

Working in and Working to Principles: Penn's Lie and Hare's Myth of Universalizability



Edward W. James

Ethics, Vol. 83, No. 1 (Oct., 1972), 51-57.

Stable URL:

<http://links.jstor.org/sici?sici=0014-1704%28197210%2983%3A1%3C51%3AWIAWTP%3E2.0.CO%3B2-N>

Ethics is currently published by The University of Chicago Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/ucpress.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

Discussion

WORKING IN AND WORKING TO PRINCIPLES: PENN'S LIE AND HARE'S MYTH OF UNIVERSALIZABILITY*

Edward W. James

Bates College

When William Penn was faced with the loss of Pennsylvania, he lied by promising the Crown that the Pennsylvania Quakers would take part in defense measures—something he knew they would never do. Assuming Penn was morally right to have lied, could he also have been morally right if he had instead told the truth?

A widespread assumption in ethical philosophy would lead us to answer negatively. This assumption, which R. M. Hare calls the principle of universalizability, states that when two similar persons are in similar moral situations and one of them is obligated to do a certain act, so is the other. Hence, by it Penn would have been morally right only in lying. Although not clearly articulated until recently, this assumption can be seen working in the attitudes taken and questions posed by philosophers, moralists, and "ordinary people." It can be seen in Plato's Socrates asking how we (all) should live; in Hobbes, Bentham, Mill, and Moore telling us to do, in a given situation, that (one) act, or that (one) type of act, which will promote the greatest good; in Prichard and Ross saying that we should do what we (all) know we ought to do; in Baier specifying when killing is right and when it is wrong;¹ in the other Sartre, the responsible one, speaking of moral man as "a legislator deciding for the whole of mankind";² and in our own moral method, when we ask ourselves what *one*, that is, any given person, ought to do in this situation.

It is to the great credit of R. M. Hare to have seen the significance of this assumption in ethical philosophy and to have stated it as clearly as one could wish: "The judgement that I ought to do a certain thing commits me to the view that no similar person in a precisely similar situation ought to fail to do

* I wish to thank Charlene Entwistle and Professors David Smith and Carl Straub for helpfully criticizing earlier drafts.

1. Kurt Baier, *The Moral Point of View* (Ithaca, N.Y.: Cornell University Press, 1963), pp. 191-95.

2. Jean-Paul Sartre, "Existentialism Is a Humanism," in *Existentialism from Dostoevsky to Sartre*, ed. W. Kaufmann (Cleveland: World Publishing Co., 1961), p. 292.

the same thing.”³ By “precisely similar” Hare does not intend to make each moral situation unique; for moral “circumstances are classifiable into kinds,” which leads to “a similar answer to be appropriate in all circumstances of the same kind.”⁴ Moreover, Hare’s later qualification, that “it is not inconsistent to admit that there may be different ways of life, both of which are good,”⁵ does not touch Penn’s case. For different ways of life are permissible, according to Hare, only when speaking about *style*—that is, only when we are concerned neither with other people’s interests nor with obligation.⁶ In the case of Penn, however, we are concerned with both. Hence Hare’s principle of universalizability tells us that if Penn was right to have lied, he would have been wrong to have told the truth.

As significant as this principle has been, however, I find it troubling. The root of the trouble is that the fundamental moral principles—such as advance the brotherhood of man, work for the greatest good—seem to allow for two major stances toward them. One stance is to *work in* the principles, the other is to *work to* them. When working in the principles, our reasons justify types of actions by claiming that they could occur if the ideals expressed by the principles were perfectly realized, that is, if evils no longer existed. Truth telling and pacifism, for example, work in the principles. But when working to the principles our reasons justify specific actions by claiming that they contribute to the occurrence of situations which allow people to work in the principles. Such actions would involve fighting obstacles to working in the principles—as sit-ins claim to attack the prejudicial structures of society—or “removing” antitheses of the principles—like assassinating a Hitler—or improving the lot of a person or group of persons in a “questionable” way—such as lying to conceal a Jew from a Nazi or to keep Pennsylvania in the hands of pacifists.⁷

This distinction, I believe, suggests that Penn could have chosen, with good

3. R. M. Hare, *Freedom and Reason* (Oxford: Clarendon Press, 1963), p. 153.

4. R. M. Hare, *The Language of Morals* (New York: Oxford University Press, 1964), p. 158. I may freely quote both *Freedom and Reason* and *The Language of Morals*, for there are no substantial differences between them.

5. Hare, *Freedom and Reason*, p. 153.

6. *Ibid.*, pp. 152–56. Hare is not as clear on this as one might wish.

7. Neither stance, then, is necessarily more oriented to the individual or to the society than the other; both view the individual as a moral being associated with other moral beings. And neither stance seeks to achieve the ideals expressed by the principles more than the other; one seeks to achieve them by specific actions contributing to their realization, while the other seeks to achieve them by insisting that the ideals expressed by the principles must function now, in all their fullness, in our personalities. Accordingly, neither stance is more “realistic” or “idealistic” than the other; they are both realistic in recognizing that the ideals have not been achieved and that evil exists, and they are both idealistic in being governed by, or acting from, principles. Moreover, neither stance fights evil more than the other; for both “fight” evil in their own way. Finally, neither stance is necessarily more consequence oriented than the other, for both aim to achieve the ideals expressed by the principles. Hence one can justify his actions by appealing to consequences and still work in principles: Kant’s condemnation of slothful self-development, because it does not further humanity as an end in itself, appeals to consequences but works in principles, namely, by speaking of actions which would occur in the perfect society.

reasons and thus right on his side, either to lie or to tell the truth. As it turned out he lied: When the holy experiment was beset, when the attempt to work in the principles was endangered, Penn decided to work to the principles. Yet had he instead chosen to tell the truth—to work in the principles, even at the cost of losing Pennsylvania—he would have been *just as right*. For Penn would have been acting on the same fundamental principles—for example, advance the brotherhood of man—when he lied as when he, by hypothesis, told the truth. In other words, the different acts would have been equally justified by the same principles. Thus he would not and could not have compromised his principles by lying or by telling the truth; for in both cases the same principles would have been *functioning*, that is, actually determining Penn's thoughts and actions.

Of course, Penn might have erred—by making the factual error of using the wrong means to achieve his end, or the logical error of deducing the wrong act from his moral principles, or the moral error of adopting the wrong stance toward his principles.

Penn might have erred, then, by failing to see that the ideals articulated by his principles could only be achieved by, say, working to his principles. Hence, if he had told the truth, if he had worked in the principles, he would have been using the wrong means to achieve his end. This view would make the differences between Penn's (hypothetical) truth and Penn's lie, and thus between working in and to principles, not moral but factual. For the differences would center on the factually best way to practice the principles and thereby achieve the ideals implicit in them. This view, accordingly, would relegate the distinction between working in and to principles to a mere quarrel within the confines of the ends/means distinction—with each stance maintaining it is the best means—and would thus save the principle of universalizability. For then there would be no moral disagreement concerning how to act: If the factual matter could be settled, everyone could agree on one and only one way to act in a given situation.

The rub of this view, however, lies in the qualification "if the factual matter could be settled"; and the rub's abrasiveness resides in the impossibility of the factual matter ever being settled. For historical questions involving the unreal conditional—like, if Penn had told the truth, he would (not) have best advanced his ideals—suffer from a surfeit of criteria which render factual agreement impossible. Consider Penn's lie: Pennsylvania was saved—for a while. Those working to principles would emphasize the first half of that observation, those working in would stress the second half. The point is that no universally accepted criteria are available to appraise the issue, and thus each side chooses the factual criteria that will best support its moral case. Hence the issue cannot be considered a genuine factual issue but is instead the moral issue of how a person ought to act in a given situation.

Moreover, although the distinction between working in and to principles is a genuine moral distinction, it is not the source of adamant quarrel between the two stances. For the true source of the quarrel is none other than the principle of universalizability, which deems that there is only one way to fol-

low a moral principle in a given situation. This assumption intensifies the differences between the two stances, for each side must then justify itself at the cost of the other. Yet hard common sense favors the view that both stances aid each other. For the two stances complement each other by having as their strong points the weak points of the other. One stance insists on purity of life at the expense of allowing great evils to exist; while the other insists on demolishing such evils at the expense of adopting evil's ways. Hence those working to principles help their counterparts by protecting them from annihilation by a Hitler and by guaranteeing that a Hitler will meet physical pain and resistance. And those working in principles help the other stance by being a restraining force and, perhaps even more important, by giving living substance and content to the ideals expressed by the shared principles. Far from being at odds with one another, then, it seems that the two stances would do well to cooperate with each other.

Another way that Penn might have erred was by making a fallacious deduction. This assumes that moral reasoning is a matter of deducing particular moral directives from general principles. Hare himself seems to travel this route when he argues that specific moral judgments and the general principles which "subsume" them together entail specific moral imperatives. For example, the specific moral imperative, "Don't do *A*," is logically entailed by the particular moral judgment, "*A* is evil," and the general principle, "Don't do evil."⁸ Accordingly, Hare's reasoning seems to go, if two moral situations are identical, then since "every principle that is applicable to this one must be applicable to the other,"⁹ it follows, by what may be called the principle of the deducibility of particular directives, that specific moral imperatives entailed one time will be entailed the next. Hence, after Penn chose his principles it would have been logically impossible for him to validly deduce two opposing conclusions; and the distinction between working in and to principles—in maintaining that Penn could have justified, that is, logically deduced, two opposing maxims—must be rejected.

Hare's theory of moral justification borrows much, obviously, from scientific explanation. And, to be sure, there is much to borrow: Both start with particulars, seek to put them into a larger framework, and try to guide and understand. However, Hare is led astray in adopting as his model a *complete* explanation, where the description of the *particular* event is deduced from general laws and initial conditions—something which rarely happens in science and never in situations with so many parameters as the moral variety. Instead Hare should have thought in terms of the *partial* explanation, where the general class that the particular event belongs to can be deduced but not the description of the particular fact alone or even a subclass of the general class. To use Hempel's example: Freud's unstated law concerning unconscious slips—if a person has a strong desire, slips of the tongue, memory, and the like, which express that desire, will probably occur—will not explain why *this* particular slip oc-

8. Hare, *The Language of Morals*, pp. 174-76.

9. *Ibid.*, p. 177.

curred, or even why it was a slip, say, of the tongue.¹⁰ All it will do is explain why a slip of some sort took place. Much in the same way, because of the choice man has between working in and to principles, moral principles will not logically imply a particular moral directive. Instead they provide us with a framework which justifies our specific actions and maxims by being *expressed* through them—by implying the general class of the particular directive but not the subclass, let alone the directive itself. Penn's principles, for example, would not entail any specific moral directive such as, "Lie now"; but they would justify the directive if the directive expressed them.

Accordingly, the distinction between working in and to principles reveals that moral man has a crucial region of choice in many situations. When he is presented with a choice between working in and to principles, moral man will act from principles, but *he* must choose which stance to adopt. This choice gives him what may be called a *profound freedom*—a freedom which grants two men, or the same man at different times, confronting the same situation with the same moral principles, the power to decide for radically different actions.¹¹ Such decisions are exceedingly difficult, for they determine what sort of person one is to be, what struggles and hardships he is to face, and what kinds of people and experiences he is to meet. But such decisions cannot be called arbitrary or morally wrong. For they are from principles, are based on a desire to follow the principles, are grounded in a knowledge of the situation, reflect one's training and culture, and above all, stem from the fundamental moral distinction between working in and to principles.

Nor does this profound freedom hinge on the choice of the principle: advance the classless society, promote the greatest happiness for the greatest number, seek oneness with God, affirm the absolute sacredness of human life—all fare the same. In affirming the absolute sacredness of human life, for instance, one could conclude that he ought to destroy what denies life's sacredness, even if it is a part of life itself. Of course, one could adopt as his fundamental principle, never take a life, or, never lie. But these would not be fundamental in the ordinary sense of the term; for they are so specific they would not justify other moral principles adopted, for example, be considerate to others, and they would not illuminate most moral situations. Thus it seems that *either* the possibility of a general justification of a moral position is denied and specific moral rules are arbitrarily chosen, *or* the possibility of a general justification is accepted and opposing specific moral rules are acknowledged. Either moral rules make no coherent sense, or there are equally good but divergent moral rules.

One way out of this dilemma is to maintain, with Kant, that it is our stance toward the principles which is primary. By this view the distinction between working in and to principles goes awry in presenting the two stances as equally moral. One and only one stance is moral here, and Penn in our hypothetical

10. Carl G. Hempel, "Explanation in Science and in History," in *An Introduction to Philosophical Inquiry*, ed. J. Margolis (New York: Alfred A. Knopf, Inc., 1968), p. 257.

11. Hare's problem in *Freedom and Reason*—how can a man be free and yet governed by reasons?—finds its answer in this notion of profound freedom.

case would have made the moral error of adopting the wrong stance toward his principles.

The most crucial point against this position, I believe, is the logical one that the stance does not issue from nowhere but is grounded in more general ideas than itself, general ideas which equally justify the opposing stance. That is, by accepting the justification, we commit ourselves to accepting whatever else it may justify—even the opposing stance. For example, Kant held that morality rests on unconditional laws which are in turn grounded in the rationality of consciousness; that this consciousness, by allowing freedom and thus morality to occur, is an end in itself; and that the many ends in themselves group in a realm of ends. From these general ideas Kant concluded that man should act only according to those maxims which could be consistently generalized, which treat men as ends in themselves, and which could be binding in the realm of ends. Yet he could have just as well concluded that man ought to *promote specific* situations which allow him to act rationally, to act as an end in himself, and to exist now in the realm of ends. Kant's conclusions work in principles, but his ideas equally support working to them.

Because of the logical priority of the principles, whenever we have to choose between working in and to principles, we confront an inherently ambiguous situation. In such situations we ourselves make the principles dominant; for at times we choose to act in principles, at times to them, but never do we always opt for one alone. We have our tendencies, but our tendencies do not run on single tracks. Even saints at times work to principles, just as reformers work in them: Gandhi drank milk for his health in spite of his extreme vegetarianism, and Trotsky took pains to specify that killing is almost always wrong. Philosophers, too, have this choice. Brandt for example, when discussing Ross's claim that in ordinary instances we *know* we ought to tell the truth, asks whether it really is the case that we should always tell the truth to a shrewd realtor.¹² Such a question derives its force by putting us within the framework of working to principles, where we consider that this type of realtor does not care for anyone but himself, that he will take advantage of anyone he can, and so on. Yet Brandt counters act utilitarianism by arguing that it would require us to weigh the consequences of each action so that we might even have to lie, cheat, kill, maim, commit adultery, "if the net expectable utility—assuming it can be kept secret—would exceed that of refraining!"¹³ Such an appeal to moral absurdity hinges on whether one adopts the framework of working in principles, where he would never think of doing such things.

These choices—of the saint, the philosopher, the ordinary moral man—may

12. Richard B. Brandt, *Ethical Theory* (Englewood Cliffs N.J.: Prentice-Hall, Inc., 1959), p. 386.

13. *Ibid.*, pp. 390-91. Of course, it is the very point of act utilitarians that we at times *should* do such things; and their examples—such as, one should shoot a convicted murderer about to axe his saintly grandmother—derive their power by emphasizing working to principles. I believe that the distinction between working in and to principles illuminates this controversy by showing why it must be irresolvable at certain times.

be paradoxical, tension filled, and uncomfortable; but they can hardly be called inconsistent or compromising. Rather they reflect the obvious—that good and evil both exist, and that the moral problem is to keep the good and remove the evil. As Penn wrote to the Friends in Pennsylvania: “Things are not just now . . . as you may reasonably desire. . . . We must creep where we cannot go, and it is as necessary for us, in the things of life, to be wise as to be innocent.”¹⁴

In conclusion, the distinction between working in and to principles allows Penn to have been moral even if he had done the opposite of what was assumed as right and thus leads to a rejection of the principle of universalizability. Indeed even a weaker, Kantian version of the principle, act so that you could will your maxim to be a universal law, does not endure the above criticisms. A Penn could not so will, for then he would be forgetting the inherent ambiguity of the moral situation, he would be denying the profound freedom of his neighbor, he would not be heeding the expressive as opposed to the deductive nature of the moral judgement, and he would be overlooking the complementarity of the two stances. Acting according to a universalizable maxim is not essential to being moral, and willing one's maxim to be a universal law is often positively immoral.

14. Quoted in Peter Brock's *Pacifism in the United States* (Princeton, N.J.: Princeton University Press, 1968), p. 86.