



Hare's Golden-Rule Argument: A Reply to Silverstein

Michael H. Robins

Mind, New Series, Vol. 83, No. 332 (Oct., 1974), 578-581.

Stable URL:

<http://links.jstor.org/sici?sici=0026-4423%28197410%292%3A83%3A332%3C578%3AHGAART%3E2.0.CO%3B2-7>

Mind is currently published by Oxford University 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/oup.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.

Hare's Golden-Rule Argument: A Reply to Silverstein

MICHAEL H. ROBINS

In *Freedom and Reason* Hare tells us that something does not count as a moral principle unless we are prepared to prescribe it for ourselves in a hypothetical case, in which we imagine ourselves in our victim's or adversary's shoes. This characterization of moral reasoning has frequently come under attack as being incoherent, most notably by C. C. W. Taylor in his review of *Freedom and Reason* (*Mind*, lxxiv (1965), 280-298). H. S. Silverstein, in his recent note on Hare (*Mind*, lxxxi (1972), 448-450), has contributed to this general line of criticism by adducing a 'special kind of case' in which 'the agent knows or believes something which is not known or believed by those affected by his proposed action and in which this fact itself is regarded by the agent as being relevant to his decision' (p. 448).

The case in question is the 'deathbed promise' in which someone makes a lying promise to a dying man, the intention being to make him 'happy'. The essential thing is that the promisee's happiness depends upon his not knowing, or ever discovering, that the promise is deceptive. To test his principles, the promisor is to put himself in the dying man's place; but this immediately reveals a difficulty concerning *what* the promisor, in his hypothetical role, is supposed to imagine. Silverstein writes: 'He cannot imagine himself, . . . , knowing or learning that the promise made to him is deceptive, for this falsifies a crucial feature of the actual case. Yet neither can he imagine himself not knowing or learning that the promise is deceptive, for then a crucial feature of the case would be left out; from his hypothetical point of view as a dying man the case would so far contain only a promise, not a deceptive promise, . . .' (p. 449). Silverstein concludes that such a case is 'morally Gödelian' and that the application of the golden-rule to it is incoherent: '[for] it is impossible to construct a hypothetical case which is not, with respect to the actual case, either false or incomplete' (*ibid.*).

In this note I wish to show (1) that the first horn of the dilemma is a bogus one, i.e. that it is false that he cannot, in the hypothetical case, conceive of himself knowing or learning that the promise is deceptive; and I shall want to maintain this even though the promisee's ignorance is a crucial feature in the actual case. I shall merely show that it is not a crucial feature of the hypothetical case.

(2) This leads to a general point concerning a common misunderstanding of the 'golden-rule' argument, attributable to just about everyone who has leveled the 'incoherence' criticism against Hare, namely the mistaken view that all of the relevant aspects of one's adversary in the actual case must be imagined, intact, in the hypothetical case. Silverstein's formulation expresses this in a rather pointed way: 'If A is considering

doing *x* which affects *B* in a certain way, *A* must imagine a hypothetical case which is *exactly similar to the actual case*, except that *A* plays the role of *B*.¹ (Italics mine.) What is essential, he claims, is that the action must affect *A* in the hypothetical case exactly as it affects *B* in the actual case (p. 448).

I wish to show that this formulation misconstrues the form and the purpose of the argument, that the role-shift is a much more limited one than Silverstein and others would lead us to believe. This can be shown by first considering a more common kind of case, which, according to Silverstein's formulation, would also be 'morally Gödelian'. I shall then try to show how Hare's distinction between 'ideals' and 'interests' intimates the form of the argument in such a way that all of these difficulties, including the 'deathbed promise', no longer arise.

Let us begin by considering a more common case in which two people have conflicting desires and ideals and in which this conflict must be viewed as a relevant feature of the case: let us suppose that a Zionist is venting his moral outrage at the recent acts of Arab terrorists. According to Silverstein's characterization, we would be invited to believe that for the Zionist to universalize his moral judgment by putting himself in the place of the terrorist, he would have to imagine himself to have both the Arab's 'inclinations' and 'ideals'. But this would indeed make the imaginative role-shift incoherent if its whole point is to be a challenge of the Zionist's ideals and principles, which, *ex hypothesi*, conflict with the Arab's ideals. For it would lead us to believe that to judge an act from the point of view of someone else's principles (which are contrary to our own) is somehow a test of the morality of your own principles. If, on the other hand, he were to imagine himself, in the hypothetical role, *not* believing in the Arab's principles, he would have 'falsified a crucial feature (of the Arab's situation) in the actual case' (for it is largely on account of such ideals that the guerrillas are led to such acts of desperation in the first place).

Now this appears to be a difficulty only because we have lost sight of the original *point* of the imaginative role-shift, which (stated one way) is to serve as a test of our professed moral views by determining whether they are held as a matter of principle, or are based upon our 'inclinations', *ad hoc* considerations, etc. Accordingly, it is supposed to challenge those views by asking us if we could still assent to them if we were to imagine ourselves in a situation, not with our own inclinations, but with those of our adversary. Accordingly the role-shift regarding the above case is supposed to take place only to the extent that the Zionist imagines that he has acquired the Arab's 'desires', but not his 'principles'. On the contrary, it implies (a) that the Zionist's *actual* moral beliefs must be brought to bear on the hypothetical case, and to that extent the latter is never exactly like the former; (b) similarly, the Arab's moral principles cannot be imagined as his own in the hypothetical case, however important they may be in the actual case.

It is important to see *why* the role-shift takes place only concerning desires, etc. and not principles (independently of its being incoherent if it included both). Let us assume that the agent must already be *aware* in the

actual case of all the relevant features of his adversary's situation (otherwise he will not know *what* he is supposed to imagine as his own). Now once we already know this, the only features of the case with regard to which there is any *further point* in imagining as our own is our opponent's 'desires', not his ideals. For it is one thing to *know* your opponent's desires, as belonging to him; it is another to appreciate their significance by imagining them as your own. But it is evident that the *cogency* of ideals, and the considerations on which they are based *qua* ideals, is not appreciated in this fashion, i.e. once we already *know* what our opponent's ideals are, they do not thereby become more persuasive by the additional belief act in imagining them as our own. What the role-shift of *persons* suggests, then, is that 'desires' are tied to a person's *identity* in a way that 'principles' are not; the *point* of the *role-shift* is only the limited task of finding out whether our professed principles are really based upon considerations independent of our 'identity' and its trappings; it is not supposed to perform the larger task of adjudicating (except in this way) between rival moral principles.¹

Now as for the deathbed promise, it would seem that the same considerations would establish a similar disparity between 'desires' (happiness) and 'cognitive states'. Thus I shall want to argue that when the agent puts himself in the dying man's place, he need do so only with regard to his 'happiness', not his 'ignorance'. Rather I shall show that he must bring to bear on the hypothetical case his own knowledge in the actual case (that it is a lying promise), just as he must his own moral principle. I emphasize his *own* knowledge (in *propria persona*) to forestall the objection: how can the *dying man* be happy if *he knows* that the promise is deceptive? For the agent is not *imagining* himself knowing, *as the dying man*, but rather that his knowledge is not one of the features supplanted in the role-shift, but, like his principles, remains part of his original equipment that he brings with him to assess only his imaginary happiness. Conversely the promisee's ignorance, though essential to his happiness, enters the picture, *not* as a state the agent *imagines as his own*, but as the ignorance of *somebody*, entailed by his actual knowledge, in whose context only the desire-shift takes place.

It might be thought (and this gets to the thrust of Silverstein's and Taylor's criticism) that there is something artificial, in fact 'schizophrenic', about shifting desires in this way without at the same time shifting cognitive states and ideals on which they may (logically) depend. But this only brings to light the fact that we appreciate the significance of desires vs. cognitive states in radically different ways—that we accord to our own desires a special status that we do not ordinarily accord to our cognition. For that reason the role-shift has a *point* only concerning the former. To see this, let us consider the same argument with regard to cognitive states as such that we considered with regard to principles: if the agent is already *aware* of his adversary's ignorance of a crucial feature of the actual case, what *further point* is there in *imagining himself not*

1 That is why this is a very weak account of moral reasoning unless it is buttressed with Hare's version of 'utilitarianism' (see *Freedom and Reason*, p. 123).

knowing, as there is a further point imagining himself happy? Unlike desires, the significance of other people's cognitive states is not diminished by knowing that they belong to other people.

Thus the golden-rule may appear 'schizophrenic' because it is but an application of the old fashion idea of splitting the personality into its 'appetitive' and 'cognitive' aspects; imagining oneself in the place of others thus drives home the point that only in the appetitive, egocentric aspect are moral principles 'no respecter of persons'. This is indeed but a small part of what is involved in justifying moral principles (see *supra*, p. 580, n. 1). It becomes 'incoherent' only if one ascribes to it more than this limited task; but since in that case it also becomes *pointless*, one can make the charge of incoherence only at the cost of knocking down a straw man.

BOWLING GREEN STATE UNIVERSITY