

Universalizability and R. M. Hare's Argument for Utilitarianism

Kirk Ludwig
Philosophy Department
University of Florida
Gainesville, FL 32611-8545
(352) 374-8236
kludwig@phil.ufl.edu

Abstract

Universalizability and R. M. Hare's Argument for Utilitarianism

This paper is concerned with R. M. Hare's argument from his account of the logical character of moral judgments to their universalizability, principally in *Freedom and Reason*, and then from their universalizability, so grounded, to act preference utilitarianism, principally in *Moral Thinking*. I argue that there are two so far unnoticed, but fatal, problems with Hare's argument. The first problem is that Hare's account of the logical character of moral judgments does not rationally require universalizability. The second problem is that the sort of universalizability Hare's arguments would lead to will not yield classical preference utilitarianism when combined with the rest of his argument, but, instead, to a quite counterintuitive position which requires summing over preferences of merely possible individuals.

Universalizability and R. M. Hare's Argument for Utilitarianism*

I

R. M. Hare argues famously, and ingeniously, for utilitarianism from his moral prescriptivism¹ and the requirements of rationality. The striking nature of this argument is brought out in this statement of Hare's about it:

... if we assumed a perfect command of logic and of the facts, they would constrain so severely the moral evaluations that we can make, that in practice we would be bound all to agree to the same ones. ... the freedom which we have as moral thinkers is a freedom to reason, i.e., to make rational moral evaluations; and the rules of this reasoning, which are determined by the concepts occurring in the questions we are answering, bring it about that, over the most important part of morality, we shall, if we are rational, exercise our freedom in only one way. (MT, p. 7)

If Hare is right, we can settle moral disputes in principle on the basis of nothing but clear-headed agreement on the moral language (philosophical logic) and the actual facts. This takes moral

* I wish to thank David Copp, John Biro, R.G. Frey, Alastair Norcross, Marina Oshana, Crystal Thorpe, Jon Tresan, the audience at a presentation of an earlier version of this paper at the 2004 CD APA, and the audience at an presentation of a longer paper, "Hare's Argument for Utilitarianism," from which this material originates, at the June 2003 conference on Ethics at Bled, Slovenia, for valuable discussion.

¹ I will use 'prescriptivism' specifically for Hare's account of the dual nature of moral judgments as having a descriptive meaning and a prescriptive meaning. See section II for details.

argument out of the realm of appeal to intuitions and grounds it on linguistic and non-linguistic facts.

Hare's argument rests on the claim that the dual nature of moral judgments² as prescriptive and descriptive requires that (1) if we make any moral judgments, we be willing to universalize them, that is, to make them no matter what our position might be in the circumstance in which we make them (FR, §§2.2-2.8). He combines this with the claim that (2) the right moral judgment to make is that made under conditions of full information about the relevant consequences of doing what one prescribes in making the moral judgment (MT, pp. 86-90). Then he argues that (3) to have full information about the relevant consequences, one must know what it would be like for anyone who would be affected by doing what one prescribes be done in the circumstances (MT, pp. 91-2). Further, (4) to know what it would be like for anyone who would be affected by doing what one prescribes be done in the circumstances, Hare argues, one must have certain conditional preferences corresponding to those of others affected in the circumstances. Specifically, if someone has a preference that p with strength s , then one must have a preference with strength s that if one were in his circumstances with his preferences, then p (MT, §§5.3-5.4). This, combined with the requirement that we universalize our prescriptions, that is, be willing to make them from any point of view in the circumstances, is supposed to lead us to "treat *other people's* prescriptions ... as if they were our own" (MT, p. 17), i.e., to lead to a

² Throughout I will use 'moral judgment', following Hare, as a term of art to refer to certain speech acts performed using certain sentences containing moral terms. I won't here attempt to delineate the relevant class of sentences. They will include such sentences as 'Gluttony is a sin', 'What you did was wrong', 'You shouldn't have lied', 'Helping her was the right thing to do', 'Charity is a virtue', and so on. The picture becomes more complicated when we move to molecular sentences such as 'If you said that, then you did something wrong', which, on Hare's view, would involve a conditional prescription. My discussion throughout will be restricted to atomic sentences.

judgment that in effect sums the preferences of everyone affected in the circumstances in a way that yields as the right judgment the one which maximizes preference satisfaction over all affected individuals.

In this paper, my concern is with Hare's argument from prescriptivism³ to universalizability, and then from universalizability to utilitarianism. I will argue that there are two so far unnoticed, but fatal, flaws in Hare's main argument. The first is that the dual nature of moral judgments as descriptive and prescriptive does not, as Hare argues in *Freedom and Reason*, rationally require universalizability. The second is that the sort of universalizability Hare's argument would lead to will not yield classical utilitarianism when combined with the rest of his argument, but, rather, to a quite counterintuitive position. It appears otherwise only because Hare uses in his argument from universalizability to utilitarianism a restricted version of the principle which his argument in *Freedom and Reason* yields. In arguing that Hare has not established that universalizability follows from his view about the content of moral judgments, and that the form of universalizability he argues for would not yield the result he aims for with the rest of his argument, I am not arguing that moral judgments are not universalizable. I am arguing rather that the resources that Hare brings to bear to establish this from the standpoint of his account of the dual nature of moral judgments as descriptive and prescriptive is not adequate for this. If this is right, then even if moral judgments are universalizable, and in just the way that

³ Hare describes his position as 'universal prescriptivism' (FR, §2.5), which builds in the universalizability of moral judgments. One might think therefore that it is trivial that prescriptivism entails universalizability. However, universal prescriptivism is "a combination ... of universalism ... and prescriptivism" (FR, p. 16). And the universalizability of moral judgments is argued for on the basis of other features of them, and, in FR, specifically, as discussed in II below, their having descriptive meaning.

Hare needs, it is not clear that Hare can help himself to that fact, absent an account of how that can be so compatibly with his basic account of the nature of moral judgment.

II

Hare's most extended argument for the universalizability of moral judgments is given in *Freedom and Reason* (§§2.2ff). My primary concern will be with this argument, which Hare cites in his fullest defense of the thesis that preference utilitarianism follows from prescriptivism in *Moral Thinking* (pp. 115, 219). I will return later to the suggestion of a different argument to be found in *The Language of Morals* (§8.2).

Hare's prescriptivism holds that a moral judgment has two independent components to its meaning, one prescriptive, and one descriptive (LM, §§5-7; FR, part I; OWF, §2; P). In saying that such and such is evil, I am attributing to it a property, and making a statement about it. But I am also, according to Hare, making a third person prescription, which can be expressed by an utterance of 'Let it be the case that such and such not be done'. In Hare's words, "[w]e say something prescriptive if and only if, for some act *A*, some situation *S* and some person *P*, if *P* were to assent (orally) to what we say, and not, in *S*, do *A*, he logically must be assenting insincerely" (MT p. 21; LM §§1.7, 2.2; SE §1.6). The effect of the inclusion of a moral predicate in a sentence is to induce a dual reading of its content. As Hare puts it at one point, "moral statements are a hybrid, sharing some characteristics of both pure descriptions and of pure prescriptions" (OE, p. 95). When someone sincerely utters a sentence such as [1], he performs two speech acts, which can be represented with the contents of sincere utterances of [2] and [3], where 'is *F*' stands in for a purely descriptive predicate that expresses the property which 'is evil' does but without the prescriptive meaning Hare assigns to the latter (SE 1.7). This solves the problem of why there seems to be a deep connection between making sincere

moral judgments and motivation: the connection is made through the sincerity condition on the prescriptive speech act.

- [1] Torturing cats for fun is evil
- [2] Torturing cats for fun is *F*.
- [3] Let it not be the case that cats are tortured for fun.⁴

The universalizability of moral judgments is supposed to follow from this account of the content of moral judgments in a straightforward way (FR, 2.2-2.8). The argument is sketched in the following passage:

We must now notice the connexion between the fact that some judgements are descriptive and another feature which it has become the custom to call, when we are speaking of moral judgements, universalizability. It is important to emphasize that moral judgements share this feature with descriptive judgements, although the differences between them in other respects, are, as we shall see, sufficient to make it misleading to say that moral judgements are descriptive. Nevertheless, in so far as moral judgements do have descriptive meaning, in addition to the other kind of meaning which they have, they share this characteristic, which is common to all judgements which carry descriptive meaning. (p. 10)

⁴ In conversation in the early 1990s Hare accepted this rendering of the prescriptive force of such a moral judgment. I will suggest a different way of construing the prescriptive force in section V, which, however, would have [3] as a “consequence” given [1].

The argument, in short, goes as follows. Descriptive judgments are universalizable trivially.⁵

Moral judgments involve a descriptive component. Therefore, moral judgments are universalizable as well. What exactly does this come to though?

If I assert sincerely that a thing is red, then I am committed to anything which is relevantly similar to that also being red, where relevant similarity has to do with the features of the thing on the basis of which I sincerely asserted it was red. I am committed in particular to its being red independently of what position I may have in the circumstances. If I judge a rose on the table to be red because of those visual features it has in good light, then I am committed to anything like that being red, irrespective of whether I occupy a different position, or any position, in such a qualitatively similar situation—for the numerical identity of the various things present in the circumstances is irrelevant to the features of the thing on the basis of which it is called red.⁶

On Hare's account, moral terms have descriptive as well prescriptive meaning. So when I say, for example, that excessive pride is evil, I attribute to this state a certain property, call it

⁵ Hare reserves the term 'descriptive judgment' for judgments whose whole content is descriptive (FR, p. 26)

⁶ See note 6, the quotation from FR, p. 11. The universalizability of descriptive judgments applies only to predicates which are purely descriptive and not to predicates whose application conditions include reference to particular individuals, such as 'x is talking to Donald Trump'. If I am Donald Trump, then clearly whether I occupy a position in the situation makes a difference to the applicability of the predicate. On the other hand, no such difficulty arises for 'x is talking to y'. What about perspectival relational predicates, however, such as 'x is to the left of y'. Judging correctly of two objects *x* and *y* that *x* is to the left of *y* depends on the perspective from which the judgment is made. In the same qualitative situation, from one perspective the claim may be true, but from another false. For perspectival predicates, the requirement is that one be committed to making the same judgment in relevantly qualitatively similar situations from the same perspective.

being evil.⁷ I thereby commit myself to saying of every instance of excessive pride that it has the property of being evil. If I say that Job's pride was evil, then I commit myself to anyone in a state relevantly similar to Job's also having a state with the property of being evil, where relevant similarity has to do with those features of Job's pride on the basis of which I said that it was evil, no matter whether it is Job or someone else, even myself.

Judging that a state or act is evil is also, on Hare's view, making a prescription—let it not be or be done. Hare argues that in virtue of being committed to universalizing the descriptive component of the judgment, we are likewise committed to universalizing its prescriptive component.⁸ That is, we are committed to universalizing the moral judgment *per se* in virtue of

⁷ I will continue to underline the phrase 'being evil' to emphasize that the intention is to use 'evil' as it appears in this phrase in a purely descriptive sense. On the intended use, one can say of something that it has the property of being evil sincerely without making any prescription. I have heard it denied that Hare thinks that 'is evil' and other evaluative terms are used to ascribed properties. But Hare responds explicitly to this misunderstanding of his position (OWF, p. 68): "To explain the distinction more fully: the descriptive meaning is simply the truth-conditions of statements containing moral words, i.e., the criteria for their application. Anybody who acknowledges that they have such truth-conditions obviously cannot be accused of thinking that moral statements are not true or false. So if there are non-cognitivists, I am not one of them. Those who think I am should quote chapter and verse." In MT he says (p. 216): "Moral judgements are about properties which actions have (e.g. the property of resulting in someone's death). The properties of the actions are the reasons for the judgement; that is another way of stating the requirement of universalizability, which is generated by the logical properties of the moral concepts."

⁸ Hare says, "It is impossible consistently to maintain that moral judgements are descriptive, and that they are not universalizable" (FR, p. 16). For, where *u* is the thesis that moral judgments are universalizable, and *p* the thesis that they are prescriptive, and *d*' the thesis that "moral judgements ... do have descriptive meaning" (FR, p. 17), Hare says: "As we have seen, *d*' entails *u*." Again: "The truth in naturalism is that moral terms do indeed have descriptive meaning. It is not the only element in their meaning, and it is therefore misleading to refer to it, as do the naturalists, as *the* meaning of a moral term; but *in virtue of possessing this descriptive meaning moral*

being committed to universalizing its descriptive component, and so are committed to universalizing in the same sense the prescriptive component. Thus, I am committed to prescribing what I do in any relevantly similar circumstances, independently of the numerical identity of individuals in the situation. If I judge that some act of another is evil, then if I do something relevantly similar, I am committed to judging of my act that it is likewise evil, and so I am committed to making the prescription in my own case as well. I take Hare to intend, more precisely, that making a moral judgment entails a commitment to making the same judgment about all cases identical with respect to those universal properties *on the basis of which one makes the moral judgment*,⁹ with the *same* strength of preference. This is a very strong

judgements are universalizable, and naturalism has the merit of implying this” (FR, p. 21; emphasis added). And: “Any judgement which has descriptive meaning must be universalizable, because the descriptive meaning-rules which determine this meaning are universal rules” (FR, p. 39).

⁹ “Moral judgements are,” Hare claims, “universalizable in only one sense, namely that they entail identical judgments about all cases identical in their universal properties” (MT, p. 108). But he also makes clear elsewhere that that this is a consequence of something stronger: “by calling a judgment universalizable I mean only that it logically commits the speaker to making a similar judgment about anything which is either exactly like the subject of the original judgment or like it in the relevant respects. The relevant respects are those which formed the grounds of the original judgment” (FR, p. 140). And the grounds for the original judgment were the grounds for its holding its descriptive component. “If a person says that a thing is red, he is committed to the view that anything which was like it in the relevant respects would likewise be red. The relevant respects are those which, he thought, entitled him to call the first thing red: its red colour” (FR, p. 11). Hare affirms this as the ground for the universalizability of moral judgments, for he says: “a philosopher who rejects universalizability is committed to the view that moral judgements have no descriptive meaning at all” (FR, p. 17). See also the quotation in the previous note from FR, p. 21.

requirement, and Hare in fact focuses on just those cases identical with respect to *all universal properties*, not just those on the basis of which one makes the moral judgment.¹⁰

III

Hare's argument in *Freedom and Reason* from the commitment to the universalizability of the descriptive component of a moral judgment to the universalizability of its prescriptive component has a deceptive simplicity. But the argument could succeed only on an assumption which Hare is committed to denying. If universalizability cannot be secured as a rational commitment on the basis of Hare's prescriptivism, then the argument for utilitarianism from prescriptivism, which depends crucially on this, will fail.

The commitment to the universalizability of moral judgments is grounded in the universalizability of their descriptive component. Clearly one is committed to the descriptive component being universalizable. Why think one must express this commitment to universalizing the descriptive component with the dual package of descriptive and prescriptive content? There would be a reason to think this if the prescription were grounded in the descriptive content, in the sense of being determined and explained by it. But this is something that Hare denies (MT, p. 16; SE §3.5). This denial plays an important explanatory role in his

¹⁰ Hare's view of the relation between the descriptive and prescriptive component of moral judgments is complex, for he holds that the prescriptive component, at least in the case of "thin" moral predicates like 'ought', 'right', and 'wrong', is in a sense dominant (see FR, p. 24; P). This allows the descriptive component to shift while we think of ourselves as still using the moral terms in their primary sense. This is what is supposed to allow for the possibility of genuine effective moral criticism of prevailing norms (FR, p. 25). But it is crucial to Hare's view of this as a rational process that moral prescriptions be universalizable. For that is what provides the leverage for getting people to change their views about what is right and wrong. But that rests in turn on arguing from the universalizability of the descriptive component to the universalizability of the prescriptive component.

ethical theory. It explains why in different groups, periods, and cultures, the moral terms are applied in ways that suggest they have different, incompatible extensions, as in the case of Hare's example of the missionary and the cannibals (LM §9.2). It is because they have different descriptive meanings. It is because they are all also alike in having a prescriptive meaning that they count as moral terms, and it is this that is supposed to provide for the possibility of rational discourse and moral persuasion despite what look like basic disputes about what to count as falling under 'good' or 'right' (see LM, §9.2; MT §4.2; RAD; SE §4.3). This requires the independence of their descriptive and prescriptive components. For if the descriptive meaning determined their prescriptive meaning, one could not come to attach to the same descriptive meaning a different prescriptive meaning; but this is what is involved, on Hare's account, in changing one's moral views.

But if the commitment to universalizing the prescriptive component of a moral judgment is not grounded in its specific descriptive content, the commitment to universalizing the descriptive component of a moral judgment is not itself, and does not require, a rational commitment to universalizing the prescriptive component. Suppose that in some circumstance I want to say that a contemplated act of mine has the property of being evil, but I don't wish to prescribe that I not do it. Suppose further that I have on another occasion judged that some relevantly similar act of another was evil. What am I rationally required to do in virtue of that earlier judgment? The descriptive component of it commits me to judging likewise about my contemplated act, but likewise specifically with respect to its falling under the descriptive content of the judgment. What rationality requires of me now (in light of my earlier judgment and present desire not to prescribe that I not do what I contemplate) is not that I judge my contemplated act to be evil, but only that I not express the commitment that it has the property of

being evil by using the English term ‘evil’. I am rationally required, if I haven’t changed my mind about what features are relevant, holding that this act has that property on the basis of which I judged the other’s act to be evil. But there is no rational requirement on me to likewise condemn it because I condemned that similar act by another. The fact that one way of making the same descriptive judgment would require, if I made it sincerely, that I condemn the act, only means that if I do not wish to condemn it, I not attribute the property in that way. If our vocabulary limits the ways we can express these things, then we may feel a need to universalize our prescriptions. But this would involve a kind of linguistic sleight of hand. It might have a kind of subliminal psychological force on us, but it would not have rational force on us. And, as Hare reminds us, we could always resort to the use of inverted commas to say what we want without the usual prescriptive force (LM, p. 124-5, FR, p. 189-90; MT, p. 24). It is instructive to think here of the ways in which derogatory attitudes are propagated by the use of negative evaluative words like ‘fag’ and ‘kike’. It is clear from this case that the commitment to universalizing the descriptive component of these words does not carry any rational commitment to universalizing their expressive component. How could it? Thus, even granting prescriptivism, and that we are committed to using the moral language, we are not further rationally committed to universalizing our *prescriptions*—at least, not on the grounds which Hare provides in *Freedom and Reason*.

IV

I have so far concentrated on the argument that Hare gives in *Freedom and Reason* (§2.2 ff) from the dual descriptive-prescriptive character of moral judgments to their universalizability. If I am

right, this argument is unsuccessful. The argument in the form I have discussed first clearly appears in *Freedom and Reason*. In *The Language of Morals*, Hare also endorses the universalizability of moral judgments. He offers grounds which at least superficially appear to provide a different basis of universalizability. The way Hare explains why moral judgments are universalizable in *The Language of Morals* seems largely to drop out of sight in later work. Still, if it provides an alternative route to universal prescriptivism from that advanced in *Freedom and Reason*, then even if the argument in *Freedom and Reason* is unsuccessful, Hare's full argument for utilitarianism from universal prescriptivism may still succeed. It will be worthwhile, therefore, to examine the argumentation in *The Language of Morals* closely.

The argument originally appears in the context of non-moral evaluative judgments, such as the judgment that that is a good car (LM, §§8.2-3); Hare then argues it can be extended to moral terms (LM, c. 9). The gist of the idea is conveyed in the following passage:

Now since it is the purpose of the word 'good' and other value-words to be used for teaching standards, their logic is in accord with this purpose. We are therefore in a position at last to explain the feature of the word 'good' which I pointed out at the beginning of this investigation. The reason why I cannot apply the word 'good' to one picture, if I refuse to apply it to another picture which I agree to be in all other respects exactly similar, is that by doing this I should be defeating the purpose for which the word is designed. I should be commending one object, and so purporting to teach my hearers one standard, while in the same breath refusing to commend a similar object, and so undoing the lesson just imparted. (p. 134)

In the terminology of *The Language of Morals*, the standards associated with a value-word give its descriptive meaning (in the following ‘value-word’ is to be understood as restricted to terms that have associated standards). Hare also speaks of the standards as providing the reason for calling a thing good. Later, he cites the giving of reasons for a judgment as what grounds a supervenience claim about it, i.e., that the same claim is to be made in all situations alike with respect to the ground for it (SUP, p. 67).

The argument here goes as follows. The function of value-words is to teach standards. The logic of value-words must “accord with this purpose,” that is, it must be compatible with this function of value-words. If value judgments like ‘this is a good car’ were not universalizable, that is, if the same judgment were not called for as a matter of its logic in all circumstances qualitatively exactly alike, then value-words could not be used to teach standards. Therefore, value judgments like ‘this is a good car’ are universalizable.

This argument, I believe, is no more successful than the argument we reviewed in the previous section. First, it is not clear that what is required for the logic of value-words to be compatible with value-words being used to teach standards is that value judgments like ‘this is a good car’, or any component of them, be universalizable, as opposed to simply not ruling out using them consistently. If one can use value-words consistently in application to things, that is, on the basis of some features of what they are applied to, then this would be enough for them to be used to teach standards. Second, the conclusion of the argument is supposed to be that value judgments, ultimately moral judgments, are universalizable in the sense that making one in certain circumstances rationally requires making the same one, *ceteris paribus*, in relevantly similar circumstances, specifically, circumstances to which the same standards apply (in which the features on the basis of which the original judgment was made are present). But even if the

logic of value-words does involve an element which is by its nature universalizable, which would surely suffice for them to be used to teach standards, this will not be enough to support the conclusion that Hare needs. For as we saw in the previous section of this paper, Hare's view that moral judgments have a descriptive and prescriptive meaning entails that they have a component of meaning which is universalizable, trivially. But this does not entail that making a moral judgment in one case rationally requires making the same moral judgment in all relevantly similar cases.

To drive the point home, let us revert to the example discussed in the previous section of derogatory words. These words undoubtedly do have a function in the communities in which they are used, and, indeed, probably a number of functions, social, and linguistic. One of their functions undoubtedly is to teach certain standards of behavior, in the sense of providing guidance for behavior. The mechanism for this involves the exertion of peer pressure or social authority to conform to the attitudes which are understood to be conventionally expressed by the use of such words in application to objects fitting their descriptive content. It does not follow from this, however, that judgments using such words are universalizable in the sense Hare has in mind: that is, in virtue of having made such a judgment on one occasion on the basis of the descriptive content (the standard for its use), one is not rationally compelled to make the same judgment *in toto* in the future when the descriptive content applies to someone. Similarly, that the moral terms are used in judgments which are intended to teach standards does not entail that it is a part of their logic that they carry a rational commitment to making the same judgment in every situation qualitatively alike in the relevant respects.

A final point that should be made is that this argument, like the one in *Freedom and Reason*, yields, if successful, a principle that requires that one make the same moral judgment in

all circumstances alike with respect to the relevant standards (alike in fitting the descriptive meaning) of the moral terms. This is stronger than the requirement that one make the same judgment in all circumstances exactly similar. The latter is entailed by the former, though not vice versa.

V

Why is it not open to Hare to respond to these criticisms by holding that the prescriptive component of a moral judgment is already universalized, in the sense that it is a universal prescription?

To answer this question, we must take a closer look at the prescription to be associated with a moral judgment, and how we will understand a universal prescription. Since Hare does not say enough about his views for one to know precisely how he would spell out the prescriptive component, the following is an attempt on Hare's behalf to construct an account that makes the prescription associated with a moral term universal.

Moral judgments may be universal or particular in the sense that they may be about all things (perhaps of a certain kind) or about particular things. Thus, contrast [1], repeated here, with [4], where Tom is a particular cat.

[1] Torturing cats for fun is evil.

[4] Sam's torturing of Tom for fun is evil.

[1] is a universal moral judgments and [4] is a singular moral judgment. [4] follows from [1] together with the premise that Tom is a cat. ‘is evil’ appears in both [1] and [4]. This is to carry the prescriptive meaning. It must be the same meaning in each. Hare is concerned to say that all moral judgments are universalizable, including judgments such as [4]. In all circumstances alike with respect to the (purely qualitative) properties of Sam’s torturing Tom for fun I am committed to making the same sort of moral judgment. So if things are exactly the same except that it is Bill torturing Tabby rather than Sam Tom, one is supposed to be rationally committed to judging that Bill’s torturing Tabby is evil. If we seek the universal element not in the descriptive component, but in the prescriptive component, then even in singular judgments like [4] we must find the prescriptive component to be universal in form. This can be supplied by taking the prescription to be about the property attributed in accordance with the descriptive meaning of ‘is evil’.

Let, then, ‘ $M(x, t)$ ’ be a term expressing the descriptive meaning of a moral predicate with specifically prescriptive meaning, where ‘ t ’ is a variable restricted to times and time intervals. Let ‘ $!^+$ ’/‘ $!^-$ ’ represent a mood marker for a positive/negative third person imperative sentence (‘let it be the case that’/‘let it not be the case that’). Let me henceforth focus on positive third person imperatives. A universal prescription will bind the variables in ‘ $M(x, t)$ ’ with universal quantifiers, but there will be two readings depending on whether ‘ $!^+$ ’ takes wide or narrow scope with respect to the quantifiers:

$$(W) \quad !^+[(\forall x)(\forall t)M(x, t)]$$

$$(N) \quad (\forall x)(\forall t)!^+[M(x, t)]$$

In addition, there is the question of how to handle the semantics of the mood setter itself. For a simple second person imperative like ‘Open the door’, we can represent the semantic function of the mood setter by representing utterances of the sentence as have a distinctive sort of compliance condition, constructed from the declarative core of the sentence. Where ‘IMP(ϕ)’ represents an imperatival form of a future directed declarative ϕ (see SN):

For any speaker s , time t , IMP(‘You will open the door’) is complied with relative to s at t iff $\text{ref}(\text{‘you’}, s, t)$ makes it the case that ‘You will open the door’ is true relative to s at t .

For a third person, rather than second person, imperative the speech act performed using it is not addressed to a particular audience (we can vary the strength of the operator as well, of course, but this parameter will not affect any of the reasoning that follows, so I will leave it as ‘make it the case that’). In the case of the wide scope reading (W), we would appear to have two basic options. First, the compliance conditions require *someone* (or someone meeting certain requirements) to make it the case that the declarative core is true. Second, that the compliance conditions require that *everyone* (or everyone meeting certain requirements) make it the case that the declarative core is true.

In the case of the wide scope reading (W), it seems implausible to think that a prescription that might be associated with a moral judgment would require either everyone or someone to make it the case that everyone at all times promotes something. This is implausible because it is so clearly beyond the power of anyone. On the other hand, it also seems implausible that it should be directed at any particular audience.

In the case of the narrow scope reading (N), the universal quantifier interpretation of the scope marker likewise seems inappropriate. It is not plausible that morality requires, with respect to each relevant thing and time, that each person should promote it. This is clearly not possible. Therefore, the existential reading is to be preferred. Thus, we can represent the content of a universal prescription of the sort to be associated with a moral judgment as:

For every x, t , let someone make it the case that $M(x, t)$.

Then, if it turns out that one is oneself the person best placed, by some action open to one, to make it the case for some value of ' x ' and ' t ', that ' $M(x, t)$ ' obtains, then one will be motivated oneself to do it.

Suppose then that moral judgments which involve prescriptions involve universal prescriptions of this sort. The question is whether this would be enough for Hare's purposes. The answer, I think, is no. The difficulty is this. If we are able to argue that moral judgments are universalizable in the sense that one is rationally compelled to make the same moral judgment in all situations alike with respect to their non-moral features, and more specifically alike with respect to the descriptive meaning of the moral terms, then in simply recognizing that something has the same properties, one is called on to make the same prescription. This is a very powerful requirement. For it means that if one uses the moral language at all, one is required by rationality consistently to make the same moral prescriptions in the same circumstances (barring shift in the descriptive meaning of the moral terms). This plays an important role in Hare's argument to preference utilitarianism. For it is this that commits one in particular circumstances

to consider how one would judge in all circumstances alike with respect to the features on the basis of which one makes the judgments, regardless of one's own position.

However, granting that the prescriptions involved in moral judgments are universal in form, it does not follow (from this alone) that one is rationally committed to making the same prescription in all cases similar with respect to the properties ascribed by the descriptive meaning of the moral judgment on the basis of their being similar. This would be rationally required only if my sentiments did not differ. But if they did, the fact that the circumstances were the same would not require me to make the same judgment. And the fact that I now have certain sentiments and am willing to make a certain universal prescription does not rationally require my sentiments to be the same in circumstances relevantly similar. It is true that there can be an inconsistency between two utterances of universal imperatives if their compliance conditions conflict. So if I judge on one occasion, 'For any x, t , let someone make it the case that $M(x, t)$ ', and later judge, 'For any x, t , let no one make it the case that it is not the case that $M(x, t)$ ', the two judgments are in conflict because they cannot be both satisfied. Yet there is nothing irrational about making the one judgment at one time and the other at a different time. The beauty of grounding universalizability of moral judgments in the descriptive meaning of moral judgments was that the satisfaction of the descriptive meaning was supposed to bind the speaker to making the same prescription on recognition of the circumstances being relevantly similar. But the prescriptions having a universal form cannot bind one to make the same prescription in all relevantly similar situations. Nor, for reasons we have discussed, can an associated descriptive content, if it does not determine the prescription. In particular, that if I make now in these circumstances a certain universal prescription, my quantifiers range over actual individual in the actual world only. The prescription I make carries no rational commitment to making the

same prescription were I to occupy a different position in the current circumstances, for that would be a counterfactual circumstance. Without that rational commitment, Hare's argument for utilitarianism cannot go through.

VI

Grant for the sake of argument that moral judgments are universalizable *tout court*, in the sense required by Hare's argument either in *Freedom and Reason* or *The Language of Morals*—that is, that if one makes a moral judgment one is rationally committed to making the same *moral* judgment in *all relevantly similar* circumstances. Does this yield, with Hare's other assumptions, preference utilitarianism? No. But to see why we must review the basic structure of Hare's argument.

The argument rests on three claims (MT, ch. 5; SEM, pp. 183-190). The first is that the correct moral judgment is the one we would make on the basis of knowledge of all relevant features of the situation, abstracting away from limitations on our reasoning powers.¹¹ The second is that the relevant features of the situation include what it would be like to be anyone who might be affected by what we prescribe. Hare importantly connects this with the universalizability of moral judgments. For we must be prepared to make the same judgment were we in each other affected person's place.¹² The third is that to know what it is like to be

¹¹ Hare does not think one must be committed to using the moral language (FR, p. 100; MT, 10.7-10.8). He only argues that if one is, then one is committed to making judgments which are the judgments which would be endorsed by preference utilitarianism.

¹² "It is irrational," Hare argues, "to prescribe anything without regard to what, concretely, we are prescribing; and this involves cognizance of ... what its execution in this concrete situation would entail. ... When the prescription is

someone else, one must acquire certain conditional preferences mirroring his preferences. For each preference for p of strength s which someone S has in the circumstance, one must acquire a preference of strength s that if one were in S 's position with S 's preferences, then p . Thus, requirement of full information about relevant consequences is supposed to require that one acquire preferences corresponding (in the sense just indicated) to each preference which anyone in the circumstances has. This, together with the requirement that one be willing to make the same prescription no matter what one's position in the circumstances, is supposed to lead one to sum over all of the conditional preferences one has acquired in deciding which moral judgments are acceptable, in effect choosing that one which maximizes preferences satisfaction over all affected individuals.¹³

Questions can be raised at a number of places in this argument. But my present concern lies specifically with the role played by universalizability, as Hare argues for it. In his discussion of the implications of universalizing a judgment, in the light of the full information about relevant consequences requirement, Hare focuses on what knowledge we would have to have about the effects on others in all *exactly* similar cases, that is, those that are the same with respect to *all* their universal properties. We are to know what it would be like to be each of the actual persons who might be affected, because we are committed to prescribing what we do were we to be in each of their situations, construing this as having their experiences, given their preferences.

However, universalizability, given *how* Hare argues for it, requires us also to judge the same not just in all situations that are alike with respect to all their universal properties, but in all

... universalizable, ... we are required to satisfy ourselves that we can accept the *universal* application of the prescription; and this includes its application were we in the other's position. So the facts we need to be cognizant of will include facts about his position as it affects him with his preferences" (MT, p. 89).

¹³ See HAO for a reconstruction of the details of this argument.

situations alike with respect to *the properties on the basis of which one makes the prescription*.

This will be a subset of all the universal properties instantiated in the circumstances in which one is considering making a judgment. One is then committed to making the same judgment no matter how circumstances vary compatibly with those relevant universal properties being instantiated.¹⁴

In being rationally committed to making the same judgment in each such case compatible with the instantiation of the universal properties on the basis of which one makes the judgment, one is rationally committed, according to Hare, to getting all relevant information about them. This includes what it would be like to be those people in each of the possible cases compatible the relevant universal properties being instantiated. That is, one must know what it would be like to be anyone possibly affected in any possible situation alike with respect to the relevant features. This requires in turn, according to Hare, acquiring conditional preferences corresponding to the preferences of people who would be affected in each of those possible cases. Thus, if there are few restrictions imposed by the descriptive content of a moral judgment on the preferences of people in circumstances compatible with it, one will have to acquire a very

¹⁴ At one point in *Moral Thinking* Hare notes that his argument in *Freedom and Reason* commits him to something stronger than just that one is committed to making the same judgment in when confronted with a thing “similar in the relevant respects” (p. 62). But he goes on to say, “... our present argument has no need of a definition of universalizability in terms of relevant similarity. In this book, we shall be appealing, in our account of critical thinking, only to exact similarity, and shall not need, therefore, to say, before the argument begins, what is and what is not relevant” (pp. 62-3). Hare seems not to notice that if the actual requirement of universalizability he argues for is stronger, it is the one which is operative in critical thinking about moral judgment. If critical thinking requires one to take into account *p*, and *p* entails *q*, then critical thinking requires one to take into account *q*; but it does not then follow that it does not require one to take into account *p*.

great many corresponding preferences in addition to those Hare identifies to be in a position to make the correct judgment about the actual situation.¹⁵

An immediate consequence of this is that, even if the rest of Hare's argument goes through, the result will not be an account of correct moral judgment which aligns with that given by preference utilitarianism. Preference utilitarianism takes into account only the preferences of actual people who would be affected.¹⁶ But universalizability, given what this comes to on Hare's account, and his argument that one must know what it would be like to be any person affected in a relevantly similar case, will require taking into account how people would feel who had preferences very different from those that any actual person has, or ever will have. The preferences of those actually affected have no primacy and do not call for special consideration in the calculation. Consequentialism is concerned with the consequences our actions actually have. Making merely possible preferences relevant will in effect make moral judgments, if we follow out Hare's reasoning rigorously, sensitive to consequences of what we would be doing in other possible worlds.

Moreover, the result, which Hare does not intend, is plainly counterintuitive. Suppose some action I contemplate will not frustrate the preferences of anyone who may actually be

¹⁵ One consequence of this is that it will make the problem of figuring out what the correct moral judgment is vastly more difficult, and there may be some worries about whether it makes sense to suppose one person could acquire all the preferences this would require, not because of limitations of intellect, but because it looks like it would likely require one person to have distinct preferences identical in content but different in strength, which is impossible.

¹⁶ We may allow also that it is relevant what the preferences would be of people who would come into existence in the future, given certain actions now. But those required by universalizability of the sort Hare's argument requires would not be restricted to actual people or those who would come into existence in the future, given certain contemplated actions now.

affected by it. Nonetheless, it may be morally proscribed, if we follow out the actual consequences of Hare's argument, because I must be willing to make the same judgment in circumstances alike in respect of the descriptive basis but in which it would have overall a seriously negative affect on the sum of preference satisfaction—either because preferences are different, or the results of the contemplated action are different, or both. Intuitively, these facts should not be relevant, given that as things are, what I contemplate doing will not be harmful.¹⁷

The difficulty arises only if the descriptive properties associated with moral terms allow for some leeway in what preferences people in the circumstances have. Therefore, one might try to defend Hare by arguing that the descriptive content associated with moral terms will in fact so restrict what preferences are compatible with the instantiation of the relevant properties that either only the actual preferences are compatible, or what additional ones are compatible, if taken into account, will yield the same judgment as taking into account only the preferences of actual people who would be affected.

It seems unlikely that there is a plausible defense along these lines. Certainly, to the extent to which Hare makes suggestions about the descriptive meaning attached to moral terms,

¹⁷ One way to bring out the problem is to note how it exacerbates Hare's difficulties with the fanatic, that is, someone who has strong enough preferences for intuitively bad outcomes that his preferences end up carrying the day in an impartial summing over all relevant parties (FR, pp. 105-6; MT, pp. 170-82). Hare agrees the moral conclusions we would make through critical thinking, if there were fanatics, would be counterintuitive. But he argues, though they are possible, they are not actual. So our actual intuitions, which express what would be the results of critical reflection about what principles it is best to act in accordance with given our general circumstances and our limitations in acquiring information and thinking through consequences, should yield the verdict that moral judgments influenced by fanatics are not correct. But if we should take into account non-actual preferences, it is likely we will have to take into account those of merely possible fanatics as well.

it would not restrict very much how we filled in the picture about the people who might be affected by an action undertaken in context in which we judged something to be good or bad, or right or wrong.¹⁸ But even if one could mount a defense along these lines, it would already have a cost. The reason Hare does not spend much time on the exact descriptive meaning of moral terms is that it is not supposed to matter. His argument is supposed to be insensitive to this. It is supposed to yield preference utilitarianism no matter what the descriptive content initially associated with a moral term. If the reasoning above is correct, then it follows that this is not correct, and that the right result would indeed depend crucially, if it could be achieved at all, on the descriptive content of the moral terms.

What descriptive content could be associated with a moral term that would restrict its application in such a way as to guarantee that no preferences other than the actual ones were to be taken into account? One could require that its descriptive content fix all the universal properties of what it is applied to. But this is untenable because we apply moral terms across circumstances which differ in their universal properties. For the same reason we cannot require that it fix all the actual preferences in a given context of use, for we apply the same moral terms across circumstances in which preferences vary. Nor would it be plausible to argue that the descriptive content should be restricted in some way that ruled out any but actual circumstances,

¹⁸ In *The Language of Morals*, we learn of the culture of British India that for an Army Major of the old school to call a subaltern “an awfully good man” would mean that he “played polo, stuck pigs with *élan*, and was not on familiar terms with educated Indians” (p. 146). In Hare’s story about the missionary and the cannibals, the missionary’s term ‘good man’ is said to apply to men who are “meek and gentle and do not collect large quantities of scalps,” while the corresponding term for the cannibals applies to men that are “bold and burly and collect more scalps than average” (p. 148).

for we can clearly apply moral terms confidently to what we know to be counterfactual circumstances in which people have preferences that no actual people do.

Another possibility is to argue that the descriptive content of a term such as 'right' as applied to actions is (something in the ballpark of) 'maximizes preference satisfaction'. This would allow complete leeway with respect to what preferences people had in the circumstances, and it would avoid the difficulties raised above. For to be willing to make the same judgment in all circumstances alike in relevant respects would not require one to take into account in making the present judgment any preferences other than those of people affected by the action. However, there are two significant difficulties with this proposal.

First, Hare is committed to explaining differences in moral judgment across communities, times, and cultures as a matter of their attaching different descriptive meaning to moral statements, though the same prescriptive meaning. In any such community, rational reflection is supposed to lead to judgments in accordance with those prescribed by a preference utilitarian calculus. But the above proposal would make this possible only if all such communities started out with the same descriptive meaning after all.

Second, if this proposal were right, then there would be no need for an argument from prescriptivism to preference utilitarianism in the first place. It would fall out of the descriptive meaning of moral statements. The force of the argument was supposed to derive from its showing us how the logic specifically of the prescriptive component of moral judgments, together with general requirements on rational judgment and requirements on gaining relevant knowledge, would have the serendipitous result of leading us to make judgments in accordance with preference utilitarianism. This defense of the argument makes it superfluous.

It might be suggested that while initially the moral terms may have various descriptive contents, they will converge on a descriptive content that accords with preference utilitarianism, on the grounds that Hare's own argument shows that rational moral agents, because they are required to universalize their moral judgments, will come to make judgments in accordance with those prescribed by preference utilitarianism. What prescriptions rational moral agents find they must make on the basis of the logical of the prescriptive function of moral judgments, since this is, according to Hare, in the case of general or thin moral concepts, dominant, will require that the descriptive content not be in conflict with them. It might be argued that the only descriptive content that would be guaranteed to secure this is identical with the content preference utilitarianism would associate with the relevant term. In the end, then, the descriptive content will be of the right sort to avoid the difficulties that seem to be presented for Hare's argument by attending to the strict force of his argument for universalizability, but the argument can still be seen to be doing the main work.

However, the suggested argument is ineffective against the present criticism. For, everything else aside, it presupposes that the judgments that rational moral agents reach would accord with those endorsed by preference utilitarianism. But that is so only on the assumption that we can limit the universalizability of prescriptions to the positions of actual agents affected by an action one might undertake. That is what the proposal to identify the descriptive content with 'maximizes preference satisfaction' was supposed to secure. Without that assumption, which is given up here, there is no reason to expect that the judgments that rational moral agents would make would be those endorsed by preference utilitarianism, and every reason not to expect that.

We have not found a plausible view of what the descriptive content of moral terms could be that would both have the right result and not render Hare's argument unnecessary. Let us therefore try a different tack. If, in being willing to make the same moral judgment in any circumstance compatible with the descriptive content of it, we must take into account many more preferences than those of actual people who would be affected by an action, we may find ourselves faced with an impossible task. There may be no end to the number of preferences we would have to take into account, and, indeed, it might not be decidable. In the light of this, one might argue that rationality requires that we imposed some additional restrictions to make the task more tractable.¹⁹ In fact, to make the problem tractable, it might be argued, we should restrict attention to the positions of actual people affected with their actual preferences.

One difficulty with this response is the assumption that the rational response to finding that it is impossible to reach a conclusion about what to judge, or that it is undecidable, on the basis of the requirements of the logic of the judgment being made, and general requirements on rationality and knowledge, is to make a judgment on a different basis than that the content of the judgment requires. Consider an epistemic parallel. It is a norm of epistemology that one not form a belief about a matter which one's evidence underdetermines. Suppose that one's epistemic position makes the decision about whether *p* intractable or undecidable. It would not be a rational response to restrict what evidence one pays attention to so as to make the decision a tractable one. In this case, suspension of judgment is the appropriate response. Likewise, if we are unable to reach a decision about the right moral judgment, on the basis dictated by the content of the moral judgment, the right response would be to suspend judgment, rather than to shift from the required basis to some other basis.

¹⁹ This response was suggested to me by Alastair Norcross.

However, even if we waive this difficulty, if we allow that rationality requires that when faced with an intractable or undecidable question, we restrict in some way the problem so as to make it solvable, there will not be a unique solution. Rationality will not dictate that we pay attention only to actual preferences of actual people affected. There are clearly many other ways of restricting what we consider that will make the problem tractable. Some will involve restricting, perhaps, the closeness of possible worlds we consider to the actual world. Some may involve restrictions to the actual world, but not ones that will result in judgments in accordance with preference utilitarianism. For example, we might restrict attention to the preferences of people affected within ten minutes (ten meters) of the action contemplated, or twenty, and so on. The requirement that we make the problem tractable will not be a rational requirement that singles out a solution that accords with preference utilitarianism.

VII

To summarize, Hare's argument from his prescriptivism to preference utilitarianism depends crucially on his argument from prescriptivism to universalizability. I have argued that there are at least two flaws in Hare's argument. First, the argument from prescriptivism to universalizability fails because it depends upon transferring the rational requirement of universalizability of the descriptive component of a moral judgment to its prescriptive component. But there is no rational requirement on universalizing the prescriptive component unless that is determined by its descriptive component, which Hare denies, and which denial plays a crucial explanatory role for him in explaining the extent of moral relativity. Second, were the first stage of the argument to go through, the principle that would be supported is not

the one that Hare needs for the next stage of his argument. The principle that Hare needs is that one must make the same prescription in all circumstances exactly alike with respect to their universal properties. The principle which the argument from prescriptivism would support, if it were successful, is that one must make the same prescription in all circumstances alike with respect to the universal properties on the basis of which the judgment is made. This must be a subset of all the universal properties, for we apply the same terms across changes in universal properties, and across changes in preferences. It cannot be itself the property of maximizing preference satisfaction both because that conflicts with Hare's view and because it would render his argument superfluous. But then plausibly the commitment will extend to making the same judgment in circumstances in which there are persons with different preferences than any actual persons. In light of this, there is no reason to think the calculation, if possible at all, would yield the same results as maximizing preference satisfaction of the actual persons affected. Prescriptivism then fails at two levels to provide a rational grounding for preference utilitarianism, as Hare argued.

REFERENCES

Original publication date in brackets if different from cited version.

- HAO: Gibbard, A. (1988). Hare's Analysis of 'Ought' and its Implications. In D. Seanor & N. Fotion (Eds.), *Hare and Critics: Essays on Moral Thinking* (pp. 57-72). New York: Clarendon Press.
- LM: Hare, R. M. (1952). *The Language of Morals*. New York: Oxford.
- FR: Hare, R. M. (1963). *Freedom and Reason*. New York: Oxford.
- MT: Hare, R. M. (1981). *Moral Thinking: Its Levels, Method and Point*. Oxford: Clarendon Press.
- SUP: Hare, R. M. (1989 [1984]). Supervenience. In *Essays in Ethical Theory* (pp. 113-44). New York: Clarendon Press.
- RAD: Hare, R. M. (1989 [1986]). A *Reductio ad Absurdum* of Descriptivism. In *Essays in Ethical Theory* (pp. 113-44). New York: Clarendon Press.
- SEM: Hare, R. M. (1989). The Structure of Ethics and Morals. In *Essays in Ethical Theory* (pp. 175-190). New York: Clarendon Press.
- OE: Hare, R. M. (1996). Ontology in Ethics. *Canadian Journal of Philosophy*. *Supplementary volume 21*: 83-98.
- OWF: Hare, R. M. (1996). Off on the Wrong Foot. *Canadian Journal of Philosophy*. *Supplementary volume 21*: 67-77.
- SE: Hare, R. M. (1997). *Sorting Out Ethics*. Oxford: Clarendon Press.
- P: Hare, R. M. (1998). Prescriptivism. In E. Craig (Ed.), *Routledge Encyclopedia of Philosophy*. London: Routledge.

SN: Boisvert, D. and K. Ludwig (2006). Semantics for Nondeclaratives. *The Oxford Handbook of the Philosophy of Language*. Oxford: Oxford University Press.