

# Kant's Derivation of the Formula of the Categorical Imperative: How to Get it Right

by Jacqueline Mariña, West Lafayette

In this paper I will discuss a recent controversy over Kant's derivation of the formula of a categorical imperative. Bruce Aune<sup>1</sup> and Allen Wood<sup>2</sup> have both called attention to a "gap" in Kant's deduction of the formula from its mere concept. If such a gap exists, then Kant cannot move from the rational but uninformative demand that a maxim conform to universal law yielded by the *concept* of a categorical imperative, to the action guiding requirement that a maxim be able to include itself as a universal law. The problem has been treated at length in a recent article by Henry Allison,<sup>3</sup> in which he concedes that a significant gap does exist in Kant's derivation of the formula in the *Foundations of the Metaphysics of Morals*. However, Allison counters the weight of this objection by noting that Kant presents a successful deduction of the formula in the *Critique of Practical Reason* through his injection of an additional premise, that of transcendental freedom. In this paper I will show that no such additional premise is needed, and that the deduction that Kant presents in the *Foundations* is successful as it stands. The paper will be divided into three parts: part one provides a general account of Kant's derivation of the formula of the categorical imperative; part two is devoted to developing the problem of the gap and Allison's proposed solution to it, and in part three I show what is wrong with Allison's proposal, as well as develop my own account of how Kant's derivation in the *Foundations* is not flawed by a gap.

## I

A crucial feature of Kant's ethics is his contention that the moral law must be determined independently of our inclinations and desires. A principle of action formulated in order to achieve a *given* end is binding only insofar as that end is desired, and if one were to give up the end, there no longer would be any need for the corresponding practical principle. Since such a practical principle specifies that which is practically necessary in order to achieve a particular goal, once that goal

<sup>1</sup> Bruce Aune, *Kant's Theory of Morals*, (Princeton: 1979), pp. 29–30.

<sup>2</sup> Allen Wood, *Hegel's Ethical Thought*, (Cambridge: 1990), pp. 161–166.

<sup>3</sup> Henry E. Allison, "On a Presumed Gap in the Derivation of the Categorical Imperative," *Philosophical Topics*, Vol. 19, No. 1, pp. 1–15.

is no longer aimed at, the need for those actions leading to it specified by the practical principle disappears. Consequently, if all practical principles were formulated in the service of inclination, there could be no moral law necessarily binding all rational natures, since in such a scenario, whether an imperative was valid for a will would be contingent, dependent upon whether or not a given end was held in view. It would be quite possible for an individual to claim him or herself free from the constraints of those principles commonly taken to be moral ones: given that s/he does not have the corresponding sorts of desires, these principles are not binding on him or her. For Kant, on the contrary, “a law, if it is to hold morally, i. e., as a ground of obligation, must imply absolute necessity.”<sup>4</sup> Since a ‘constitutional fact’ about the desires we in fact have cannot be the basis of a necessary law having universal scope, such a law, if it is to necessarily bind all rational wills, must be established *a priori*. Key, then, to Kant’s project is the derivation of an *unconditioned* law that can be action-guiding. Insofar as such a law is understood as constraining a will whose principles are not necessarily in accord with it, it functions as a command, and its formula is that of a categorical imperative.

If Kant’s project is to succeed, then, it must be possible to derive an action-guiding formulation of the moral law completely *a priori*. Such an *a priori* derivation can occur only through an analysis of the *concept* of an unconditioned practical law, since only an unconditioned practical law is by definition free of any empirical content. All premises playing a significant role in the derivation must be therefore strictly entailed by the concept. Were Kant’s deduction to fail, the whole of his ethical theory would be invalidated, since only such an action-guiding formula can determine that which is unconditionally and practically necessary.

Kant’s deduction of the formula of the categorical imperative from its “mere concept” really contains only two major premises from which the formula is itself derived; these premises are based upon that which is contained in the concept of a categorical imperative. The derivation of its formula occurs in a very difficult passage in the *Foundations*:

If I think of a hypothetical imperative as such, I do not know what it will contain until the condition is stated [under which it is an imperative]. But if I think of a categorical imperative, I know immediately what it contains. For since the imperative contains besides the law only the necessity that the maxim should accord with this law, while the law contains no condition to which it is restricted, there is nothing remaining in it except the universality of law as such to which the maxim of the action should conform; and in effect this conformity alone is represented as necessary by the imperative.

There is, therefore, only one categorical imperative. It is: Act only according to that maxim by which you can at the same time will that it should become a universal law.<sup>5</sup>

<sup>4</sup> *Foundations of the Metaphysics of Morals*, trans. by Lewis White Beck, (New York: 1985), pp. 5–6; [389]. References to the Prussian Academy edition pagination are noted in brackets.

<sup>5</sup> *Ibid.* p. 44; [420–421].

The first premise in the deduction is simply an elucidation of the nature of imperatives which cannot be categorical, namely, hypothetical imperatives, whose bindingness is contingent, relative to given ends, and which are thus not necessary in and of themselves.<sup>6</sup> The second premise is the definition of a categorical imperative as containing a law, necessary in itself and thus binding on all rational wills. Since hypothetical imperatives are conditioned in such a way that desired ends, undetermined as to their practical necessity, restrict the scope of their validity, a categorical imperative must abstract from all such ends if it is to be absolutely binding as such. As Kant notes, one cannot know what a hypothetical imperative will contain until the condition under which it is practically binding is stated. On the other hand, since the categorical imperative is necessarily binding in itself, it must be possible to know that which it contains before any such empirically given condition, dependent on what an agent in fact has as its end, is stated. But once all such ends are abstracted, we are left only with the form of a categorical imperative, i. e., its form as an unconditioned law, namely, necessity and universality.<sup>7</sup>

But how can the mere form of lawlikeness constitute an imperative? This question has troubled many of Kant's readers, who have accused him of an empty formalism in his ethical theory, one which has no determinate content, and which, as such, cannot be action-guiding.<sup>8</sup> But in fact the formal requirement that the

<sup>6</sup> For a discussion of the nature of hypothetical imperatives, see Henry E. Allison, "Morality and Freedom: Kant's Reciprocity Thesis," *The Philosophical Review* 95 (1986), pp. 399–400, as well as Thomas E. Hill, "The Hypothetical Imperative," in *The Philosophical Review* 82 (1973), pp. 429–450, reprinted in Hill's book, *Dignity and Practical Reason*, (Ithaca: Cornell University Press, 1992), pp. 17–37.

<sup>7</sup> The essential feature of law is universality; precisely insofar as such a law applies to all possible cases and admits of no exceptions, it is a priori and necessary. This essential feature of laws, which abstracts from their content, is what Kant calls the *form* of law.

<sup>8</sup> The charge that Kant's conception of the moral law is empty has been put forth by many of Kant's critics, most notably Hegel and John Stuart Mill. These charges, however, usually concern the formula of universal law itself, and have been shown to be unsubstantiated in recent scholarship, for instance, by Christine Korsgaard in her article "Kant's Formula of Universal Law," *Pacific Philosophical Quarterly* 66 (1985), pp. 24–47, and by Onora Nell (O'Neill) in her book *Acting on Principle* (New York: Columbia University Press, 1975). Allen Wood, in his excellent article "The Emptiness of the Moral Will," *The Monist*, Vol. 72, No. 3, July 1989, pp. 454–483, focuses on another feature of Hegel's charge against Kant which he believes to be more central: every maxim must have an end, and insofar as the end is taken as the determining ground of the will, it threatens the purity of the moral motive. Hence the description of a morally worthy action as being from duty alone excludes a description of that action as geared toward a particular end. But failing our ability to describe an action as geared to a particular end as a *relevant* act description for its counting as morally worthy, just any particular act will do. As Wood puts it, "Thus, to have a pure will is to act in a way that is so alienated from every concrete consideration that it must be quite indeterminate what I do." (pp. 477–478). This more subtle critique also ignores the force of thinking of the CI as second order principle which selects among maxims, an understanding of the CI which I will develop more fully below. As such a second order principle, it tests a maxim and the end towards which it is directed to see whether it can include itself as a universal law. Clearly, it is both the particular maxim and the end it

categorical imperative imposes upon the will, namely that its maxims be universalizable, while itself devoid of all material content, does not come up empty handed in its capacity to determine the will to action. This is because it assumes the agent's subjective principles of action (i. e., maxims) as its "matter" and requires that only those maxims that are universalizable be acted upon.<sup>9</sup> A categorical imperative, insofar as its mere concept is concerned, is one which commands: 'Conform your will to universal law as such.' Such a command seems empty enough by itself. However, once the premise is added (a premise that Kant takes for granted in the passage quoted above) that the will in fact acts according to subjective principles or maxims, i. e., self imposed general policies through which the will exercises its causality, then it becomes clear that such a categorical command functions as a second-order principle which selects among an agent's maxims. In commanding that the will conform to universal law as such, it commands that only those maxims that are universalizable be acted upon. The demand for universalizability in the subjective principles of the causality of the will, as deduced from the mere concept of a categorical imperative, is thus given its most precise expression in Kant's first formulation of the categorical imperative: "Act only according to that maxim *through* which you can at the same time will that it should become a universal law."<sup>10</sup>

## II

However, Bruce Aune<sup>11</sup> and Allen Wood<sup>12</sup> have given a new twist to the emptiness charge. According to Aune, while Kant's formulation of the categorical imperative itself may be action guiding, there is a crucial gap in Kant's derivation of the formula from its mere concept. Because such a gap exists, Kant cannot legitimately move from the concept, itself unable to provide a guide to action, to the more fruitful formulation of the categorical imperative. If this gap cannot be bridged, then Kant's ethics remains empty, unable to provide any significant criterion for moral action.

contains which is tested; hence the description of both the maxim and the end *are* relevant factors in determining the moral worth of a principle on which an action is performed.

<sup>9</sup> As H. J. Paton notes, "Kant at once makes it clear that there is no question — as is sometimes supposed — of deducing particular duties merely from the empty form of universal law. On the contrary, we have to consider the *matter* which has to be fitted into this empty form. The matter consists of our ordinary material maxims based on inclination for definite objects; and what we have to do is to accept or reject these maxims by the principle of universality", H. J. Paton, *The Categorical Imperative*, (Chicago: The University of Chicago Press, 1948), p. 73.

<sup>10</sup> *Foundations, op. cit.*, p. 44; [421].

<sup>11</sup> Bruce Aune, *op. cit.*, pp. 29–30.

<sup>12</sup> Allen Wood, *op. cit.*, pp. 161–166.

Aune's analysis focuses on Kant's transition from the claim that since no condition restricts the validity of a categorical imperative, nothing remains to be thought in it "except the *universality of law as such to which the maxim of the action should conform*," to his full blown formulation of the categorical imperative. The requirement that a maxim conform to universal law as such, according to Aune, is not equivalent to the imperative that a maxim of action be universalizable. Aune questions what it means for a maxim to be *consistent* with universal law, showing that the perspicuity of the phrase is not something that can be taken for granted. He claims that in order for Kant's derivation to work, it must be established that a maxim is consistent with universal law "when and only when its *generalization* could be willed to be a universal law."<sup>13</sup> The problem, as he sees it, is a certain disparity between the requirement that a maxim conform to universal law and the requirement that the *generalization* of that maxim conform to universal law; only the latter will provide a guide to action. While a generalized maxim's conformity with universal law guarantees that the maxim itself will conform to such a law, the converse is not the case, that is, a maxim may conform to universal law while its generalization may fail to do so. So Aune: "for all we know, the only general law bearing on the maxim *m* might be to the effect that if a person *x* acts on *m*, at least one person *y* must not do so. In this case *m* itself (or a particular action on *m*) may be perfectly consistent with universal law even though *G(m)* is patently incompatible with it."<sup>14</sup>

In other words, the requirement that a maxim conform to universal law could be interpreted in such a way that it comes up empty as a guide to action. We may, for instance, merely be requiring that the maxim be consistent with universal laws of logic and of nature; a maxim's conforming to universal law in *this sense* will not guarantee that its generalization will conform to universal law. It is, however, only the latter, stronger requirement that provides moral guidelines, and hence a gap in Kant's derivation remains.

In his article, "On a Presumed Gap in the Derivation of the Categorical Imperative", Allison has sharpened and brought into focus the exact nature and potential seriousness of this gap. He criticizes Aune's analysis of the problem while agreeing with him that the requirement that a maxim conform to universal law as such is not *prima facie* equivalent to the imperative that a maxim of action be universalizable: the failure of this equivalence rests on the equivocal nature of the meaning of the requirement that *a maxim conform to universal law*. Allison charges that Aune does not have the problem quite right on two related counts. First, "If a maxim is to be tested for its conformity to law by means of its generalization, then, presumably, the law in question must be conceivable as the product of the generalized maxim. Otherwise ... there would be no way to relate the maxim to the law."<sup>15</sup> In

<sup>13</sup> Aune, *op. cit.*, p. 88.

<sup>14</sup> *Ibid.*, p. 89.

<sup>15</sup> Allison, *op. cit.*, p. 7.

defense of Aune, it might be pointed out that what is at issue in the first place is precisely whether the requirement of a maxim's conformity to law is *equivalent* to the requirement that it conform to law by means of its generalization. Allison's second point is, however, much more devastating to Aune's critique: Kant's categorical imperative does not require that one be able to will the *generalization* of one's maxim as a universal law, it tests, rather "whether one's maxims are those 'through which' one can 'at the same time will' that they should become universal laws."<sup>16</sup> The two are not equivalent; in fact, Kant's formulation of the categorical imperative provides us with a much stronger requirement than Aune's. Were we to understand the categorical imperative in Aune's weaker sense, we would be unable to derive a contradiction in conception from the generalization of the maxim of false promising: if everyone made false promises, the practice of making promises would simply die out. On the other hand, there is a contradiction in acting on the policy of false promising while at the same time willing that such a policy be adopted by everyone else, since one would be acting on a policy that, when universalized, would make impossible the employment of that very policy itself.

Yet as Allison notes, these shortcomings in Aune's analysis do not impair the force of his claim that there is a gap in Kant's derivation. The problem is expressed more adequately by Wood: at this point the universality requirement demands only "universality of applicability", that is, it requires only that an agent adopt only those policies which s/he would concede are rational for any other agent to adopt in similar circumstances. The requirement that a maxim conform to universal law may amount to no more than the rational but uninformative demand that maxims of action express the *objective* practical sufficiency of means to ends. For example, a rational egoist acting on the maxim to lie when in difficulties may easily concede that since such a policy often does serve to extricate people from hardship, it is a rational one for any individual in relevant kinds of binds to adopt. The rationality of such a maxim, and hence its putative universality, rests on its expression of the objective practical sufficiency of means to ends.

But the universality demand that Kant requires in order to arrive at the formula of a categorical imperative is of a different kind. As Allison notes, the claim that Kant needs to justify in order to arrive at the formula is the requirement that one's "maxims be capable of including themselves when regarded as universal laws,"<sup>17</sup> in other words, no contradiction should result if the maxim is taken as a practical principle adopted by everyone *while at the same time* being adopted by the agent.<sup>18</sup>

<sup>16</sup> *Ibid.*, p. 7.

<sup>17</sup> *Ibid.*, p. 7.

<sup>18</sup> The point that the word "zugleich" in Formula I of the categorical imperative must be taken seriously has been made by numerous commentators, including Allison (*Ibid.*, p. 7). In his article "The Categorical Imperative," Thomas Pogge, for instance, explains the significance of the term when he notes: "... the presence of the word [zugleich] must indicate a further point, which is, quite literally, that the agent must be able to will his maxim *together with* its universal permission: *M*'s universal availability must not make it impossible to will (to adopt) *M*." Although this citation also reflects Pogge's rather atypical reading of the categorical imperative as requiring that an agent be able to will that everyone be *permitted*

This is a much stronger universality requirement, one which, according to Allison, cannot be derived from the weaker demand that maxims conform to universal law. While it is true that a maxim capable of including itself when regarded as a universal law conforms to such a law, not all maxims conforming to universal law in the weaker sense outlined above are capable of such self-inclusion.

It may be objected that the weaker requirement, Wood's "universality of applicability" is surpassed by Kant's earlier insistence that a categorical imperative must abstract from all ends which make an imperative only conditionally binding, i. e., its bindingness cannot depend on the matter of willing. But this abstraction from all material ends may still not serve to fill the gap in Kant's derivation, for, as Allison notes: "For all that we have seen so far, an unconditioned practical law might require nothing more than that one's maxim be one on which it would be reasonable for any agent to act in relevantly similar circumstances. As a universal requirement, this principle applies to all agents, independently of the empirical content of their maxims; but it is quite distinct from the substantive requirement expressed in the categorical imperative that one's maxim be able to include itself as a 'principle establishing universal law.'" <sup>19</sup>

Allison admits that a problematic gap exists in Kant's derivation of the formula for a categorical imperative as it is presented in the *Foundations*. He solves the problem by referring to the derivation of the formula as presented in § 5 and § 6 of the second *Critique*. According to Allison, Kant there derives the formula of the categorical imperative from the premise of the concept of a categorical imperative *along with* the premise of transcendental freedom. Very briefly put, a transcendently free agent is not only not causally determined by the mechanism of nature; s/he is also free in the sense that s/he can determine the will *independently* of motives having to do with the agent's sensuous nature. Mere contracausal freedom is not enough, since it implies only that an agent is free in its choice of means to ends which are themselves empirically given. <sup>20</sup> What is required as a premise in order to fill the gap in the derivation is the much stronger sense of freedom—transcendental freedom, i. e., independence from sensuously conditioned motives in the determination of the will. This independence implies that agents do not simply *have* ends, but rather, these ends are incorporated into their maxims through a transcendently free act. Allison notes that if ends are themselves freely set, then the mere

to adopt her maxim, which is not what is at issue here, it does elucidate quite nicely what it means to say that a maxim must be able to include itself when universalized. Thomas W. Pogge, "The Categorical Imperative," in *Grundlegung zur Metaphysik der Sitten: Ein kooperativer Kommentar*, Otfried Höffe (Frankfurt am Main: Vittorio Klostermann, 1989) pp. 174–175; Cf. Onora O'Neill, *Acting on Principle*, p. 78 and Christine Korsgaard, "Kant's Formula of Universal Law", pp. 24–27.

<sup>19</sup> *Ibid.*, pp. 9–10.

<sup>20</sup> See Allison, "On A Presumed Gap ..." *op. cit.*, p. 10. Cf. Allison's account of practical freedom in "The Concept of Freedom in Kant's 'Semi-Critical' Ethics," in *Archiv für Geschichte der Philosophie*, Band 68, Heft 1, 1986, pp. 96–115, as well as chapter 3 of his *Kant's Theory of Freedom*, (Cambridge, Cambridge University Press: 1990).

presence of a sensuous inclination “is not a sufficient or justifying reason for acting on the basis of it. On the contrary ... such a reason could be provided only by a universally and unconditionally valid rational principle of choice, that is to say, a practical law.”<sup>21</sup> In other words, if we are capable of acting independently of desires that we happen to have, then the mere presence of a desire cannot provide a fully adequate account of *why* the end that fulfills these desires is of value, and thus why such an end can or should be chosen. If a course of action is chosen only because it relates to a desire that a person happens to have, its worth is relative only to such a person and its relation to others is left undetermined. Only a universally and unconditionally valid rational principle can determine the worth of both maxims and actions relative to the intersubjective community of rational knowers, that is, only the testing of a maxim by such a principle can provide a sufficient reason for action.

Armed with the premise of transcendental freedom, Allison proceeds to complete the deduction of the formula for a categorical imperative. Given this premise, a maxim is rationally justified just in case it is adopted, at least in part, *because* it conforms to universal law. But his requirement rules out, or rather surpasses, the weaker universalizability requirement that a maxim merely express the objective practical sufficiency of means to ends. So Allison: “... if I am required to adopt a maxim at least in part because of its conformity to universal law or, equivalently, an unconditioned practical law, then, clearly, this maxim must be able to include itself as a “principle establishing universal law,” which is just to say that the maxim must have legislative form”.<sup>22</sup>

Yet, as Allison himself notes, the rational egoist may still claim that part of the reason he adopted a certain maxim was that it was expressive of means objectively sufficient to his end. However, this objection disregards the basic thrust of Allison’s argument, which turns on the nature of the *rational justification* for the adoption of a maxim. Whether or not a maxim conforms to universal law is a function of whether its having worth can be determined intersubjectively, that is, whether its value is not determined merely on the basis of the desire of a particular agent. Thus the rational justification requirement for the adoption of a maxim guarantees that a principle of action is not adopted *simply* on the basis of a given desire, since it requires that everyone could will it. Only such a universalizability requirement tests a maxim to see whether sufficient reasons can be given for its adoption. If they cannot be provided, then its value is *merely* conditioned, dependent upon given desires. For instance, as Allison observes, a maxim of false promising possessing Wood’s universality of applicability “is deemed reasonable in the first place only because of certain presupposed ends, which derive whatever justification they might possess from the agent’s desires. Consequently, these desires and not the intrinsic reasonableness or lawfulness of the policy function as the ‘determining ground of

<sup>21</sup> *Ibid.*, p. 11.

<sup>22</sup> *Ibid.*, p. 11.



the will' ”<sup>23</sup> If on the other hand, a maxim can be universalized, its value has been shown not to depend *simply* on the given desires of the particular agent in question, but rather, has been shown to be capable of being willed by all.

### III

There are, however, some difficulties with Allison's analysis. It ignores the full significance of Kant's claim in the second *Critique* that “though freedom is certainly the *ratio essendi* of the moral law, the latter is the *ratio cognoscendi* of freedom.”<sup>24</sup> If Kant is correct in making this claim, then in injecting transcendental freedom into the picture in order to derive the formula for a categorical imperative, Allison has added an unnecessary extra premise. If transcendental freedom is already implied by the *concept* of an unconditioned practical law insofar as it is a condition of the possibility of such a law, then it is not an additional premise needed in order to overcome the gap. If I am justified in making this claim, then no such gap exists in Kant's derivation of the formula as presented in the *Foundations*.<sup>25</sup>

Allison cannot have it both ways. He cannot admit the validity of Kant's claim that “freedom and unconditional practical law reciprocally imply one another,”<sup>26</sup> while at the same time conceding that a gap exists in Kant's derivation of the formula in the *Foundations*. If there is any gap at all in Kant's derivation, it occurs at the level of his claim that the very possibility of an unconditioned practical law requires transcendental freedom. In other words, only an analysis of the *nature* of an unconditioned practical law (in terms of an analysis of the conditions of its possibility) can establish such freedom as one of its conditions. Either Kant has failed in showing that transcendental freedom is a necessary and sufficient condition of a will that is bound by a categorical imperative, in which case his deduction of the formula for a categorical imperative fails *both* in the *Foundations* and in

<sup>23</sup> *Ibid.*, p. 12.

<sup>24</sup> *Critique of Practical Reason*, trans. by Lewis White Beck, (New York: 1956), p. 4; [5].

<sup>25</sup> It may be readily objected that my own analysis, which uses some of Kant's formulations in the second *Critique* to defend the validity of Kant's deduction in the *Foundations*, ignores the controversy over whether a major reversal in Kant's thinking occurred between the time that he wrote the *Foundations* and the time he wrote the second *Critique*. Whereas in *Foundations* III Kant establishes the reality of freedom in order to establish the validity of the moral law, in the second *Critique* Kant begins with the moral law as a fact of reason, and from this fact proceeds to establish the reality of freedom. This controversy, however, does not touch the issue with which we are concerned here, and with which Kant is preoccupied in *Foundations* II. There Kant shows that the *possibility* of a valid unconditional practical law presupposes transcendental freedom. It is with this question that we are here concerned, and Kant's position on this point is relatively the same in both the *Foundations* and the second *Critique*.

<sup>26</sup> *Critique of Practical Reason*, *op. cit.*, p. 29; [29].

the *Critique of Practical Reason*, or, if Kant has been successful in showing that transcendental freedom is a condition of the possibility of the validity of a categorical imperative, then both deductions succeed.

Kant's categorical imperative is itself a test for maxims; as such a test, it is a meta-level principle. Now, what Allison has shown is that a *maxim* such as that of false promising, possessing only universality of applicability, is conditioned since its adoption depends *only* on the presence of given desires; if we assume that an agent is transcendently free, the individual acting on such a maxim does not have sufficient reasons for his or her behaviour. However, if we are going to show that transcendental freedom is a condition of the bindingness of an unconditioned practical principle, we must work at the level of the meta-principle itself, that is, we must show how the *unconditioned* nature of a categorical imperative can be fleshed out such that it will yield a meta-level principle requiring maxims to be able to include themselves when regarded as universal laws. Indeed, only when a maxim passes this kind of universalizability test can the agent who adopts it have sufficient reasons for action.

As noted above, the categorical imperative is a meta-level principle. So, for that matter, is Wood's requirement that a maxim possess universality of applicability; both principles are maxim testers. What we need to show clearly is how an analysis of the concept of an unconditioned law is going to yield the stronger meta-level requirement that a maxim be able to include itself when regarded as a universal law. This is the only way to overcome the gap once and for all. All we have to work with, then, is the precise meaning of *unconditioned* and how it applies to each of the two meta-principles that are possible candidates as unconditioned laws. We need to be able to show that *only* the meta-principle that a maxim be able to include itself when regarded as a universal law is unconditioned.

Both the stronger universalizability requirement and Wood's requirement of universality of applicability are capable of selecting among maxims. Moreover, both requirements seem to be unconditioned, independent of the particular ends that an agent adopts. The requirement that means be adequate to one's end applies whether one wishes to extricate oneself from a bad situation through a false promise or whether one wants to become a doctor in order to cure the sick. Its applicability thus applies regardless of the particular end in question. Yet while the stronger universalizability requirement is capable of throwing out those maxims having *only* conditioned worth, the weaker is not, but only throws out maxims that do not specify means adequate to their end. This tells us that the former requirement is independent of *given* ends adopted on account of sensuously conditioned desires in a way that the latter is not. In fact, the latter requirement *cannot* select among ends precisely because it is not really unconditioned but only seems so; the *given* end upon which it depends corresponds to a second-order desire, that is, the desire to fulfill one's desires. It thus has the form of a hypothetical imperative: if you desire to fulfill your desires, then your means should be objectively practically sufficient to your end. The fact that it depends on such a second order desire is what makes it

*seem* unconditioned, since adoption of the principle would not seem to depend upon any *particular* end. Maxims possessing only “universality of applicability” *are thus binding only on agents having the end in question*; in other words, these maxims are adopted by the agent *because* of an end, and not because they conform to universal law. On the other hand, the mere concept of an unconditioned practical law dictates that all contingent ends be abstracted from the principle determining that which is objectively necessary in rational willing, and it is this abstraction from such ends which guarantees its unconditioned validity.

Because the requirement of universality of applicability is not unconditioned, it remains undetermined as to its *moral* worth. The end of fulfilling one's desires is of value to an agent simply because this end is desired;<sup>27</sup> but this gives us very contingent grounds for assigning worth to an end.<sup>28</sup> Allison is surely right when he claims that if we are transcendently free, appeal to our sensuously conditioned desires cannot provide sufficient reasons for action; only appeal to an unconditioned practical law can provide these reasons. And this once again brings to the fore the mutual interdependence of the concepts of transcendental freedom and unconditioned practical law: only a transcendently free will can be bound by the latter.

If my discussion above is correct, then Kant's analysis of an *unconditioned* practical law is sufficient to move to the stronger universality requirement expressed in his formula of a categorical imperative. It is also the implications of the *unconditioned* nature of such a law which imply that if such a law is to be possible, we must then be transcendently free. Since an unconditioned practical law contains no condition to which it is restricted, its bindingness cannot depend on any contingent ends. But this in turn implies that such a law can be valid of our wills just in case pure practical reason can determine the will, i. e., just in case the universality of law as such can be the ground of our adoption of a maxim. As Kant notes in the *Foundations*, an unconditioned practical law is one “the conception of which must determine the will without reference to the expected result.”<sup>29</sup> And this means

<sup>27</sup> This is stressed by Kant: “The ends which a rational being arbitrarily proposes to himself as consequences of his action are material ends and are without exception only relative, for only their relation to a particularly constituted faculty of desire in the subject gives them their worth” *Foundations, op. cit.*, p. 52; [427].

<sup>28</sup> In his article “Kant's Argument for the Rationality of Moral Conduct,” Thomas Hill makes a related and illuminating point. If rational agents were to act only on hypothetical imperatives, “selecting their ends according to their inclinations or for no reason at all, [they] would never act for *sufficient* reasons, as ... Kant thought a rational free will could do.” [p. 14] In *Pacific Philosophical Quarterly* 66 (1985), pp. 3–23; italics mine; reprinted in *Dignity and Practical Reason, op. cit.*, pp. 97–122. Cf. Andrews Reath, “Kant's Conception of Practical Rationality,” in *The Monist*, Vol. 72, No. 3, July 1989, esp. pp. 394 ff., where Reath discusses how Kant's conception of practical rationality is tied to the CI and how the CI involves justifications that are complete.

<sup>29</sup> *Foundations, op. cit.*, p. 21; [401–402]. I am grateful to Henry Allison for having called my attention to this passage.

that if an unconditioned practical law is to be binding on our wills, we must be transcendently free.<sup>30</sup>

While Allison's analysis does show that maxims possessing only universality of applicability cannot receive full rational justification, it is not at all clear that he overcomes the gap at the appropriate meta-level. And while his insightful analysis of transcendental freedom helps to uncover the deep structure of unconditioned practical laws, he injects it as an additional premise in his derivation. I hope to have shown that such a premise adds nothing that is not already contained in the concept of an unconditioned practical law. If my own analysis is correct, then Allison has conceded too much both to Aune and to Wood in admitting that there is a substantial gap in Kant's deduction as it is presented in the *Foundations*.

<sup>30</sup> An earlier version of this paper was presented at a colloquium on the categorical imperative at the 1993 Pacific Division Meetings of the A. P. A. in San Francisco, CA in March 1993. I am grateful to Henry Allison, Manfred Baum, Manfred Kühn, and Franklin Mason for their insightful comments on earlier drafts, as well as for the suggestions of an anonymous commentator.