Action from Duty But Not in Accord with Duty

John Hardwig

Action from Duty But Not in Accord with Duty

John Hardwig

In the *Foundations*, Kant draws a distinction between action which is in accord with duty and action which is done from the motive of duty. This is a famous distinction, of course, and there are many interesting issues concerning it and its implications for ethical theory. In this paper, I wish to focus on just one noteworthy feature of Kant’s use of this distinction. Like any distinction between logical compatibles, this one yields four logically possible classes of action: (1) actions which are both in accord with duty and from duty; (2) actions which are neither from duty nor in accord with duty; (3) actions which are in accord with duty but not from duty; and (4) actions which are from duty but not in accord with duty. What interests me about these four possibilities is that, to the best of my knowledge, Kant never considers or even mentions the last of these possibilities: action from duty but not in accord with duty. This is perhaps surprising in a philosopher with Kant’s interest in logic and passion for thoroughness. One would have thought that he would mention this logical possibility, even if only in order to discount it as not really possible.

Beginning with the idea that there are cases of action from duty but not in accord with duty, I argue in this paper that Kant could not have admitted that there can be actions of this kind, for their existence undermines an important aspect of his ethical theory—his view that moral principles and judgments must be both autonomous and universally valid.

Content can easily be given to three of the four classes of action I have distinguished, and Kant does so quite explicitly. Action which is both in accord with duty and from duty is morally correct action; indeed, it is morally praiseworthy action or, as Kant puts it, action which has “moral worth.” Action in a moral context which is neither in accord with duty nor from duty is immoral action. And, although there may be problems in determining the moral status which Kant assigns to action in accord with duty but not from duty, it is again reasonably clear what sorts of actions fall into this class; they are actions which correspond to the precepts of ethics, even if only accidentally, but which are motivated by immoral or at least nonmoral motives.
But what kind of actions could be classified as actions from duty but not in accord with duty? Obviously, these would be actions which are motivated by the desire to do one’s duty, but which do not correspond to the requirements of morality; they would be cases in which the agent is mistaken about what his duty is. The agent in question intends to do his duty; in fact, he does what he takes to be his duty and he does it because it is seen to be his duty. Thus his action must be described as action from duty. But he is mistaken about what duty (in general) is, about the morally relevant facts of the context of his action, or about how the general requirements of moral action apply to the context of his action. Because he is mistaken, he is in fact not acting in accord with duty, despite what may have been a conscientious attempt to formulate the correct moral judgment and despite his belief that he is doing his duty because it is his duty. He is, then, acting from duty but not in accord with duty.

Now it seems clear—beyond the need for argumentation, really—that there are actions of this sort. We are not dealing here with a mistaken assessment of consequences, for Kant’s ethics is not a consequentialist ethics. Rather, for Kant, immoral actions are actions based on maxims which cannot be universalized without contradiction. But contradictions need not be apparent to one who holds contradictory beliefs, and they do not always reveal themselves to agents who conscientiously look for them. In Kant’s terms, then, an action from duty, but not in accord with duty would be an action which the agent believes he can will to be a universal law but which cannot in fact be universalized without contradiction. The inquisitor torturing a heretic in a last-ditch attempt to save his soul; a goodwilled Nazi (perhaps Himmler himself) engaged in a desperate struggle to save Western civilization; a social scientist marshaling arguments to convince his students that morality is nothing more than social mores; a woman destroying her potential and her life in an attempt to honor, love, and obey her husband—these could be examples of action from duty but not in accord with duty. In fact, I would argue that anyone acting as Kant recommends in “On the Supposed Right to Lie from Altruistic Motives” would merit precisely this description.

If there clearly are cases of action from duty but not in accord with duty, why is it that Kant overlooks them? Is this simply an oversight? Perhaps Kant does not discuss this kind of action because he believes that there are no actions of this kind. To say that there are no actions from duty but not in accord with duty is to say that there are no mistaken moral judgments, perhaps even that the whole idea of a mistaken moral judgment is itself a mistake. If there are no mistaken moral judgments, then there is no moral problem of knowledge. Could it be that this is Kant’s view? It seems implausible that it is, given his repeated insistence

1. That Himmler himself is an example is the view of Jonathan Bennett (see “The Conscience of Huckleberry Finn,” *Philosophy* 49 [1974]: 128–34).
on the rationality and objective or universal validity of moral legislation and judgment. The theory which Kant develops for determining whether action is in accord with duty is so intricate that not even an ethics professor can be sure that he has it right. Can Kant conceivably have held that these difficulties arise only for philosophical reflection on morality, but not for moral action itself? Can he have thought that moral agents always arrive at correct moral judgments? Perhaps that ordinary moral agents, like native speakers of a language, implicitly follow rules that are extremely difficult to formulate? At the risk of exploring a view of Kant’s ethics that is totally implausible or one-sided, at best, I wish to suggest that Kant may very well have held that there is no moral problem of knowledge and hence no actions from duty but not in accord with duty.

This much, at least, is clear. For Kant, as opposed to Plato, the paradigm problem of moral action is not a problem of knowledge, but a problem of volition. The paradigm moral situation is not one in which an agent may act wrongly because he mistakenly believes that he is doing what he ought, nor one in which the agent struggles with the problem of discovering what he ought to do; rather it is the situation in which the agent struggles to do what he knows he should do. It is also clear that Kant thought that the primary role for a philosophical theory of ethics is to resolve the theoretical problems of the philosopher, not to guide action by illuminating the practical difficulties of the moral agent. On Kant’s view, everyday morality, like Newtonian physics, suffers not so much from internal difficulties as from a lack of adequate philosophic foundation and from skeptical assaults launched from theoretical positions. There are certainly any number of mistaken metaethical theories, on Kant’s view, but the existence of mistaken moral judgments does not follow from the existence of mistaken theories of moral judgment.

Nevertheless, it is a long way from the view that the problems that besiege ethicists are philosophical problems and that the main problem of the moral agent is not a problem of knowledge, to the view that there is no moral problem of knowledge and no actions from duty that are not in accord with duty. Does Kant hold this more extreme view? His writings are not consistent on this point, but there are passages in almost every one of his major writings on ethics in which Kant does seem to maintain that there is no moral problem of knowledge.2 Perhaps the most emphatic of these is in the second Critique: “What is required in accordance with the principle of autonomy of choice is easily and without hesitation seen by the commonest intelligence. . . . That is to say, what duty is, is plain

of itself to everyone. . . . The decision as to what is to be done in accordance with [the moral law] must not be so difficult that even the commonest and most unpractised understanding without any worldly prudence should go wrong in making it.”3 At other times, however, Kant wishes to insist on the need for moral enlightenment and on the practical importance of philosophical discussions of ethics. I do not wish here to pursue further the attempt to discover whether Kant really does hold the view that there are no mistaken moral judgments. Rather, I will try to show that he must hold this view. Given his ethical theory, Kant cannot consistently admit that there could be actions from duty, but not in accord with duty.

We can understand how the denial that there can be mistaken moral judgments is integral to Kant’s ethical theory if we focus, not on his doctrine of moral rationality per se, but on a more neglected side of his ethical theory: his doctrine of autonomy and his view that only autonomous moral legislation and judgment can yield categorical or moral obligation. As Kant puts it, autonomy of the will is the capacity of the will “by which it is a law to itself.”4 Kant’s own formulations of the principle of autonomy intimately connect it with the principle of universality or moral rationality. However, these two metaethical principles are not simply synonymous. They are two, distinct principles and Kant, in fact, recognizes that he must argue that the principle of universality implies the principle of autonomy and vice versa. As we shall see, if there are actions from duty but not in accord with duty, these arguments cannot succeed. In order to show this, it will be useful to separate, momentarily, Kant’s doctrine of autonomy from his doctrine of moral rationality. If we sever the principle of autonomy from its putative connections with the principle of universality, we can formulate it as follows: “a person can be morally obligated only by a law which he gives himself” or “a person can be morally obligated only by his own conception of the moral law.” Kant offers strong arguments in support of this principle of autonomy.

Kant argues, of course, that the legislation which yields moral obligation must be autonomous because any law or principle which does not originate in the self presupposes an “incentive”—the hope of a reward or the fear of punishment—in order for the law to be adopted by the self as a principle of volition. The need for incentives means that external legislation can, by definition, yield only hypothetical, not categorical or moral imperatives. Accordingly, Kant maintains that external objects or goods cannot be the source of moral legislation. More important for present purposes, the principle of autonomy also means (and Kant is clear about this) that moral obligation cannot be grounded in the laws of the state or, more generally, in other people’s principles, nor can it be grounded in the legislation of a Supreme Being, his infinite wisdom and benevolence notwithstanding. Because moral obligation must be categorical obligation,

any external, any other-than-self source of moral principles is precluded. As Kant puts it repeatedly in *The Metaphysical Principles of Virtue*: “Ethical legislation cannot be external.”

Further, although Kant’s discussions of autonomy focus on the autonomy of the moral legislator, the autonomy of the moral judge can be established by the same arguments. If an agent’s concrete moral obligation is to obligate *him* categorically, *he* must be the judge who applies the moral law to the particular context of his action. This is clear if we consider the case in which someone else determines what an agent’s duty is, the agent disagrees with this judgment, and attempts to persuade him fail. In this case, Kant’s argument immediately applies—some incentive would be required to induce the agent to accept the other’s moral judgment as a principle of volition for him. If the agent accepts the inducement which the incentive supplies and does what the other person thinks he ought, he is acting only on a hypothetical, not a categorical, imperative; and hypothetical imperatives cannot, for Kant, express or ground moral obligation.

Given this view that the autonomy of the moral legislator and judge is necessary for moral obligation, we are in a position to see why Kant cannot admit that there are actions from duty but not in accord with duty. For, as is well-known, he also maintains that moral legislation and judgment must be objectively or universally valid—valid for all rational beings—and a mistaken moral principle of judgment obviously would not be universally or objectively valid, simply because it is mistaken. To admit that there can be actions from duty but not in accord with duty is to admit that autonomous moral judgments need not be objectively or universally valid and also that objectively or rationally valid moral judgments are not always reflected in a given agent’s autonomous judgments.

To put the problem in its most dramatic form, the moral agent who holds mistaken moral views could have no moral obligations. For there are, for Kant, two necessary conditions for moral obligation—autonomy and universality. In order to obligate anyone morally, a principle or judgment must be universally valid or valid for all rational agents; in order to obligate *me* morally, it must be my own principle or judgment. Consequently, if I hold principles which I take to be moral principles or formulate judgments which I take to be moral judgments, and if these principles or judgments are mistaken, I have no moral obligations. For, in that case, there is no principle or judgment which meets both of the necessary conditions for moral obligation. My principles or judgments are not universally valid, and universally valid principles or judgments are not *my* principles or judgments. Admittedly, they would be mine if I were perfectly rational and made no mistakes. But my adherence to mistaken moral principles and judgments proves that I am not. If moral

legislation must be universally valid legislation, I could perhaps be said to have temporarily lost my capacity to be a moral legislator. And because "ethical legislation cannot be external," if I have lost the capacity to be a moral legislator, there is no moral legislation that is applicable to me.

The possibility of mistaken moral judgments thus reveals a conflict or tension between two fundamental metaethical principles which Kant thinks are united: the principle that moral laws and judgments must be objectively or universally valid and the principle that they must be autonomous. Kant is quite aware that his theory requires the union of these two principles; in fact, he takes this to be the key to his solution for the philosophical problems of ethics. He proclaims:

If we . . . look back upon all previous attempts which have ever been undertaken to discover the principle of morality, it is not to be wondered at that they all had to fail. Man was seen to be bound to laws by his duty, but it was not seen that he is subject only to his own, yet universal, legislation, and that he is only bound to act in accordance with his own will, which is, however, designed by nature to be a will giving universal laws. For if one thought of him as subject only to a law (whatever it might be), this necessarily implied some interest as a stimulus or compulsion to obedience because the law did not arise from his own will.7

This union of the principle of autonomy and the principle of universality lies at the very heart of Kant’s ethics, and he cannot give up either without destroying the core of his theory. He cannot give up the principle of universality, for without it he cannot claim that moral principles are objective, valid for all rational beings. And he cannot give up the principle of autonomy, for even rational principles will not yield categorical obligation unless they are products of the agent’s will, unless they are his principles. External principles may be impeccably rational, but they can never, as we have seen, yield categorical or unconditional obligation, precisely because they are external.8 Kant therefore needs to maintain that moral principles and judgments are both autonomous and universally valid. But he can maintain that they are both, only by denying the possibility of mistaken moral judgment. This is the fundamental reason why Kant must claim that there is no such thing as an action from duty but not in accord with duty.

6. Ibid.
8. Strictly speaking, I think it could be shown that two different senses of "categorical" or unconditional obligation are conflated in Kant’s ethics. One of these senses corresponds to the requirement that moral principles be objectively valid, valid for everyone, without conditions. The other sense arises from the fact that only the agent’s own principles can obligate him without presupposing some incentive, or condition. The argument of this paper would also serve to show that there is a tension between these two senses of categorical obligation. When I speak of categorical obligation here and elsewhere in this paper, I am referring to the second of these two senses of the term.
But actions from duty but not in accord with duty exist anyway. And because they exist, we see that there is a discrepancy between the principle of autonomy and the principle of universality; some autonomous moral judgments are not universally valid and some universally valid moral judgments are not reflected in the autonomous judgments of all agents. Can Kant escape the difficulty that this conflict implies for his ethics? Two possibilities present themselves: a reinterpretation of the principle of autonomy and a reinterpretation of the principle of universality. Neither seems satisfactory.

In the first case, Kant could modify the principle of autonomy by arguing that a judgment cannot really be autonomous unless it is also objective or universally valid. Mistaken moral judgments simply are not autonomous, for they are not deliverances of pure practical reason. This is the tack that is implicitly taken by many commentators on Kant’s ethics. But it will not solve the problem. For if we redefine autonomy in this way, the agent who holds mistaken moral principles or who acts on mistaken moral judgments must be held to be ignorant of or mistaken about his own autonomous judgment; otherwise he would not embrace moral judgments which are not universally valid. But an autonomous judgment which the agent is unaware of will not obligate him categorically. If you know, as I do not, what my autonomous judgment is and if persuasion fails and I cling to my mistaken moral judgment, then you must introduce incentives if I am to act on my autonomous judgment. But in this case, my action is based on a hypothetical imperative not a categorical or moral imperative. An agent can be unconditionally or categorically obligated only by what he believes to be his own moral judgment. An agent’s unconditional obligations can derive only from the moral vision he accepts.

Alternatively, we might try to solve Kant’s difficulty by modifying the principle of universality. Some contemporary ethicists, perhaps sensing Kant’s difficulty and yet desiring to unify freedom and reason in their accounts of the moral point of view, have tended to water down Kant’s principle of universality to a principle of universalizability. A principle can be said to be universalizable if the agent who acts on it believes, on the basis of a careful and conscientious attempt to arrive at an ethical judgment, that he has discovered a principle which is universally valid. A universalizable principle is universal or objectively valid if what he believes is true—that is, if his moral judgment is not mistaken. Then, if we say that the moral point of view requires that moral judgments be universalizable, but not necessarily universally valid, the difficulty Kant is in is appreciably decreased. But it is not completely eliminated unless the principle of universalizability is itself a universal principle—“universal,” this time, not only in the sense of “valid for every moral agent,” but also in the sense of “accepted by every moral agent.” If there are moral agents who do not accept the principle of universalizability (and I believe that there are), then there is a tension between even the more modest principle of
universalizability and the principle of autonomy. For there will then be moral agents for whom even universalizable moral judgments will not yield categorical or unconditional obligation.

To summarize, the existence of actions from duty but not in accord with duty reveals a difficulty which lies at the core of Kant's ethics: there is a conflict or tension between the requirement that moral judgments be universally valid and the requirement that they be autonomous. This is not, however, only a problem for Kant; rather, we touch here on what I take to be a fundamental problem for many theories of ethics. The importance of something like Kant's principle of universality—perhaps only the weaker principle of universalizability—has been so widely recognized that it is often taken to be a defining characteristic of the moral point of view. The principle of autonomy has been largely neglected in contemporary discussions of ethics. But it ought not to be. For with it Kant is pointing out what is clearly true: the notion of principled moral action in the full sense presupposes that the principles an agent acts on must be his own. Moreover, if Kant is right in thinking that moral obligation is unconditional or categorical obligation, then any account of moral obligation will require something like the principle of autonomy. Because actions from duty but not in accord with duty exist, ethical theories cannot consistently maintain that moral principles and judgments must be both autonomous and objectively or universally valid. And yet we seem to need to say that they must be both.