

Andrew Kope

Professor Hurley

Philosophy 2700F

26 November 2009

The Practical Implausibility of Kant's Universalizability Test

Modern ethical theory is typically divided into two main schools, namely teleological and deontological. Teleological ethical theories (e.g. utilitarianism) have a focus on the outcomes of actions, whereas deontological theories focus on the rules and duties that serve as imperatives for one's actions. The foremost deontological ethical theory is Kant's deontology, which states that an action "has its moral worth not in the purpose attained by it, but in the maxim according to which the action is determined" (Kant, 199). The maxim of an action can be defined as "the principle of volition according to which . . . the action has been done" (Pojman, 199). In the Kantian view, every intentional human action has a maxim. For example, the maxim guiding the action of a woman preparing to sleep could be that of 'when I am tired, and it is appropriate to do so, I sleep.' This maxim is what is used to judge the moral worth of an action.

Based on this idea, Kant proposes the idea of a Categorical Imperative, which "declares an action to be of itself objectively necessary without reference to any purpose" (Pojman, 204). The Categorical Imperative is first drafted by Kant as the Formula of Universal Law. The Formula of Universal Law judges whether or not an action is good in itself, and whether we are making an exception of ourselves in holding that maxim. It states that one should "act only according to that maxim whereby you can at the same time will that it should become a universal law of nature" (Pojman, 205).

The Formula of Universal law, by extension, provides the *Universalizability Test* for the

maxims of one's actions. The Universalizability Test involves imagining the maxim of an action as a universal law of nature, and then assessing whether that universal law contradicts itself in one of two ways. This contradiction can either be a contradiction in conception, where the maxim cannot even be thought of as a universal law of nature without contradicting itself, or a contradiction in will, where the willing of the maxim would contradict itself if the maxim were to be a universal law of nature. Maxims that fail the Universalizability Test through a contradiction in conception result in a *perfect duty* (one we must always do when the opportunity arises, typically negative – i.e. a duty not to do something). Maxims that fail the Universalizability Test through a contradiction in will result in an *imperfect duty* (one we must perform at least some of the time, usually positive – i.e. a duty to do something).

Whether the Universalizability Test of the maxims of our actions is a valid test that provides the “right” moral results is still a matter of debate. Opponents of Kant's view have posited two main objections to the Universalizability Test; namely that it can produce false positives and false negatives (Feldman, 227). It is my position, however, that when using the *true maxim* of an action the Universalizability Test is in theory a valid moral test. Secondly, I posit that though theoretically valid the Universalizability Test does not in practice provide us with the “right” moral results, because the true maxims of our actions are not available to us.

To argue these two theses, I will first define what I mean by the *true maxim* of an action (in contrast to an *observed maxim*), and show that it is not available to us. I will then demonstrate that when using the true maxim of an action, the Universalizability Test gives us the “right” moral results (results that are intuitively plausible). I will then explain how the examples demonstrating the claims that the Universalizability Test does not give us the “right” moral results (because it produces false positives and false negatives) rely either on a misapplication of

the test, or on the use of inaccurate observed maxims. Finally, to defend my second thesis I will argue that because we cannot know the true maxim of an action, the Universalizability Test is (although theoretically valid), an irrelevant consideration when deciding on morally “right” courses of action.

The *true maxim* of an action can be defined as the actual, objective, and unequivocal maxim of the action. The true maxim of an action is the only maxim among a multitude of plausible alternatives that actually captures the true principle of volition behind the action. The true maxim of an action can also be said, by definition, to be impossible to determine in any practical setting, as “we can never, even by the strictest examination, completely plumb the depths of the secret incentives of our actions” (Pojman, 202). Although we can specify the true maxim of a theoretical action in a thought experiment, because we lack the ability to objectively assess our own internal states, we can only speculate on the true maxim of a real action (this speculation produces what we might call an *observed maxim*: an apparent maxim for an action that though plausible, is not known to be the objective, unequivocal *true maxim*). For example, the observed maxim of a man speeding in traffic from could be ‘whenever I am in a rush, I will do whatever I can to reach my destination faster, even at the cost of the safety of others.’ The true maxim of his action, however, could be ‘whenever my wife is in urgent need of medical attention, I will do whatever I can to get her to a hospital faster, while putting others in as little danger as possible.’

The relevant distinction between the observed maxim and the true maxim of an action, as I have defined them, is that the Universalizability Test does not necessarily give the “right” moral result when using an observed maxim of an action, but can be assumed to give the “right” result when used with the true maxim of an action. Though impossible to prove conclusively

(given the artificial nature of all constructions of a true maxim for an example action), it seems intuitive that the true maxim of our action would be the appropriate one to extend to a universal law of nature, and then assess whether that universal law contradicts itself (the algorithm of the Universalizability Test). To demonstrate this point, I will use two examples, one of an action whose true maxim passes the test and whose observed maxim does not, and another of an action whose true maxim does not pass the test and whose observed maxim does.

In my first example, a man (Rick) kills another man in an alleyway. The observed maxim of his action by a passerby was ‘I will kill those who make me angry.’ This maxim would result in a contradiction in will (Rick would not want others who he makes angry to kill him), and would therefore fail the Universalizability Test, and his action would be morally wrong. In actuality, the true maxim of his action was ‘I will kill those who make me angry only in self defence.’ The true maxim of his action, unlike the observed maxim, passes the Universalizability Test and is morally acceptable.

In a second example, a man (James) makes a large amount of money in high finance. The observed maxim of his action to his colleagues was ‘I will fight as hard as I can to make money in the stock market.’ This maxim would pass the Universalizability Test. The true maxim of his action, however, was ‘I will manipulate the financial system and steal from my associates to make money in the stock market.’ The true maxim of his action, unlike the observed maxim, fails the Universalizability Test (a contradiction in will; James would not want others to steal from him) and is morally impermissible.

Having demonstrated that the Universalizability test using the true maxim of an action gives us what appears to be the “right” moral result, it can now be shown that the claims that the test produces false positives and false negatives rely on either misapplying the test (false

positives) or on the use of an inaccurate observed maxim (false negatives).

A false positive of the Universalizability Test passes a maxim that is the principle of volition for an intuitively immoral action. For example, a university student with the maxim 'I will buy my term paper whenever doing so will not result in a change to the school system in general' (Feldman, 226) would seem at first to pass the Universalizability Test. On further examination, however, it becomes clear that the Universalizability Test would not actually pass this maxim. This maxim would actually fail as a contradiction in conception, because either university students as a whole would all cheat, or no one would cheat (the 'some of the time' stipulation of the maxim translates to an all-or-none binary when everyone is acting to serve their own best interest ignorant of the cheating of others). In the first case, if everyone cheats, the school system in general would be affected, and in the second case, if no one cheated, then the maxim would propose an action that could never happen (a contradiction in will and a contradiction in conception, respectively). As such, one can not imagine this maxim as a universal law without self-contradiction, and it fails the Universalizability Test, eliminating the false positive.

A false negative of the Universalizability Test occurs when it fails a maxim that is the principle of volition for an intuitively innocuous action. For example, someone playing tennis with the maxim 'I will play tennis at 10:00 when everyone else is at church and the courts are not busy' (Feldman, 227) would seem to be morally in the right. When universalized, however, this maxim results in a contradiction in conception, as if everyone played tennis at 10:00 on Sundays, the courts would be incredibly busy, and one can not imagine this maxim as a universal law without self-contradiction (contradiction in conception). I would argue, however, that this maxim results in a contradiction in conception indicating that the action is morally impermissible

because it is not the true maxim of the action. Instead, the true maxim of the action could be ‘I will play tennis when the courts are not busy’ (this does not result in an all-or-none situation as in the cheating example above because the playing/not playing status of others is clearly apparent, unlike their cheating/not cheating status). This new maxim that is likely the true maxim of the tennis player passes the Universalizability Test, and eliminates the false negative.

Although the Universalizability Test of the Categorical Imperative has been shown not to produce false positives or negatives as others have claimed, it is however still ultimately ineffective when deciding on morally “right” courses of action. Its inadequacy in determining the morally “right” course of action is derived from the fact that it relies on universalizing the *true maxims* of our actions. As demonstrated previously, the *true maxims* of our actions are not available to us (outside of the context of a thought experiment). As such, even if the *true maxim* of an action can be assumed to always pass or fail the Universalizability Test according to the moral character of the action (in terms of the Categorical Imperative), because we cannot ever know the true maxim of an actual action, the Universalizability Test is an irrelevant consideration when determining morally “right” courses of action. In theory, the Universalizability Test is a valid test of the moral character of a given action, though only because we can specify the true maxim of that action. In practice, the Universalizability Test can often fail to give us the “right” moral results, because it is forced to rely solely on fallible observed maxims, and is therefore an irrelevant consideration when actually deciding on morally “right” courses of action.

My opponent might argue that the Universalizability Tests lack even theoretical validity, and its failing does not rest in our inability to assess the true maxims of our actions. For example, Barbara Herman offers the following thought experiment,

Suppose I know that you will gamble away money saved for your child’s

education, a deed you will later regret. You will also loan the money to me, if I tell you that I am in trouble and need it. So I decide to ask you for a loan, though I do not in fact need the money. I promise to repay the loan, but do not intend to. I will instead keep the money for your child. (141)

My opponent would suggest that this action seems morally permissible (one's intuition is that gambling away money to be used for a child's education is wrong), however would fail the Universalizability Test based on its observed maxim of 'I will borrow money and not repay it if it suits my goals' (a contradiction in conception resulting in no one lending money, as there will no longer be any reason to believe it will be repaid). In a similar way to the false negative argument addressed earlier, Herman's objection is based on the observed maxim of the relevant action, and not its true maxim. What is more likely the true maxim of the action is 'I will save you from yourself when I know you will thank me for it later, to promote the welfare of your family.' This potential true maxim of the action does pass the Universalizability Test, once again indicating that the test gives the intuitively "right" moral results when using the true maxim of an action.

I have argued that when using the *true maxim* of an action the Universalizability Test is a valid moral test of maxims, however also that it does not *in practice* actually provide us with the "right" moral results, because the true maxims of our actions are not available to us. Having argued against the impractical nature of the Universalizability Test, I would like to conclude with the proposition that arguments about whether it gives us the "right" moral results and is valid in and of itself seem ultimately to be moot – regardless of whether the test can give us the correct moral results when used with a true maxim, the true maxims of our actions are not available to us and the test therefore never practicable, and is in fact largely irrelevant.

Works Cited

Herman, Barbara. *The Practice of Moral Judgement*. Harvard University Press, 1993.

Kant, Immanuel. "The Foundations of Ethics." In: Louis, Pojman. *Moral Philosophy*.

Indianapolis: Hackett Publishing Company, 2003.