your situation.

What is tempting to conclude from this, but what is now widely recognized to be a mistake, is that in seeing that moral principles are in the above senses *universalizable*, we can conclude that they are universal principles which apply to everyone alike and allow no exceptions, such that if I ought to do *Y*, you ought to do so as well. Bernard Mayo, for example, maintained in his *Ethics and the Moral Life* that a moral judgment "must be *universalizable* in the sense that it applies not only to me but to you; not only to you but to me; not only to us but to everybody. . . ."⁷ But this *universalist* claim

for *universalizability* plainly is not on. You can't get from *universalizability* to *universalism*. Some moral principles are meant to apply only to children or to old people or to soldiers or to doctors or to religious people or to Roman Catholic priests. The moral principle enjoining priests not to divulge what is told in the confessional, a doctor's Hippocratic oath, the duties of parents to children, the obligations of a psychoanalyst to his patients are very special and situational. The hat trick the *universalizability principle* does not pull off is that of grinding out universal substantive moral principles for us.

There is a sense in which all moral principles, indeed all principles *sans phrase*, are *universalizable*, but not in the sense that they are universal or that in general that they apply to everyone alike. A principle applies to everyone who falls within its scope, but that is a different matter. It is not even true that all moral principles apply to everyone irrespective of who he is. Some, as we have seen, apply only to priests, to doctors, to children, to parents, and, in some rare instances, only to a certain individual, e.g. to Jesus. "Moral Principles", as Don Locke has nicely put it, "may be universal in the sense of applying to everything of the particular sort . . . they are not universal in the sense of being applicable to everyone. . . ." ⁸

It is true, of course, that they are in a sense generalizable even when applying, *in fact*, only to one individual. Thus, if Jesus should have accepted death on the cross to save humankind, it is true that anyone relevantly like Jesus and in a relevantly similar situation should do so as well. Even acts of supererogation are universalizable. If a person with relevantly similar commitments and other relevantly similar qualities and in a relevantly similar situation hove into sight, then he/she should also act as Jesus did. We cannot, of course, where the act is really an act of supererogation, insist that he must; 'must' is a stronger moral term than 'should': not all oughts are obligations. Typically very admirable acts, as heroic or saintly acts, are not required of us, though we, if we have considerable moral integrity, may require them of ourselves. If acts of supererogation were things society can rightly hold us to, as something we *must* do, then they would not, indeed could not, be acts of supererogation. But we still correctly, in the above circumstance, are to say that he should so act. We don't require it of him but we think his doing so is desirable. It is, to return to Jesus, an historical accident that only Jesus was in that situation. It is not a logical truth or any kind of conceptual truth, that no one else could be. Similarly, though parents have special obligations to their children, anyone who comes to be a parent has such obligations. This is what the universalizability thesis was trying to catch in claiming that moral principles applied to everyone alike. Still, in the sense

specified above, moral judgments or principles do not all apply to everyone alike, and in that important straightforward sense, they are not all *universal*, though they are *universalizable*.

III

Can we, in the teeth of what I have argued above, derive the *principle of impartiality* ("All human beings have an equal right to the fulfillment of their interests") from the *principle of universalizability* in the way we have formulated it, either by immediate inference or in conjunction with some purely factual premises or from factual premises plus analytic premises?

My statement of the *principle of universalizability* was this: If an action or attitude is right, if something ought to be done or to be the case or if something is good, then anything relevantly similar to it is also right, ought to be done, or ought to be the case, or is good in relevantly similar circumstances, and for (where this is relevant) relevantly similar persons. But someone could plainly accept that and assert that we are so different that we do not all have an equal right to the fulfillment of our interests. Some of us are cruel, some kind, some lazy, some industrious, some avaricious, some generous, some without integrity, some with it, some concerned principally to feather their own nests, some with an intense concern for others, and most of us, in these and other respects, are at neither extreme of virtue or vice. I am not saying that the person who claimed that because of these manifest differences between us we do not all have an equal right to the fulfillment of our interests is justified in making that claim or that he should make that claim. What I am saying is that that claim is at least not self-contradictory or that there is anything *conceptually* incoherent or untoward about the assertion of it or that it violates the *principle of universalizability*. Something can be morally anathema yet conceptually in order, even perfectly intelligible, as a bit of moral discourse.

Yet such an assertion contradicts the *principle of impartiality* while remaining perfectly compatible with the *principle of universalizability*. Thus, the *principle of impartiality* cannot be derived from the *principle of universalizability*. An elitist or social Darwinian can believe that if something is right for him to do, it is also right for others to do who are relevantly similar to him and similarly placed — and thus accept universalizability — and still perfectly consistently deny that the great masses of people share a right to do what he and his elite peers have a right to do even when doing it would aid in the fulfillment of their interests, for he sincerely believes that such

run-of-the-mill people are not relevantly similar to him. He accepts the *principle of universalizability*, but rejects the *principle of impartiality* (as we have defined it) and does this quite consistently. We may — I do, at any rate — wish to set ourselves in struggle against such elitists or social Darwinians, but we have no good grounds for thinking they are being inconsistent or have said anything conceptually untoward. Morally untoward, perhaps, but not conceptually untoward.

The same point can be made if we accept Don Locke's formulation of *universalizability*. Suppose we say that a person cannot sincerely and consistently make a moral judgment unless that person is prepared to accept "the same moral judgment in any relevantly similar situation, including those which affect you personally differently".⁹ The elitist can say quite consistently, while accepting Locke's formulation, that he is so different and his relation to people is so different that he is not in a relevantly similar situation to that of the vast majority of other men. Since this is so, it would be wrong to claim that they had, compared to him and his peers, the same equal rights to fulfillment of their interests. There is no denial of universalizability in Nietzsche's doctrine of the *Übermenschen* and *Untermenschen*. But there is a rejection of the *principle of impartiality*. I deplore such a rejection of the egalitarian attitude. But I do not think that it is an inconsistent position.

IV

Paul W. Taylor, in an unfortunately neglected essay, "Universalizability and Justice", argues that universalizability logically compels an agent to assent to the following general principle: "If it is wrong for another to discriminate against him (the agent) on the ground of a difference he (the agent) does not acknowledge to be relevant, it must also be wrong for him (the agent) to discriminate against another on the ground of a difference the other does not accept as relevant."¹⁰

To probe this, consider the following rather desert-islandish case: Suppose the agent is a Jew living in an anti-Semitic society with strange religious beliefs that forbid inoculation for diseases including infectious diseases. He, in this anti-Semitic religious society, is not allowed to hold a professorship in any university on the grounds that he is a Jew. He rightly resents this and rightly enough does not acknowledge this as a relevant ground for discrimination, though it is all the same widely regarded in the society in which he lives as being a morally and legally acceptable relevant ground for such special treatment. Indeed it is part of the law of the land. He immigrates to a saner

society. There there are no barriers, legal or otherwise, to his becoming a professor, but eventually the idea loses its allure and he becomes instead an immigration officer in his newly adopted land. A typically religious member of his former homeland wants to visit his new-found homeland, but refuses to get inoculations for a disease that he might quite likely be the carrier of or very well might contract and subsequently spread. He will not do so because it is against this putative tourist's religious principles to have such an inoculation. Indeed, as he sees things, to do so would be to act immorally. As an immigration official, our former professorial aspirant bars this person from entry into his country on the grounds that he has not had these inoculations. (Let us suppose he has legal discretion to do so or not to do so.) He does so on the ground – indeed a not implausible ground (to understate the point) – that the aspiring tourist may spread a very dangerous infectious disease. But the aspiring tourist, let us assume, no more regards this discrimination against him as relevant than he, the immigration officer, regards his being barred from teaching in the universities of his former homeland on the ground that he is a Jew as relevant. Yet, it is very likely the case that his action in barring the would-be tourist was the morally correct decision to make. Surely it was not an unreasonable decision to make. Yet, if this is so, it violates Taylor's principle that he takes to be derivable from the *principle of universalizability*. A (the immigration officer) discriminates against B (the religious fanatic) on a ground that B does not acknowledge as relevant, but A regards his action as justified (morally right), but he insists that B is wrong to discriminate against him (A) on grounds that he (A) does not acknowledge as relevant. The symmetry of Taylor's principle is broken; there is no reciprocity here as Taylor's principle requires. The *principle of mutual acknowledgement*, a principle central to Taylor's argument, is the principle that a relevant difference, which would justify treating one person differently than another, must be such that it will be freely acknowledged by all whose interests are affected by such proposed differential treatment. Yet, it surely seems at least to be a principle which would be rightly rejected or at least overridden in such a context. Surely it is anything but evident that there must, in such circumstances, be the reciprocity of mutual acknowledgement that Taylor's principle requires for there to be a morally valid decision. It is not evident that morality requires it, let alone that not so acting is 'morally unintelligible', because it violates 'the logic of normative discourse'. Even if the moral situation were more anomalous than I believe it actually is, A's actions are still intelligible as moral actions. If they are incompatible with the *principle of universalizability*, then we should be suspicious of the claim that there

can be no moral claims which are not universalizable. However, I think *A*'s actions are plainly as compatible with universalizability as are actions in accordance with Taylor's principle of mutual acknowledgement. If this is so, Taylor has not shown that if we accept universalizability, we must accept his principle. *A* can (to show that it is so), as well as a person committed to the principle of mutual acknowledgement, believe that if *X* is right for *A* to do in situation *Y*, it is right for any relevantly similar person to do in a relevantly similar situation. The value replacing the variable *X* in *Y* could be 'prevent people not inoculated for diseases *L*, *M* and *N* entering his country at time *T*, whether or not they regard that discrimination as relevant'. He will quite consistently believe that *B* should also so act in situation *Y*, even if *B* does not see the situation in those terms. So, someone can quite consistently accept universalizability and still accept the negation of Taylor's principle, so it cannot be the case, as Taylor contends, that one is logically compelled to accept Taylor's principle if we accept universalizability. Taylor has not succeeded in showing that there is anything in the very nature or logic of moral discourse or normative discourse which compels us to reject criteria of relevant differences, not acceptable to all whose interests are affected by them. He has not shown how a clairvoyant understanding of the principle of universalizability must lead us to acknowledge that "criteria of relevant differences, not acceptable to all whose interests are affected by them, cannot be used as grounds for excluding anyone from having an equal right in deciding what criteria are to govern those social practices that affect his interests".¹¹ It is not clear that *A* and *B*, in my above example, have an equal right in deciding criteria of relevance for who can enter a country and who cannot. And even if I am somehow mistaken on this substantive moral point, I am not mistaken because I have said something which is incompatible with the *principle of universalizability*, while my interlocutor has consistently followed out its logic. What is morally mistaken or untoward need not be morally unintelligible or incoherent.

V

Getting substantive moral principles or maxims out of the *principle of universalizability* is like getting blood out of a turnip. There is a kind of ethical rationalism — often a strain of Kantianism, though, as John Rawls has shown, not the only Kantian strain — that seeks the rational foundations of justice and sometimes the whole of morality in an *a priori* or formal *principle of universalizability* which tries to establish itself as a principle which is rationally

undeniable for anyone who can think clearly and has a firm understanding of moral discourse. I have not sought to show that the *principle of universalizability* can be rationally denied (that we could give conclusive or even rationally persuasive reasons for denying it), but I have sought to show that it is rationally deniable and indeed highly improbable that such a *principle of universalizability* is the 'rational foundation' of justice or of morality more generally. It is understandable that such a formalism and ethical rationalism should have such a strong hold on us. If it were true, logic alone, pure reason alone, would, at least in a general way, tell us how we ought to live. But such a rationalism is a myth. In moral philosophy it returns again and again like the repressed.¹² For moral philosophers on the quest for certainty, it, like the ontological argument for the existence of God, holds a persistent and powerful attraction, but both conceptions are thoroughly myth-eaten, revealing, once again, the irrational heart of philosophical rationalism.

Department of Philosophy
University of Calgary

NOTES

¹ D. H. Monro, *Empiricism and Ethics* (London, England: Cambridge University Press, 1977), Chapters 13–16.

² R. M. Hare, *The Language of Morals* (London, England: Oxford University Press, 1952); R. M. Hare, *Freedom and Reason* (London, England: Oxford University Press, 1963); and Marcus Singer, *Generalization in Ethics* (New York, NY: Alfred Knopf, 1961). The criticisms have been myriad and varied but some of the classic ones include: Don Locke, "The Trivializability of Universalizability", in *Readings in Ethical Theory*, eds. Wilfrid Sellars and John Hospers, 2nd edn. (New York, NY: Appleton-Century-Crofts, 1970), pp. 517–527; David Keyt, "Singer's Generalization Argument", *Philosophical Review* 72 (1963), pp. 466–476; George Nakhikian, "Generalization in Ethics", *Review of Metaphysics* 17 (1964), pp. 436–461; Peter Winch, "The Universalizability of Moral Judgments", *Monist* 49 (1965), pp. 196–214; and Joseph Gilbert, "Neutrality and Universalizability", *Personalist* 53 (1972), pp. 438–441.

³ J. L. Mackie, *Ethics: Inventing Right and Wrong* (Harmondsworth, Middlesex, England: Penguin Books Ltd., 1977), p. 83.

⁴ R. M. Hare characterizes universalizability as follows: the thesis of universalizability requires that if we make any moral judgment about a determinate situation, we must be prepared to make the same judgment about any precisely similar situation (p. 42). To say that moral judgments are universalizable is to say "that they entail identical judgments about all cases identical in their universal properties" (p. 108). Hare further remarks, "... if we make different moral judgments about situations which we admit to be identical in their universal descriptive properties, we contradict ourselves" (p. 21).

R. M. Hare, *Moral Thinking: Its Levels, Method and Point* (Oxford, England: Clarendon Press, 1981). The universalizability thesis, so stated, seems to me manifestly a correct conceptual thesis about moral discourse. The interesting question is whether it can generate any substantial moral norms or provide, as Hare believes, rational guidance in moral reasoning. My contention is that it does not. Where the universalizability thesis can be established simply by careful attention to the logic of moral discourse, it is trivial. Where it is given a richer reading it is not a thesis which is required by reason.

⁵ Don Locke, "The Trivializability of Universalizability", p. 520; Paul Taylor, "Universalizability and Justice", in *Ethics and Social Justice*, eds. Howard E. Kiefer and Milton K. Munitz (Albany, NY: State University of New York Press, 1968), pp. 142–163; and Paul Taylor, "On Taking the Moral Point of View", *Midwest Studies in Philosophy* 3 (1978), pp. 35–61.

⁶ Don Locke, "The Trivializability of Universalizability", p. 524.

⁷ Bernard Mayo, *Ethics and the Moral Life* (New York, NY: Macmillan, 1958), p. 91.

⁸ Don Locke, "The Trivializability of Universalizability", p. 527.

⁹ Don Locke, "The Principle of Equal Interests", *The Philosophical Review* 90 (October 1981), p. 535.

¹⁰ Paul Taylor, "Universalizability and Justice", p. 161.

¹¹ *Ibid.*, p. 167. R. M. Hare's latest book came into my hands too late to include a discussion of it in the body of my text. However, it does seem to me that his latest effort at achieving morally interesting results from universalizability runs afoul of the very same counter case I directed against Taylor. Moreover, it is absurd to claim, as Hare does, both that universalizability is a logical requirement for those who are going to reason morally and that it also "demands that we treat other people's prescriptions (i.e. their desires, likings, and, in general, preferences) as if they were our own" (p. 17). That latter claim is not on as a logical thesis. It would only gain plausibility if it had a different logical status and if we had made and justified the *substantive moral claim* that people are sufficiently alike to justify such similar treatment or that, alike or different, they are to be so similarly treated in these respects. But there are no conceptual guarantees for such claims which would require us to accept them on pain of logical error. R. M. Hare, *Moral Thinking*, p. 17. I should add here that the most uncompromising form of ethical rationalism is not in Hare's work, but in the work of Alan Gewirth. I have examined it in some detail in my "Against Ethical Rationalism", in *Gewirth's Ethical Rationalism*, ed. Edward Regis, Jr. (Chicago: The University of Chicago Press, 1984).

¹² I have, in various ways, tried to criticize such rationalism in my "On Needing a Moral Theory", *Metaphilosophy* (1982); "Grounding Rights and a Method of Reflective Equilibrium", *Inquiry* (1982); "Problems of Westermarck's Subjectivism", in *Edward Westermarck: Essays on His Life and Works*, ed. Timothy Stroup (*Acta Philosophica Fennica* 33, 1982); and "Baier on the Link Between Immorality and Irrationality", *Nous* March 1982).