



DISCUSSION NOTE

*IS KANTIAN ETHICS SELF-REFUTING?
A REPLY TO MILLGRAM*

BY JOSHUA GLASGOW

JOURNAL OF ETHICS & SOCIAL PHILOSOPHY

DISCUSSION NOTE | SEPTEMBER 2008

URL: WWW.JESP.ORG

COPYRIGHT © JOSHUA GLASGOW 2008

Is Kantian Ethics Self-Refuting?

A Reply to Millgram¹

Joshua Glasgow

FRIENDS OF KANTIAN ETHICS have made significant strides in reconciling a theory that was once perceived as alien to ordinary human agency with the complex realities of everyday life. However, it has been recently argued that one element of this reconciliation project renders a core Kantian principle self-refuting. The target principle is Kant's requirement that we act only on those maxims that can be consistently universalized. As part of this consistency requirement, the universalized maxim must be compatible with rational agency and, according to the "New Kantians," with the agency of ordinary humans. Elijah Millgram argues that, so construed by New Kantians, a maxim of abiding by this principle fails its own test: in a world where everyone acts according to the universalizability principle, everyone will act on policies that admit no exceptions, and ordinary human agency requires that those exceptions be made.

Importantly, this is explicitly *not* supposed to be a rehashing of the tired (and false) accusation that the universalizability principle *generates* exceptionless, or absolute, moral rules. In Millgram's own words, "Kant has sometimes been accused of having a moral theory that generates exceptionless rules (that everybody has to abide by); that was a misperception of a different feature of the view, which is that it *operates on* exceptionless rules (but different ones for different agents at different times)" (Millgram 2003: 533). And because of this feature, it fails its own test, at least as construed by the New Kantians.

Here I defend Kantian ethics against this objection. What I hope to show, though, is not that the universalizability principle is not self-refuting, but that it is self-refuting in the way suggested by Millgram only if it *does* generate absolute moral rules. Put otherwise, if Millgram is correct that the charge of absolutism is *not* properly directed at the universalizability principle, then the self-refutation charge will also fail. But I won't argue here for the truth of the antecedent, in part because Millgram already endorses it.

The principle in question is the Law of Nature formulation of Kant's Categorical Imperative (FLN), which sets the following moral permissibility test (roughly and briefly put). If it is not possible to coherently conceive of a world in which everyone wills your maxim as if it were a law of nature – if universalizing your maxim generates a "contradiction in conception" – then it is impermissible. Furthermore, even if you can conceive of such a world, if it is nevertheless inconsistent with willing any maxim that is rationally required, then it generates a "contradiction in the will," and the (original)

¹ Previous versions of this paper were included in talks given at the University of Otago and Victoria University of Wellington. I am very grateful to the participants in those sessions, as well as to Ramon Das, Elijah Millgram, Douglas Portmore, David Shoemaker, Mark Timmons and anonymous referees, for helpful discussion.

maxim is impermissible. Your maxim is morally permissible when universalizing it generates neither kind of contradiction.

When identifying contradictions in the will, New Kantians “exploit deep facts about specifically human agency – about the range of ends that it is reasonable to expect humans to adopt, about the inability of humans to do everything themselves [etc.]” (Millgram 2003: 530). Consider the (New and Old) Kantian argument that non-beneficence is forbidden: a world of universal non-beneficence is conceivable, but it would frustrate our inevitable need for help from others, and thus it generates a contradiction in the will – an inconsistency between nobody getting any aid and the agent rationally willing aid for herself (G: 423).² As stated, this argument would not work for a more self-reliant species; it only applies to ordinary, needy humans.³ Furthermore, the argument does not require that the *entirety* of human agency be frustrated; rather, universal non-beneficence merely must threaten “the range of projects that a rational deliberator...might well adopt” (Millgram 2003: 531), where such projects are those pursued by ordinary humans living in a world with a division of labor.

This claim of New Kantianism – that contradictions in the will can rely on facts about ordinary human agency – is Millgram’s target. His argument that making FLN-compliance itself one’s maxim would fail the New Kantian version of the FLN test is as follows (Millgram 2003: 532-534): For FLN-compliance to itself be FLN-permissible, there must be a world where everyone acts only on maxims that can be universalized without contradiction. That is, in this world everyone adopts only lawlike maxims, which means (according to Millgram, anyway) that they have no exceptions. But just as my (ordinary) agency requires that I receive some aid from others, it also requires that others make policy exceptions for me. For one mundane example, unloading luggage in a city with few public parking spaces requires the private parking-lot attendant to make a quick exception to allow the traveler (who has no parking permit) to temporarily park in the lot while unloading luggage. More broadly, since our “projects are vulnerable,” since we cannot plan for every potential policy that we will cross when pursuing our projects, since social rules are not tailored to the needs of every individual, and since we cannot even predict many of our future ends, it is unlikely that our projects and others’ policies will be tightly coordinated (Millgram 2003: 536-537). So because ordinary human agency requires pursuing projects in an imperfectly coordinated social world, exceptions will have to be made. But exceptions seemingly cannot be made if everyone only adopts maxims in a lawlike man-

² I refer to Kant’s *Groundwork* (1997) as “G,” citing the Academy pagination.

³ This argument is arguably inapplicable even for a rugged individualist in our own species: presumably even those on deserted islands exercise agency effectively, even though they are unaffected by others’ behavior. In presupposing that agents are not isolated in this way, Millgram is merely working within the premises set by the New Kantians. (For the case that Kant supplies another argument, which does not make this presupposition, see Glasgow 2001.)

ner. In short, the demands of ordinary agency rationally require me to will that others make exceptions to their policies, which is inconsistent with willing a world in which everyone follows FLN, for in that world nobody makes exceptions to their (lawlike) policies. The maxim of abiding by FLN is itself prohibited by FLN. New Kantianism is self-refuting.

Let's isolate the premises of this argument more precisely. (1) For a maxim of complying with FLN to itself be FLN-compliant, its universalization must be consistent with ordinary human agency. (2) If everyone is committed to FLN-compliance, everyone acts on maxims that they treat as lawlike policies. (3) Lawlike policies exclude exceptions. But (4) ordinary human agency requires agents to make exceptions to policies, so (5) a world of universal FLN-compliance is inconsistent with ordinary human agency, in which case (6) a maxim of FLN-compliance is not itself FLN-compliant. I want to focus on premise 2.

Millgram (2003, 532) states 2 and 3 as follows: "accepting the [maxim of FLN-compliance] involves ... understanding your actions as governed by, or as the deliverances of, lawlike policies. (I am not just bearing down hard on something Kant happened to say; I will in due course argue that the [FLN test] can work only if its inputs are lawlike policies.) For a policy to be lawlike is for it to have no exceptions. So the [world of universal FLN-compliance] is one in which, when anyone acts, he understands himself to be acting from a universal and so exceptionless policy that governs his action." But as far as I can tell (Millgram is not explicit), the argument that comes in due course defends not the claim that (2) the FLN-compliant agent must see her actions as governed by lawlike policies, but the separate claim that (3) lawlike policies must be exceptionless. Thus Millgram never seems to defend premise 2.

Premise 2 is also, I believe, false: FLN-compliant agents do not need to treat their maxims as lawlike. Consider Kant's statement of FLN: "*act as if the maxim of your action were to become by your will a universal law of nature*" (G: 421, emphasis added). FLN's thought experiment requires investigating whether my maxim *could be* lawlike, but treating it that way for a *thought experiment* does not require that I actually adopt it as law governing *my behavior*. So while FLN-compliant agents only adopt maxims that are compatible with the form of law, this commitment alone does not also commit them to treating their maxims like laws. Instead, the FLN-committed agent is simply committed to straying from those maxims only in FLN-compliant ways.

For example, when an agent changes her maxim from "To jog every Saturday" to "To hike every Saturday," that switch of equally FLN-compliant maxims – that failure to treat the first maxim like an inviolable law – does not entail that she repudiates FLN-compliance as a governing maxim in her life. Even if we grant that (3) treating a maxim as lawlike means making no exceptions to it, for our agent to permissibly decide to hike instead of jog, it only need be the case that the hiking maxim could *also* be universal law. So, since maxims are just one-off affairs (because they only represent our one-off intentions), the fact that a FLN-compliant maxim *could be* a law of nature

does not mean that it *actually is* law or *must be adopted by the agent as* law.⁴ Thus 2 is false: it is not the case that FLN-compliant agents act only on maxims that *they treat as* lawlike (whether or not doing so would require treating their maxims as exceptionless); instead, they only act on maxims that *could be* law.

But forget about jogging and hiking – what about instances like the parking-lot case? Can the attendant make the kind of flexible exception we need to enable agency? Like the jogger/hiker, the FLN-compliant attendant can switch from a “Permit holders only” maxim to a “Permit holders plus the occasional road-weary traveler” maxim, so long as each maxim is FLN-compliant. Assuming, with Millgram, that the first, absolute maxim could be law, the remaining question is whether the second, exception-inclusive one also could be law. If it and other exception-inclusive maxims could be law, then that suggests that exception-requiring agency is compatible with a world of universal FLN-compliance, which would neutralize the self-refutation argument. If they cannot, then the self-refutation objection will withstand the rejoinder as so far presented. *But for that to be so, FLN must only license maxims that do not include exceptions.* That is, it must only generate absolute or exceptionless moral prohibitions.

Put more diagnostically, if we adjusted premise 2 to make it true, we would reformulate it to say that (2*) if everyone is committed to FLN-compliance, everyone acts on maxims that could be universal law. To make the argument valid, 3 would also then have to be adjusted to say that (3*) maxims that could be universal law must exclude exceptions. And therein lies the problem: 3* is just the premise that FLN generates only absolute moral rules – that any maxim that includes exceptions will violate FLN’s universalizability test.

Thus, far from rejecting the tired absolutism objection to Kantian ethics, the self-refutation objection revives and requires it. So if Millgram is right that it is inappropriate to accuse FLN of absolutism, that is, if 3* is false, then FLN will withstand the self-refutation objection, too.

Joshua Glasgow
Victoria University of Wellington
Philosophy Programme
joshua.glasgow@vuw.ac.nz

⁴ For further discussion of the point that adopting a maxim for *this* action does not commit the agent to that maxim for other actions at other times, see Glasgow 2003; Nell 1975, 34; O’Neill 1989, 129-30. Note that this does not mean that maxims cannot *also* capture one’s long-term plans or deep commitments, such as the commitment to abide by FLN; it just means that, *qua* maxims, they are representations of *this* or *that* long-term plan or deep commitment, or present intention. This one-off kind of maxim, at any rate, seems to be the kind to which the FLN test applies; for a recent examination of various senses of “maxim” at work in Kant’s ethics, see, e.g., McCarty 2006.

References

- Glasgow, Joshua. 2001. "Kant's non-prudential duty of beneficence." In Volker Gerhardt, Rolf-Peter Horstmann and Ralph Schumacher, eds., *Kant und die Berliner Aufklärung: Akten des IX. Internationalen Kant-Kongresses* (Walter de Gruyter), Bd. III, 211-219.
- . 2003. "Expanding the limits of universalization: Kant's duties and Kantian moral deliberation." *Canadian Journal of Philosophy* 33, 23-48.
- Kant, Immanuel. 1997. *Groundwork of the Metaphysics of Morals*. Mary Gregor (Trans. and Ed.). (Cambridge: Cambridge University Press).
- McCarty, Richard. 2006. "Maxims in Kant's practical philosophy." *Journal of the History of Philosophy* 44, 65-83.
- Millgram, Elijah. 2003. "Does the Categorical Imperative give rise to a contradiction in the will?" *Philosophical Review* 112, 525-560.
- Nell, Onora. 1975. *Acting on Principle: An Essay on Kantian Ethics*. (New York: Columbia University Press).
- O'Neill, Onora. 1989. *Constructions of Reason: Explorations of Kant's Practical Philosophy*. (Cambridge: Cambridge University Press).