“So Many Formulas”: The Relations Among the Formulas of the Categorical Imperative

Robert Guay
Department of Philosophy
Binghamton University
P.O. Box 6000
Binghamton, NY 13902

rguay@binghamton.edu
I. Introduction

Kant, having identified the formulas of the supreme principle of morality, offers a succinct explanation of their interrelation. What Kant says is, “The above three ways of representing the principle of morality are at bottom only so many formulae of the very same law, and any one of them of itself unites the other two in it.”\(^1\) This claim – hereafter the “Unity Claim” – plays the role of the eccentric cousin in the family of Kant’s ethics: although glaringly present, it is little spoken of, but seldom disowned. Most commentators, at any rate, focus their attention on more important matters, such as the content of the individual formulas, the moral psychology, or the deduction of freedom. Such matters are sufficiently absorbing to leave the Unity Claim often passed over without remark.

But the Unity Claim should not be ignored. Kant does assert it, which compels us to attempt to find a place for it in his moral theory. It would seem to constrain the interpretation of the other, more momentous issues. How one interprets the content of the categorical imperative, in particular, would seem to be significantly restricted by the Unity Claim; one could not, given the Unity Claim, offer a complete interpretation of any single formula without also at least referring to the other formulas. And, as I shall argue in Part III below, the Unity Claim is no accident. Kant is committed to the Unity Claim by virtue of some basic features of his moral theory.

This paper will thus offer what amounts to an extended commentary on the Unity Claim. I shall review the various suggestions of what it might mean, and how it might, or might not, be accommodated within Kant’s moral theory. The structure of this paper will be as such. Part II will examine the two main strategies for including the Unity Claim within Kant’s moral theory, and explain why they are both inadequate. Part III will examine the other main approach to the
Unity Claim: giving up on it. There I shall explain why that approach, too, is unacceptable. Part IV will conclude with some remarks about what the impossibility of either including or rejecting the Unity Claim implies about the viability of Kant’s moral theory.

II. Inclusionist approaches

There are two main strategies for including the Unity Claim within Kant’s moral theory: the accommodative and the prescriptive. According to the former, the three formulas are, as a matter of fact, in some sense one and the same, or at least unified. The latter strategy argues that the Unity Claim is not a statement of fact, but rather an injunction to render the content of each formula so that it can be reconciled with the content of the other two, or so that the content of the three formulas can be brought into a systematic unity. I shall first review the several different accommodative approaches that have been proposed, by Kant and by others, and after discussing why each one of them fails, turn to the prescriptive approach.

Kant himself, in the very making of the Unity Claim, offers the basis of the accommodative strategy. Recall that Kant says of the three formulas that “any one of them itself unites the other two in it”(436). Each one of the formulas, in and of itself, “unites” the others “in it,” so whatever content might be present in a given formula is already present in the others; there is no need for us to do the work of reconciling the content of disparate formulas. This, of course, leaves the reader the task of determining just what such a unity within each formula could consist in. But it would seem to offer the most straightforward sense of the Unity Claim. Two – indeed, any two – of the three formulas are superfluous, because each is a complete expression of the categorical imperative.
An initial problem with this reading is that there are many more, and slightly fewer, than three formulas of the categorical imperative. There are many more than three formulas in that Kant offers more than one version of what is supposed to be a single formula. The famous Formula of Universal Law, for example, appears in slightly different versions at 402 and 421, in yet another version in the *Critique of Practical Reason*,³ and in a completely different version as “act as if the maxim of your action were to become by your will a universal law of nature”(421; emphasis removed). There are slightly fewer than three formulas in that “the third practical principle of the will”(431), known as the Formula of Autonomy, does not seem to be explicitly stated. But this seems to be a relatively mundane difficulty. Figuring out which formulas are the three formulas might call for some exclusions, and possibly some reconstructive work, but perhaps nothing terribly daunting.

A more serious problem with the most basic reading of the Unity Claim is that Kant suggests a different, incompatible way of making sense of it. This other way is based on how Kant argues sequentially for the three main formulas by adducing new considerations; as Paton argues, “we can follow the relations between these formulas only as we study both them and the argument by which Kant passes from one to another.”⁴ After the introduction of the first formula, then, Kant calls on us to take into account the purposive character of action, and by thus supplementing the first formula, we arrive at the second formula. And then only by supplementing the second formula with a consideration of the systematic totality of all maxims do we arrive at the third formula. And as Guyer points out,⁵ this sequence is not only implicit in the derivations of the various formulas, but it is also presented by Kant immediately following the Unity Claim. There Kant explains that the three formulas are divergently “expressed” [ausgedrückt](436) according to the three features of maxims: “form,” “matter,” and “a
complete determination of all maxims.”(436)⁶ Here the unity among the formulas seems to be that, together, they contribute to a single progression that leads up to a fully adequate characterization of the demands of morality.

So already at this superficial level there is a conflict between two different accounts that Kant gives of the unity among the formulations of the categorical imperative: on one account the unity implies that the diversity is not strictly necessary, but on the other account the diversity is necessary, at least for arriving, with the addition of supplemental considerations, at the full specification of the moral law. Although the former account might be more casually presented, and the latter might have a more substantial body of argument behind it, Kant’s other remarks on the issue at *Groundwork* 436 would seem to support the former account. According to Kant, the difference among the formulas is only “subjectively rather than objectively practical”(436). This I take to mean that the difference among the formulas consists only in how one thinks about them, and not in anything about the formulas themselves. Kant later reinforces this idea by claiming that one formula provides the “strict method”(436) of moral appraisal, whereas the others, by contrast, provide intuitive access to the same moral law: they make what is abstract easier to think about.⁷ There is, however, no significant difference in content; the difference among the formulas lies primarily in how we make use of them.

There might be some way partially to reconcile these two accounts. One could, for example, start from the latter account, and argue that the supplemental concerns that allow the progression from the first to the third formula are basic features of practical reason in general. Then one could still, perhaps, affirm the former account with only a minimal qualification. One could say, that is, that the former account, that each formula unites the others in it, is correct as long as some basic features of practical reason are presupposed. This would amount to a
concession that the former account is inaccurate: a given formula does not contain the others. But it would leave room for the claim that all one needs in order to make a given formula contain the others is some implicit assumptions that one would need to adopt in a consideration of rational agency, anyway.

I do not think that solution would work, because I see neither how the first formula could get started without a great deal about practical reason already built in, nor how those supplemental considerations could be that which would bring each formulation into unity with the others. But this is not the main issue here. The main issue is that even if there were a resolution between these two accounts of the unity of the formulations, there would still be a question to ask about what the unity consists in. Whether the unity in question is self-contained or progressive, or both, what it means for diverse formulas nevertheless to be a single law is unanswered. There must be some specific sense in which many can be one; or, to consider it from the opposite direction, some explanation of how the unity of the moral law relates to all of its expressions. The relation in question is not logical equivalence. Consider two, I believe uncontroversial, examples of formulas of the categorical imperative:

Act only in accordance with that maxim through which you can at the same time will that it become a universal law. (421; emphasis removed)

So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means. (429; emphasis removed)

Even if these imperatives were to be rendered into truth claims, they would certainly not imply one another. And even if they were mutually implying, that would still leave open the question of how they relate to the moral law. The simplest answer is perhaps to say that they are the
moral law. But that conflicts with a more obvious answer, that they are *formulas of* the moral law; but that provokes the questions of whether and how a law can have formulas different from itself. So how do these formulas relate to the moral law and to each other?

One account of these relations is provided by the Unity Claim itself. According to the Unity Claim, the various formulas are “ways of representing” the principle of morality. And this makes a certain amount of sense. The supreme principle of morality is, one presumes, not essentially in German, at least not in the way that the *Koran* is essentially in Arabic. So the moral law might best be treated as an abstraction that stands apart from any possible linguistic representation of it; the formulas might then be formulas of the same law in this sense. This, I take it, is the point of Allen Wood’s suggestion that the formulas do not provide different moral principles, “but only three different ways of regarding one and the same principle.” But this does not work: it is not clear what the “one and the same” principle is, since the formulas are not semantically equivalent. We cannot attribute the diversity of *these* formulas of the single moral law to the constraints of linguistic representation, because their differences far transcend what could be caused by the indefiniteness of language.

Allen Wood offers another metaphor for making sense of Kant’s claim that the formulas are “ways of representing” the moral law: “the formulas,” according to Wood, “arise from considering the categorical imperative from different sides.” What “sides” mean in the case of the categorical imperative is unclear, but perhaps we can take as an example Kant’s use of the terms “principle of morality,” “moral law,” and “categorical imperative.” The principle of morality considered as universally legislating is the moral law, and considered as commanding beings like us is the categorical imperative. So this seems to be a legitimate case in which three different formulas are nevertheless, under certain conditions or with certain qualifications,
equivalent. But this still does not help us with the three formulas of the categorical imperative. The principle of morality, the moral law, and the categorical imperative possess something, unavailable to the formulas of the categorical imperative, that renders an explanation of their unity easily available: namely, that they are the same thing. The problematic diversity of the formulas cannot be resolved by appeal to “ways of representing,” because the “ways of representing” appeal requires just what the formulas seem to lack: the availability of some unproblematic sense of unity.

Another suggestion of how to explain the unity among the formulas is provided by Allen Wood: “the three formulas together are supposed to complement one another.”\textsuperscript{11} Here we might think of Aristotle’s notion of happiness, which is composed of many elements, each of which is part, and not merely a means to, the final good. Similarly, Wood suggests, each formula fills out part of the principle of morality, and taken together they completely specify it. But this, so interpreted, would count as a renunciation of the Unity Claim. On this reading, no formula would unite the other two in it. Rather than each being able to stand in for the whole, each is merely a non-overlapping part of the whole. This reading thus leaves us with three formulas, each of which is independently significant precisely because of its difference from the other two.

This brings us to perhaps the most promising accommodative approach, and the one that many interpreters seem to presume is Kant’s sense of the Unity Claim.\textsuperscript{12} We might interpret the Unity Claim as meaning that the various formulas are practically (or “extensionally”) equivalent, in that they impose the same moral demands. The characterization of “practical principles” as “propositions which contain a general determination of the will, having under it several practical rules”\textsuperscript{13} seems to support this practical equivalence reading. So the unity at issue does not pertain to the formulas themselves, which are semantically quite different, but to what they
produce: the same set of particular imperatives in each case. This does not quite give us the means to cash out Kant’s claim that each formula unites the others in it. But it does at least give us a way of determining whether or not the Unity Claim is true. It is true if and only if each formula generates the same set of demands.

Arthur Melnick has offered a reading of the content of the categorical imperative that includes a very thoroughgoing accommodation of the Unity Claim. Melnick reads the Unity Claim as requiring that each formulas be understood as “straightforwardly equivalent” to the others, which requires that each formula “turn on the very same concept of rationality,” which in turn requires that the formulas all “lead to the exact same restraints.” Turning on the same concept amounts to little, and what “straightforwardly” might mean is at issue. But the formulas are devices to generate restraints – or possibly affirmative duties as well – and what unites them is that they each produce the same result. “The other two formulations contain nothing more than the first as regards … the content of morality,” and that is what is truly important.

The respect that Melnick accords the unity claim in his interpretation of the content of the categorical imperative is, I think, admirable. But however plausible or implausible a reading of the categorical imperative Melnick can then generate, his approach – and indeed, any that relies on practical equivalence – would still be inadequate to support the Unity Claim, at least without substantial revisions in Kant’s moral theory. Unity, for Melnick, means that each formula produces the same results: it approves or disapproves of the same actions. But for Kant, obeying the moral law requires more than performing actions that meet certain formal constraints. Action, for Kant, is under a description, so moral action is action according to the moral law. And with this being so, any divergence in the mere formula of the law amounts to a change in the
demands on the moral character of action. Different formulas command differently, thus contradicting the Unity Claim, even if they lead to the same result.

A way of getting at the integral importance of the formula itself is to consider positive law. In any piece of legislation, the law in its particular formulation plays an ineliminable role in adjudication. Decisions based on that law can, of course, appeal to the deeper meaning of the law, the intentions behind the law, the foreseeable consequences of the law, the tradition of interpretation of the law, and so on, but none of these could entirely supplant one basis of ultimate appeal: the law as written. The letter of the law places a limit on deliberation that is not compatible with multiple formulations. And insofar as the categorical imperative is meant to serve as the basis for moral deliberation, the same issue arises. Multiple formulations provide for multiple deliberative paths, undermining the sense of there being a single law in question. This presents an interpretive challenge to practical equivalence: it is hard to see how any security could be offered for the claim that different formulas generate the exact same restraints, when they regulate deliberation differently and have a potentially infinite number of applications. This does not, however, undermine the availability of practical equivalence, but only its plausibility.

But the issue of multiple formulas is much more acute for Kant than it is for positive law. The relevant difference here is that, for Kant, mere compliance with the law is inadequate. To act morally, one must act not merely “in conformity with,” but “on” the law: one must adopt the reason that one does at least in part because of its morality. Moral action is thus a function of the reason that one in fact adopts in acting. And reasons are delimited by the terms of their expression: except in the case of semantic equivalence, acting on a nominally different reason is acting on a different reason. So different formulas of the categorical imperative call for different reasons to be acted upon, and thus impose different moral demands.
For Kant, action proceeds according to a *maxim* (cf., e.g., 406, 421), a subjective principle of will, which at the least prescribes a course of action in a given set of circumstances.\(^\text{18}\) Maxims constitute the minimum level of rationality for a bit of behavior to qualify as an action: being governed by principle, even if that principle is adopted on account of inclination. The character of a given action is accordingly given primarily by its maxim. But full rationality of an action lies in the relation of the maxim to the supreme principle of morality. The full account of an action, which allows for a characterization of it as moral, thus includes not only the maxim-description, but also something beyond that. This has a very familiar formulation in Kant: in acting morally one acts on the categorical imperative at the same time as one acts on one’s maxim. And this is why the practical equivalence reading the Unity Claim cannot hold in any straightforward sense. However superficially similar two actions may seem, they are different if they stem from different reasons. Adopting a maxim for the sake of an abstract principle of right, for example, would bring about a different action than adopting a maxim out of respect for an end would, even if the maxim and the behavior is the same in both cases. As the basis for acting varies, so does the character of the action itself. So difference in the formula that one acts upon constitutes a difference in reason, which constitutes a difference in action. The different formulas thus command differently. The mere nominal variation among formulas implies a diversity of standards for the moral assessment of action.

One could, it appears, make a small refinement to this picture and preserve the Unity Claim. As I have presented it, Kant would characterize an action’s principled character – the ultimate reason behind it – as integral to the action: the proper description of an action would be of the form “action A as permissible according to formula F”. The moral character of an action thus does not consist in the relation between an intrinsically amoral action and something else,
but is a feature of the action itself. This is what upsets the Unity Claim, since it leaves no wiggle room for different formulas to have the same moral implication. But, one might suppose, an easy fix would be to revise this account of the relationship between the categorical imperative and individual actions. If the role of moral principle were merely to approve or disapprove of particular maxims, and the moral character were something extraneous to the description of actions, then there would be a level at which diverse formulas could be united. The Unity Claim could be preserved in much the way that Melnick suggests.

But this would require a substantial revision in Kant’s moral theory. To view the categorical imperative merely as a test of maxims that stands external to acting would take away the sense that one could act according to the categorical imperative. Moral action – if this still means acting according to maxims approved by the categorical imperative – would then just be an approved form of pathological behavior, not anything that carries any special dignity. This would subvert the distinction that sustains the fundamental feature of Kant’s theory:

For, in the case of what it is to be morally good, it is not enough that it conform with the moral law but it must also be done for the sake of the law; without this, that conformity is only very contingent and precarious … (390)

Recuperating the practical equivalence reading of the Unity Claim would make moral goodness “contingent” and “precarious,” and thus destroy the possibility of any distinctly moral value. The accommodative strategy thus seems far too costly to rescue.

Another possible way to accommodate the Unity Claim would be to concede that not only the maxim, but also the maxim adoption, belongs to the proper description of an action, but to insist that whatever differences obtain between the formulas are nevertheless insignificant. For example, one could claim that the proper account of an action is simply that it is on maxim
adopted because of the categorical imperative, and not for any more specific reason. This incorporates the moral consideration into the action itself, and, as long as the result of the categorical imperative test is the same no matter what formula one employs, remains compatible with the Unity Claim. No difference among formulas can arise in the assessment of action, since the relevance of the individual formula gets subsumed under the description of the action as on the categorical imperative. This solution fails for two reasons, however. One is that this would seem to eliminate any deliberative role for the formulas. The very point of the formulas is to serve as the basis for moral deliberation – recall Kant’s distinction between “objectively” and “subjectively” practical – but what this proposal does is eliminate any way in which the formulas might figure in the account of an action. The second failing of this proposal is that it preserves the original problem. Its basic solution is that the multiple formulas are all subsumed under the one moral law, but it still lacks an account of what it could mean for many different formulas to be one law.

A final attempt at recuperating the accommodative strategy would be to follow Onora O’Neill in claiming that the formulas are not only practically, but also “intensionally” equivalent. This would seem to be, to use Melnick’s term, the most “straightforward” way in which the formulas could be equivalent: that they have not only the same extension, but even the same content. O’Neill’s version of intensional equivalence is slightly weaker than equivalence of content, however; she claims that two formulas “might be intensionally equivalent if it could be shown not merely that they in fact yield the same results, but that this result follows from the nature of the formulations.” This is a fairly weak criterion. All it seems to demand is practical (or “extensional”) equivalence plus an explanation for the practical equivalence; there need not be any other sort of equivalence. And indeed, as O’Neill explicates the intensional equivalence
of the formulas, they turn out to be “complementary,” “mutually determining,” and manifesting “two perspectives” that “can be combined.” These are all concessions that they have different content; if I am correct about the inadequacy of practical equivalence, they thus impose different demands on action. I think that there are also straightforward reasons why the formulas propose substantially different tests for the assessment of maxims. But even if it were conceded that the formulas were “intensionally equivalent” in O’Neill’s sense, that would still be inadequate to sustain the accommodative strategy. The accommodative strategy appears to be a failure.

Even the failure of the accommodative strategy leaves hope for the inclusionist approach. The prescriptive strategy accepts the Unity Claim, but understands it as a demand to make the three formulas one and the same. Here are two statements of the prescriptive strategy:

Kant wants these formulas to be equivalent (436), and I take this not as an assertion (as if each formula had an entirely clear meaning on its own), but as a prescription: the subsidiary formulas make distinctive contributions to the clarification and specification of the categorical imperative – they gradually enrich its meaning, until at last its full import can be understood. And once fully understood, the categorical imperative can then be read back into each of these formulas so as to make this equivalent as Kant demands. (Pogge 1989: 172)

Kant’s remark that the three formulas express “the very same law” might best be read as an invitation to understand each formula, and even its application to empirical circumstances, only in light of the others. Instead of making us look for an a priori
equivelancy proof, it should place a constraint on the way we interpret all three formulas.

(Wood 1999:187)

The Unity Claim here, then, does not pertain to anything intrinsic about the formulas themselves. Rather, it is a constraint on how the agent fills out the content of the categorical imperative: one must go back and forth across formulas until each comes to the same thing.

One way of understanding the prescriptive strategy, then, is as a weak form of practical equivalence. It concedes that the formulas do not, in themselves, mean the same thing, but then argues that they can be made to allow (or demand) the same maxims. There does not seem to be much textual support for this approach, but it does try to preserve as many of Kant’s claims as possible. As a weak form of practical equivalence, however, it would succumb to the same difficulty as practical equivalence. Even if the formulas came to be understood as each rejecting the same maxims, that would still leave them making different moral demands, thus contradicting the Unity Claim.

A perhaps stronger way of understanding the prescriptive strategy is as claiming that the individual formulas are not sufficiently determinate to guide action, but that their mutual interdetermination generates enough content to be action-guiding. This is again weak in textual support, but it does appeal to a very firm intuition about the categorical imperative: that it is tremendously unclear. So, we might imagine, the point of the Unity Claim is not that each formula independently commands the same, but that no formula commands at all except in conjunction with the other two. The unity of the formulas is necessary for there to be a single command at all. There are two problems with this interpretation. One is that it is not true that the individual formulas fail to command. They may not provide as much guidance as one would like, but they each provide at least some guidance on what not to do. The other problem is that
what is needed to fill out the indeterminacies of each formula is not more formulas. The interpretive challenge presented by each formula is, at the least, determining which maxims it approves or disapproves of. What would be useful here, then, is propositions that connect the maxim-description of actions to the terms of a formula. Another formula could introduce an additional moral consideration that might focus one’s understanding of a particular situation, or even, insofar as interpretation is an activity, place a moral stricture on how one interprets the law. This might suggest a useful model of moral deliberation. But another formula does not fill out the missing content of the first formula. It provides a new consideration that replicates the same gap between principle and action. The prescriptive strategy for including the Unity Claim within Kant’s moral theory seems to be no more acceptable than the accommodative strategy.

III. Rejectionism

The other prominent approach – and it is probably the more popular approach – is to give up on the Unity Claim altogether. This approach is not always arrived at directly through a consideration of the Unity Claim itself. Thomas Hill, for example, starts from a consideration of individual formulas, and decides that they are too different to be reconciled:

> Although fine questions about the relations among the formulations must await proper interpretation of each, any reconstruction of Kant’s second formulation certainly seems to be independent of others, and, in fact, to go beyond the famous first formula … in declaring a rather substantive value judgment with significant practical implications. (Hill 1992: 55)

Paul Guyer, by contrast, approaches the issue from systematic considerations to argue that Kant misunderstood the different roles that the formulas play in explaining the possibility of rational
agency. In both cases, the Unity Claim is cast aside in order to preserve the core of Kant’s moral theory. I hope to show that this approach, too, is unacceptable.

There seem to be two main reasons why Kant commits himself to the Unity Claim. The first relates to Kant’s vision of moral value. Morality concerns that which we are fully accountable for, and we cannot be accountable for anything outside of our control. This is what generates the unqualified character of moral value: it calls for a kind of assessment which pertains exclusively to what we are capable of in our special status as rational agents, and is thus immune to any contingent features of the world. The necessity of the moral law makes available a realm of unqualified value, “so a collision of duties and obligations is inconceivable”: such a “collision” would qualify the necessity of law or moral value or both. But the formulas of the categorical imperative without the Unity Claim at least make it conceivable that there could be conflicts in duties. Formulas that command absolutely and differently could produce different results, without there being any means to adjudicate among conflicts. So rejecting the Unity Claim places the possibility of tragedy at the heart of practical reason, and thus morality.

The second main reason why Kant declares allegiance to the Unity Claim stems from the three different argumentative roles that the three formulas play. Derivations of the content of the moral law all run through the so-called Formula of Universal Law: they have as their basis the conformity to the “universality of a law as such”(421). The Formula of Humanity, by contrast, provides Kant with the basis for the “vast majority of the ethical duties he sets out” in the Metaphysics of Morals. The Formula of Autonomy, finally, is the starting point for Kant’s justification of the moral law, since the reciprocity of morality and freedom is explained in terms of autonomy. So rejecting the Unity Claim carries another dire consequence: it disconnects content, practical application, and justification. Unless the three formulas are in some sense one
and the same, the content of the law might have little to do with its application, what is justified might not be the content of the law, and application and justification might be disconnected.

A response to these difficulties seems to have had its first formulation in the work of Barbara Herman, and has since appeared, in various guises, in the work of others. According to this response, all the formulas together combine to produce a fully-developed account of rational agency (or of “rational principles”). So, for example, Paul Guyer argues that “only the four formulations of the [categorical imperative] taken together display all the conditions necessary to show that action in accord with [the categorical imperative] is a real possibility for any rational agent, a fortiori a real possibility for ourselves if we are or are capable of acting as rational agents.” Each formula, standing alone, identifies some important element or aspect, but only the several formulas as a systematic whole present enough content to make the moral law meaningful and possible. So, although the various formulas are interrelated, they must be significantly different from one another to serve their collective purpose; the Unity Claim is false. But the problem of the formulas having different argumentative functions is merely apparent, since they are unified by their common source in the nature of rational agency, which in turn links justification to content to application. And this also relieves the problem of potentially conflicting moral duties, since potential conflicts could be adjudicated at the level of systematic ground: the formulas might provide different reasons, but appeal to the more fundamental conception of rational agency could determine which reason actually generates an obligation. What ensures consistency is the comprehensive conception of rationality that lies behind the formulas of the categorical imperative, and which is only fully expressed by the formulas as a group.
I have no interest in disputing that the formulas of the categorical imperative only taken together present a fully-developed picture of rational agency. It even seems plausible to me that Kant had such an argument in mind, although his assertion of the Unity Claim would then be puzzling. But such an argument seems to neglect one important feature of its subject-matter: that they are formulas of the categorical imperative. Formulas of an imperative are themselves imperatives, and an imperative does not serve primarily to represent a conception of rational agency, or anything else; an imperative tells one what to do. Whereas multiple claims might be necessary to fill out a well-rounded theoretical conception, multiple imperatives merely give one different commands. So if a particular formula does not adequately express what it is intended to, that should not provoke additional formulas, but an admission of failure.

The problem is this. The Unity Claim seems intended to resolve the difficulties that arise with a diversity of formulas. The proposal now in question rejects the Unity Claim, and so has to resolve those difficulties by locating a more fundamental unity behind the formulas. But this does not address the difficulties. One difficulty is simply that the formulas seem to provide different and potentially conflicting grounds for action, and claiming that the formulas have a common source does not alleviate that in any way. “Go to bed,” “Do your homework,” and “Finish your vegetables” might all come from the same source, and might even collectively represent a single conception of right parenting, but they are not the same imperative. Having the same source does not imply commanding the same, or even commanding consistently; and only commanding the same is adequate to resolve Kant’s difficulty. The other way in which the current proposal fails to address the difficulties is by not articulating a solution. The conception of rational agency is left to imperatives to articulate, but imperatives are not very articulate: one does not say what something *is* by giving commands. Worse, the categorical imperative itself
seems to be left inarticulable. Each formula is the inadequate expression of an imperative for which there is no adequate expression. Principles are typically discursive, but the categorical imperative is then a principle that cannot be put into words, at least not in imperatival form. Rejecting the Unity Claim thus incurs severe costs that cannot be compensated for by other means.

IV. Conclusion

Kant presents a diversity of formulas of the categorical imperative. Each formula does not, on its face, amount to the same thing as the others. Defenders of Kant’s theory are compelled either implausibly to claim that they do amount to the same thing, or to claim that they can be made into a unity, or to claim that they are all based in the same conception of rational agency. This last claim is the most plausible, but it does not resolve the fundamental difficulty: that Kant provides several imperatives that need to all be a single imperative.

“Der kategorische Imperativ ist also nur ein einziger” (421), Kant says in introducing the universal law formula – “There is, therefore, only a single categorical imperative.” But there are two features of Kantian theory that have attracted adherence into the present day. One, which is represented by the Formula of Universal Law, is that the purity of practical reason underwrites the objectivity of the moral law. The other, which is represented by the Formula of Humanity, is that the substantial ideal of humanity serves as a basis for respect. The combination of these features leads to a powerful moral theory, but each feature is represented by a separate imperative, and what my argument suggests, I think, is that there is no way to reconcile them. The Unity Claim serves to stipulate that purity and substance can be one and the same. But, in part because the Unity Claim concerns imperatives, there does not seem to be any way to make
sense of it. Even if the formulas have complementary representational functions, they do not constitute a unified imperative.

Kant seemed to need several versions of the supreme principle of morality because none by itself was adequate. Kant’s insistence on the purity of the moral law creates a divide between the normativity of practical reason, which is constituted by the agent’s own legislation, and practice, which is, apart from its subjection to moral law, ungrounded. The multiple formulas of the categorical imperative constitute, among other things, a heroic attempt to close that gap, by making both justification and a substantive good available in the moral law: precisely the purity of the law provides the richness and specificity of substantive ethical life, and thus engages with practice without any loss of rigor. But this solution depends on the Unity Claim, which stipulates that the various features are jointly available when in fact they are not. All this perhaps leaves room for a Kantian theory, but not quite Kant’s. Without the Unity Claim one seems to be left at best with a choice of either a formalism of principle or substantial ethical practice, but not both.38

---


2 As Thomas Hill notes, Paton renders this line differently. See Thomas E. Hill, Jr., “Editor’s Introduction: Some Main Themes of the *Groundwork*,” in T.E. Hill and A. Zweig eds.,
According to Paton, the Unity Claim reads, “The aforesaid three ways of representing the principle of morality are at bottom merely so many formulations of precisely the same law, *one of them* by itself *containing* a combination of the other two” (Immanuel Kant, *The Moral Law* (London: Hutchinson University Library, 1948), H.J. Paton trans., pp. 97f, emphasis added). Paton then suggests that it is the so-called Formula of the Kingdom of Ends that contains the other two formulas. But I shall not consider Paton’s translation because (a) it is incorrect and (b) it makes no difference to the main issues of this paper whether one formula contains the other two, or each formula unites the other two in it. I am interested in what it could mean for any single formula to unite the other two in it.


6 I realize that a “complete determination of all maxims” is not, properly speaking, a feature of maxims, but that is what Kant reports. In Kant’s words, “Alle Maximen haben naemlich … eine *vollstaendige Bestimmung aller Maximen*” (436).


19 Paul Guyer, in *Kant on Freedom, Law, and Happiness* (New York, Cambridge University Press, 2000), p. 139, points out that the word translated here as “precarious” also means “disagreeable or morally contemptible.”


It is not the purpose of the paper to argue this point, but even on O’Neill’s readings, the formulas seem substantially different. The Formula of Universal Law rejects maxims that “one cannot at the same time will …” while the Formula of Humanity rejects maxims that another “cannot possibly agree with.” (O’Neill, “Universal laws and ends-in-themselves,” p. 138) This suggests two important differences: that between “could agree with” and “could will,” and that between taking into account the formal character of one’s own agency, and taking into account who else is actually an agent.


37 Cf., e.g., Christine Korsgaard, “Kant’s Formula of Humanity,” in *Creating the Kingdom of Ends* (New York, Cambridge University Press, 1996).

38 I am grateful to Paul, Anna, Steven, Mary, Robert, and Thomas for their helpful discussions on the topic of this paper.