

## **ETHICS AT THE INTERSECTION OF KANT AND ARISTOTLE**

**AN INTERVIEW WITH CHRISTINE M. KORSGAARD  
BY ANA MARTA GONZÁLEZ**

In this interview Christine M. Korsgaard answers several questions on her particular interpretation of Kant's moral philosophy. More specifically, she dwells on the role of the noumenal/phenomenal distinction for a better understanding of Kant's idea of practical reason, as well as on the compatibility between Aristotelian and Kantian practical reason. In addition, she answers a couple of more general questions on the perspectives and challenges of contemporary ethical theory.

Durante el curso 2002-03, Ana Marta González fue una *Fulbright-Visiting Scholar* del Departamento de Filosofía de la Universidad de Harvard. Durante su estancia, desarrolló un proyecto de investigación sobre moral, naturaleza y cultura en la filosofía práctica de Kant, bajo la supervisión de la prof. Korsgaard, conocida especialista en filosofía moral kantiana. Al término de su estancia en Harvard, A.M. González propuso a aquella una entrevista en la que quedarán reflejados los aspectos de su filosofía que, de un modo u otro, habían ido saliendo en el curso de las conversaciones mantenidas a lo largo del año. Entre sus múltiples ocupaciones, la prof. Korsgaard respondió con su habitual claridad a un cuestionario que se le pasó por escrito. El resultado es esta entrevista en la que aparecen los puntos fundamentales de su filosofía hasta el momento. Se reproduce aquí en inglés para evitar las imprecisiones derivadas de una traducción.

**AMG.**— You are well known among Kant Scholars for being one of the leading interpreters of a new approach to Kant’s Ethics. This approach challenges the usual charges of formalism and “rigorism” that, at least since Hegel and Schiller, have been made against every rendering of Kant’s moral philosophy. Among the authors now developing a similar approach, there are other students of Rawls, such as Barbara Herman, Thomas Hill and Andrews Reath<sup>1</sup>. Taken together, you all seem to be developing a much more “practice-oriented” account of Kant’s moral philosophy, determined not to become entangled in Kant’s characteristic “dualisms”. John Rawls once said that you don’t need to keep those dualisms in order to remain a Kantian<sup>2</sup>. This is certainly an hermeneutical option: other philosophers, such as Jaspers, took Kant’s dualisms as an expression of human finitude, thereby offering an existential or metaphysical reading of Kant. I think it is important to hold both aspects together. Otherwise, the human being could easily be reduced to a mere rational agent. I do not know whether you agree with this. In any event, contemporary ethics has been trying to rescue other dimensions of the human being. Thus, for instance, against the modern stress on autonomy, an author like Levinas has emphasized passivity and vulnerability;

---

1. REATH, Herman, and KORSGAARD, eds., *Reclaiming the History of Ethics. Essays for John Rawls*, Cambridge University Press, 1997. Thomas Hill refers to Rawls’ lectures at Harvard in 1962 (HILL, *Dignity and Practical Reason*, Ithaca: Cornell University Press, 1992, p. 17); Barbara Herman refers to his lectures in 1977 (HERMAN, *The Practice of Moral Judgment*, Cambridge, Harvard University Press, 1993, p. 50). A version of Rawls’s lectures is now available in John RAWLS, *Lectures on the History of Moral Philosophy*, edited by Barbara Herman. (Cambridge, Harvard University Press, 2000).

2. “Kant’s view is marked by a number of dualisms, in particular, the dualisms between the necessary and the contingent, form and content, reason and desire, and *noumena* and *phenomena*. To abandon these dualisms as he meant them is, for many, to abandon what is distinctive in his theory. I believe otherwise. His moral conception has a characteristic structure that is more clearly discernible when these dualisms are not taken in the sense he gave them but reinterpreted and their moral force reformulated within the scope of an empirical theory. One of the aims of *A Theory of Justice* was to indicate how this might be done”. “A Kantian Conception of Equality” (1975), in *Collected Papers*, ed. S. Freeman, H. U. Press, Cambridge, Ma., London, England, 1999, p. 264.

against our stress on external action, Iris Murdoch liked to emphasize what she called the interior “fabric of being”. You have a very powerful interpretation of Kant’s ethics along Rawls’ lines. My question is: does your interpretation of Kant allow for seeing the human being in a way other than as a moral agent?

**CMK.**— There are two assumptions behind this question, I think. One is that the various dualisms you mention add up to one big dualism, with the necessary, the formal, the rational, and the noumenal on one side; and the contingent, substantive (i.e., content-ful), desiderative, and phenomenal on the other. The other is that if we accept this basic dualism, Kant leaves us no option but to identify the essentially human with the necessary, formal, rational, noumenal side of things, and so to see human beings as bare moral or rational agents. So the first thing I want to say is that I don’t accept the first of those two assumptions, and therefore I need not accept the second either.g

I think that the distinction between the noumenal and the phenomenal is not an ontological distinction, but rather marks the difference between activity and passivity. We regard ourselves as noumenal insofar as we think of ourselves as active beings, and place ourselves among those forces that generate the world of appearances; we regard ourselves as phenomenal insofar as we know ourselves through inner and outer appearance. I see no reason to reject that distinction. Quite the contrary: the basic Kantian insight that the mind acts on something given to it depends on it. To put Kant’s point in Aristotelian terms, the mind imposes form, especially the form of law, on a matter or content that is given to it. This is true in both theoretical and practical philosophy. Kant thinks we contribute to the production of the world of experience in part by imposing the form of causal law on the material given to sensation, and that we create norms and values by imposing the form of law on maxims whose content is given to us by experience. But for that very reason, the distinctions between form and content, and reason and desire, should not be regarded as *dualisms*, since in these cases you have two items that are not

rivals of some sort, but rather play *functionally* different roles. We need them both. Desire proposes ends: reason decides whether they are worth pursuing or not; it does that by seeing whether the substantive maxims of acting on those desires pass a test given by the formal maxim, the categorical imperative itself. I have tried to bring this out by arguing that Kant holds what I call “the constitutional model” of the soul<sup>3</sup>. According to this model, the function of reason is to legislate for the good of the whole; the function of desire is to make legislative proposals, prompted by the person’s empirical needs and interests. It is confused to think of a “dualism” between two distinct parts that are functionally related. So this is a different way of overcoming the supposed dualisms than that of Levinas or Murdoch, who basically accept the claim that these are dualisms but insist that in important ways human beings fall on both sides of them.

To answer your final question more directly, then, my view not only allows but absolutely requires that we be something more than moral agents, since morality needs material to work on, and that comes from our empirical nature. But in one way, I go beyond Kant in insisting on the importance of our empirical identities. Kant thinks that desires are suitable to provide the content of our maxims because their satisfaction makes us happy. I think that the desires and interests that give content to our maxims spring from what I call our empirical or contingent “practical identities,” the roles and relationships in terms of which we define ourselves and find our lives to be worth living. (The notion of practical identity is introduced in *The Sources of Normativity*, §3.3.1; in *Self-Constitution*, currently in Lecture Four, I try to explain why human beings must have such identities.) These identities give rise to duties and obligations; they are sources of necessity in our lives, just as the moral law is. So it is essential to my story that each of us

---

3. In “Self-Constitution in the Ethics of Plato and Kant” (*The Journal of Ethics*, 3, 1999: pp. 1-29) and in Lecture Five of the current version of *Self-Constitution: Action, Identity, and Integrity*. The current version consists in a set of lectures given in 2002 as the Locke Lectures at Oxford; a book version is forthcoming from Oxford University Press.

has, and must have, individual and culturally situated ways of identifying herself that are important to her.

**AMG.**— Very much in line with Kant’s –and Rawls’- own doctrine of the primacy of practical reason, you have defined Moral philosophy as “the extension and refinement of ordinary practical deliberation, the search for practical reasons”<sup>4</sup>, and have taken Kant as a prominent example of this way of doing moral philosophy: “Kant’s arguments are not about us; they are addressed to us”<sup>5</sup>. It is from this perspective, I think, that you have reviewed the traditional approaches to the Categorical Imperative, trying to develop other aspects perhaps neglected in the past: for instance, the particular role of the Hypothetical Imperative in practical deliberation or its complementarity –rather than opposition- to the Categorical Imperative<sup>6</sup>. You have also shown how the different formulas of the Categorical Imperative –especially the formula of Humanity- account for significant moral differences in the way we deliberate about particular actions. Yet, on the whole, your focus on practical reason seems intended to highlight the central role of the concept of “rational nature” in Kant’s moral philosophy. This has been criticized<sup>7</sup> as not being very Kantian. What do you have to say about this?

**CMK.**— In the passages from Jerry Cohen’s commentary to which you refer, he criticizes me for saying that morality is derived

---

4. KORSGAARD, C., *Creating the Kingdom of Ends (CKE)*, Cambridge University Press, 1996, p. xii.

5. *Ibidem*.

6. Especially in *Self-Constitution*. See also HILL, T., “The Hypothetical Imperative”, in *DPR*, pp.1-17

7. See COHEN, G. A., “Reason, Humanity, and the Moral Law”, in *The Sources of Normativity*, Cambridge University Press, 1996, p. 172. (*The Sources of Normativity* is available in Spanish translation as *Las Fuentes de la Normatividad* translated by Laura Lecuona and Laura E. Manríquez with assistance from Faviola Rivera, Mexico City, Universidad Nacional Autónoma de México, 2000.) See also REGAN, Donald H., “The Value of Rational Nature”, in *Ethics*, 112 (2002), pp. 267-291.

from human or rational nature. It is the “nature” part of that phrase to which Cohen objects. He appeals to the well-known passages in the *Groundwork* in which Kant says that the argument for the moral law must not be derived from anything specific to human nature or even human reason. I don’t think that I make the mistake Kant is discussing there - appealing to contingent psychological facts about human beings. When I said that morality springs from our nature, my intention was to emphasize Kant’s view that the laws of reason are *our* laws, the laws of the human mind. Perhaps I can best put the point historically. The dogmatic rationalists (e.g. Wolff, Clarke) followed the Stoics in thinking of the laws of reason as laws of the structure of the world - reason in us responds to, or tracks, a rational structure in the world, perhaps one imposed upon it by God. The sentimentalists (Hutcheson, Hume) thought that morality was a projection of human sentiments onto the world. Kant falls between these two groups, since he thinks that the laws of reason are themselves projections of human standards onto the world.

Regan’s criticism is a little different. He does not claim that it is unKantian to appeal to the value of rational nature, but rather that we cannot derive a system of values from the value of rational nature. His point is not just against me but against Kant. I think that human beings confer value on things, and express our sense of the value we place on ourselves when we do so. Regan thinks this makes our values arbitrary and so not worth having. He thinks that we should value things that are intrinsically valuable. It’s a little hard for me to respond to this, short of repeating the many criticisms I have made both of realist positions generally and of the idea of intrinsic values in particular throughout my work. So I will just say that at bottom, I think that the temptation to hold such views is based on an unwillingness to face up to the contingencies of the human situation, and the desire for a (delusory) sense of secure foundations. It’s not as if we really have some form of epistemic access to facts about intrinsic values, facts that supposedly exist independently of our natural concerns. Rather, we are tempted to call things “intrinsically valuable” only because we

care about them. So why shouldn't we admit that our own concerns, and our willingness to reflectively endorse them, is the real foundation of our values? And in that sense, rational nature is the foundation of all value.

**AMG.**— Among the contemporary Kant scholars, you, along with Nancy Sherman, are perhaps the ones most interested in Aristotle. This is certainly not because of Rawls' influence. Originally, the topic of your doctoral dissertation was going to be a comparison between Aristotle and Kant's accounts of practical reason as the foundation of ethics. You have also published two papers comparing certain aspects of Kant and Aristotle<sup>8</sup>. In your way of disclosing the dynamism of Kantian practical reason, going beyond the usual debates on the Categorical Imperative, it is difficult not to think of the influence of Aristotle's deliberative ethics. How Aristotelian is your Kant?

**CMK.**— I have never understood why people regard the views of Aristotle and Kant as being opposed to each other. While they do not agree on every point of detail, it has always seemed to me that their views are broadly compatible. I have tried to explain why in one of the papers you mention, "From Duty and for the Sake of the Noble: Aristotle and Kant on Morally Good Action."

Quite generally, both Aristotle and Kant think it important to describe the workings of the mind in terms of the interplay of active and passive functions. Of course they conceive that interplay somewhat differently: for Aristotle the active intellect extracts the form (or "universal") from the matter, while for Kant, as I said above, the role of mental activity is to impose universal form on

---

8. KORSGAARD, C., "From Duty and for the Sake of the Noble. Kant and Aristotle on Morally Good Action", in S. Engstrom and J. Whiting (eds.), *Aristotle, Kant, and the Stoics. Rethinking Happiness and Duty*, Cambridge University Press, 1996, 203-236. And "Aristotle and Kant on the source of value", in *Creating the Kingdom of Ends*, Cambridge, University Press, 1996, pp. 225-248.

matter. As far as ethics is concerned, they both believe that ethics is grounded in practical reason, and this is what originally attracted me to them. They also share an important view that sets them apart from almost everyone else: namely, the view that the object of choice, and the bearer of moral value, is something expressed in a maxim or *logos* - an act done under certain circumstances for the sake of an end. Moral value does not belong merely to an act that is chosen as a way to promote some desired end, nor to an end chosen independently of and prior to the act. Instead, the act and the end are chosen as part of a package, *an act done for the sake of a certain end*, the whole of which is seen by the agent as a thing worth doing. In *Self-Constitution*, I call such a thing an *action*: an act done for the sake of an end. It is actions, in this sense, that that is judged to be lawful (either permissible or required) or noble (*kalon*). Both philosophers, I believe, think that the moral property - the property of being lawful or noble - arises from the way the parts of such a formula, the *logos* or maxim, are related to one another. Doing *this* sort of act for *this* sort of end - that, the whole thing, that's a thing worth doing, worthy of choice, noble, or not as the case may be.

Of course, Kant thinks that the categorical imperative test tells us when a maxim is lawful, while Aristotle thinks that the decision whether a *logos* is *orthos* is more a matter of judgment or perception. People make a great deal of that difference, but I think it can only be a matter of degree. Kant of course admits that judgment is always involved in the application of principles, and Aristotle thinks that we can articulate some rough general rules, so the difference between them so far as this goes is a matter of *how much*. That is, how much of a moral decision is a matter of pure judgment, how much of the reasoning can be articulated?

Most of the other differences between the two philosophers actually come from a difference in subject matter. Aristotle, because he believes that moral judgment is more a matter of perception, focuses on the question of virtue - that is, the question of what state our passive or receptive faculties have to be in if our active faculties are to work properly. If we are to "see" things

correctly, our desires have to be in a certain condition. Kant, because he believes that the grounds of moral judgment can be articulated rationally, focuses more directly on the question of moral rightness – what the criterion for it is, and how it is grounded. I believe that Kant’s theory of rightness and Aristotle’s theory of virtue are compatible, and can be grafted onto each other. In fact I think that we get a better result when they are combined in this way. Kant, for reasons I try to explain in the paper I mentioned, has a rather defective account of the emotions, and so of their role in moral judgment. This is the main source of their famous difference over the question whether virtuous action must be enjoyed. The Aristotelian account of the emotions, which treats them as perceptions rather than as mere feelings, works well with Kant’s theory of rightness. And Kant’s theory of rightness is better articulated than Aristotle’s. Where they do differ – as they do with regard to questions of human equality, for instance – it is also better.

**AMG.**— Along with the emphasis on practical deliberation, I think that we can recognize another trace of Aristotle in your interpretation of Kant, namely, your readiness to make use of teleological arguments. I first noticed this in *Creating Kingdom of Ends*<sup>9</sup>, but it is particularly clear in *Self-Constitution*. How do you ground this recourse to teleology within Kant’s moral philosophy?

**CMK.**— I don’t ground teleology within Kant’s moral philosophy, but rather within his view of conceptualization more generally, perhaps especially as it is elucidated in the third *Critique*. I believe (and think Kant believes) that the conceptual scheme we work with in everyday life is broadly teleological. This is not because things in fact have purposes (whatever that might mean) or were made for purposes, but simply because of the way we think of the world, the way we divide it into objects. Plants and

---

9. KORSGAARD, C., “Kant’s Formula of Universal Law”, in *Creating the Kingdom of Ends*, pp. 77-105.

animals may be thought of teleologically because they have self-maintaining forms; artifacts because they are our own productions. More generally we divide the world into objects by observing what goes together *functionally*. Even the everyday use of the word “cause” has a teleological element. What are we doing when we say that a certain particular object or event was “*the cause*” of another object or event? In a strictly mechanistic conception of the world, what we should really say is that one complete state of the world causes another. So when we say that one thing is “*the cause*” of another, we are holding many other things equal, and then the question arises: what determines which things we hold equal when we talk that way? Now I think that the answer to that question must be given functionally. We single something out as “*the cause*” because it is something we can use to produce or prevent an effect. It represents a place where we can enter into the causal network. This shows that the main reason we need the notion of “*the cause*” in fact springs in a deep way from the problem of action itself - we need the notion of “*the cause*” in order to formulate hypothetical imperatives.

**AMG.**— Interpretations aside, it is clear that in *Sources of Normativity* and in *Self-Constitution*, you have begun to develop your own ethical theory. In those works you draw on authors different from Kant: Hobbes, Pufendorf, and Hume... in *Sources*; Plato and Aristotle, in *Self-Constitution*. Still, in both works you remain within the framework of Kant’s ethics, trying to develop an alternative account to Hume and utilitarian inspired ethics. Is the ethical debate still defined by those two poles? Don’t you think that, in order to confront contemporary ethical issues, both ethical traditions are somehow exhausted? Isn’t it there a new moral sensibility, demanding another way of thinking?

**CMK.**— I wouldn’t say *two* poles. Hume is no utilitarian – sentimentalism and its modern descendant, expressivism, represent a distinct option. But actually, the view I have most consistently argued against is realism, not Humeanism or Utilita-

rianism. And that's because I think that realism, that is, dogmatic realism or what in *Sources* I call "substantive" realism, is essentially a defeat for philosophy – a way of giving up on the hope of explaining ourselves to ourselves. To say, *it's just a fact* that certain things are valuable, or rational, or whatever it might be, is to give up on the project of trying to explain why these things are so.

More generally, however, my strategy is not to set up two-way debates, and try to raise Kant from the ashes of his defeated opponents. Rather, what I look for, and try to articulate, is what different ethical theories have in common, how they can be fit together. That was part of the explicit aim of *Sources* (see especially §§4.5.1-4.5.5), and is a standing assumption of the way I deal with Plato, Aristotle, and Kant in *Self-Constitution*. Debate is a useful heuristic, but I am convinced that most philosophers over-emphasize its value. Surely if we can find real insight - or indeed truth - anywhere, it is at the intersection of what the really profound philosophers have thought. But this isn't just a point about methodology. It also just seems true to me that the philosophers I study have a great deal in common, and that the moral philosophers of the past say things that are utterly recognizable to a "modern" moral sensibility.

So to answer your question more directly - no, I don't think the Kantian tradition is exhausted, nor do I think it makes any sense to think of adopting new ethical theories to fit "new moral sensibilities." I think that ethics is grounded in the nature of action and practical reason, and that those in turn are grounded in human self-consciousness. These aren't things that I see as changing over time. This doesn't mean I don't think that there has been moral progress - as I wrote in *Sources*, the broad changes in outlook that characterize Enlightenment thinking seem to me to be a form of moral progress. (Some people think we are past (or *Post-*) those changes, but I think we are only just barely getting started on realizing them in practice.) But if we are ever to see a change in moral sensibility as representing *progress*, rather than as mere *change*, we must be able to understand these evolving sensibilities

within the framework of a single account of what our moral lives are all about.

**AMG.**— You have said that, systematically, the lectures on *Self-Constitution* should be read before *Sources*, although they were written later. They are indeed an extraordinary clear introduction to the structure and dynamism of practical reason. In them you argue for several controversial theses. One of them has to do with the concept of “identity”. At a certain point, you say: “in the relevant sense there is no ‘you’ prior to your choices and actions, because your identity is in a quite literal way constituted by your choices and actions”. And further: “Personhood is quite literally a form of life, and being a person, like being a living thing, is being engaged in an activity of self-constitution”. Now, assertions like these have an existentialist flavor. I think you are just trying to explain the constitution of what you call “practical identity.” I find this notion quite Aristotelian, but you use it within a Kantian framework. Accordingly you cannot speak of a substance ontologically prior to the development of any practical identity. My question is: given that Kant can hardly speak of an ontological “I”, how do you avoid an existentialist conclusion?<sup>10</sup>.

**CMK.**— Why should I want to avoid an existentialist conclusion? Say that an existentialist believes that human beings are the creators of all meaning and value, and, while we are at it, in a way also the creators of ourselves. This is something I firmly believe is true, partly because I don’t think any other hypothesis makes any sense, and partly because I don’t think any other hypothesis does any philosophical work. What I mean is that even if certain things just have intrinsic value, or if God creates certain values, this can be nothing to us unless these values coincide in some way with what *we* value. Perhaps my allegiance to existentialism shows up most clearly in §§4.4.1-4.4.2 of *Sources*,

---

10. This is actually one of Thomas Nagel’s remarks in *The Sources of Normativity*, p. 203.

where I claim that if a person ceases to value himself then for him it is true that nothing has value. That is an existentialist bullet that I am prepared to bite. What frightens people about the existentialist hypothesis is the further conclusion that if we are the creators of meaning and value we can create them anyway we like, that *anything goes*. But I don't draw that conclusion: I don't think that there is anything, including valuing and acting, that can be done anyway that we like, since every activity has constitutive rules of its own. If all my arguments go through – a large assumption of course – we end up with Kantianism. So if my view is existentialist, it is nevertheless supposed to be a kind of existentialism with rules.

Nor do I think that the ontology of the self-constituting self is something spooky. The ontological substrate of the person, that is, of the self-constituting self, is a self-conscious human animal, which for the reasons I explain in what is now Lecture Four of *Self-Constitution* is necessarily faced with the task of constructing an individual self. And the view that the ontological substrate of a person is a human animal is, as you note, simply Aristotle's view. There's nothing spooky about that.

**AMG.**— Another controversial thesis in *Self-Constitution* is the broad conception of action that you develop there. According to it, you may say that animals are able to act, only in a different way than humans. This is because at that point you are defining action through autonomy and efficacy, and not so much through "choice". Now, attributing action to animals does not sound very Kantian. You try to make it sound Aristotelian, by appealing to the Aristotelian category of the 'merely voluntary'. On the other hand, you still define the properly human action through choice. What is not clear to me in your account is whether choice is just a further grade of autonomy or efficacy along a natural line of development, or, rather, something totally new.

**CMK.**— Actually, as I notice in *Self-Constitution*, Kant wavers on the attribution of action to non-human animals. In some places he talks about an animal will; in one place he identifies the capacity for action a little too broadly with “life” (*Metaphysics of Morals*; 6: 211). The view of the third section of the *Groundwork*, on the other hand, seems to leave no room for animal action. There he *identifies* the will with practical reason. This wavering seems to me to parallel his wavering on our treatment of animals: in the *Groundwork* he categorizes animals as mere instruments or things, while in the *Lectures on Ethics* he has harsh words for those who mistreat them needlessly.

But I will leave the ethical issue aside for now. In general, I think it is a great mistake in philosophy to ignore non-human animals, either when thinking about the nature of our agency or about the nature of our cognition. Any philosopher who produces a theory that makes it seem as if non-rational agency or non-rational cognition are *impossible* has to be wrong. This is because he has produced a theory according to which animals cannot possibly exist, and they obviously do. Our forms of agency and cognition are a development out of something, something more widely shared with other conscious beings whose lives are governed by perception and action. In order to understand our own forms of cognition and agency, we need to know what that something is.

In my own view rational principle plays the role in our mental lives that instinct plays in the lives of the other animals. I characterize action from instinct as a form of autonomy because when an animal is guided by its instincts its movements are in a certain way its own. It acts from a conception or representation of the world that is shaped by its own interests in such a way as to maintain its own identity. This is what we also do when we act morally and rationally. So in that sense, human autonomy or choice is a further step along a natural path.

But we are more fully autonomous because we think about, endorse, in a way choose our own rational principles, making them more fully – and also more individually – our own. Our capacity to choose the principles of our own action results from self-

consciousness, something I see as making a radical difference between us and the other animals. After all, there has to be a radical difference - doesn't there? Because we really are very different from other living things. And I think the difference springs from self-consciousness and the strange new forms of self-control and self-creativity (and also the terrible perversions of self-control and self-creativity) that it brings. So our form of agency is also something new. So: the answer to your question whether choice is a further grade or autonomy or something new is, and has to be, *both*.

**AMG.**— Another strong thesis in *Self-Constitution* refers to the Categorical Imperative. There you maintain that not only the Hypothetical Imperative, but also the Categorical Imperative is constitutive of practical reason. From then on, you are in a position to explain the nature of bad action in terms of not fulfilling what it is expected from a rational (and free) agent. Thus, those actions are bad, which cannot be universalised, thereby constituting a unified agency. As you know I have a problem with Kant's recourse to universality. While I see that every good action is universal in some way, I don't really see that the universality proper to the good action coincides with its *subjection* to an abstract universal law. Accordingly, while I see the connection between the goodness of an action and its ability to generate internal unity in the agent, I am not sure that the universalization procedure is the right way to certify that goodness/unity. Among other things, I am afraid that this recourse does not discriminate between the internal unity required by the individual moral agent, and the unity of society as a whole. But those are two different types of unity. The requirement of universality was also one of the controversial issues in *Sources of Normativity*, for some other reasons. Raymond Geuss, for instance, argued that the notion of identity and the requirement of universality cannot be held together<sup>11</sup>.

---

11. GEUSS, R., "Morality and Identity", in *The Sources of Normativity*, p. 192.

**CMK.**— There's a lot here, so let me try to take your points one by one.

First of all, I'm not sure what you mean by an action's "*subjection* to an abstract universal law." As I understand Kant's view, an action embodies a maxim or principle: the maxim or principle describes the action. As I say in *Self-Constitution*, the relation of an action to its maxim is like the relation of a sentence to its meaning. That maxim or principle must be universal in the sense it must express a universal normative commitment on the part of its agent, a commitment that is autonomous. Neither the action nor the agent is "subject" to anything outside of itself.

Second, I think it is important to distinguish the universalizability requirement from the universalization procedure. It is one thing to say that you must be able to will your maxim as a universal law, and quite another to say that the right/best/only way to find out whether you can will your maxim as a universal law is through Kant's universalization procedure. That procedure involves performing a certain thought experiment, seeing whether you could will to act on your maxim in a world in which your maxim was universal law without generating some sort of contradiction in your will. While I don't agree with those of Kant's critics who complain that this procedure is empty, I do agree that it has limitations. (I have discussed both sides of what I just said in "Kant's Formula of Universal Law" in *Creating the Kingdom of Ends*.) It is often much more fruitful to try to determine what we can will as law by trying to work out what we can square with the value of humanity.

I argue for the identity of what unifies a particular agent and what unifies groups of people trying to act together ("society" in your example) in what is now Lecture Six of *Self-Constitution*. I think you are wrong to say these are two different types of unity. Both forms of unity, in my view, are essentially deliberative, or rather achieved through shared deliberation. If I am to be practically unified with myself over time, then what is a good reason for me now must be a good reason for me later - it must *be able* to be a good reason for me later. And if I am to be practically

unified with you, then what is a good reason for me must be a good reason for you - it must *be able* to be a good reason for you. If I am to deliberate in harmony with myself, as opposed to being at war with myself, I must take as a reason now only a consideration that will still be acceptable as a reason to me later. If I am to deliberate in harmony with you, as opposed to being at war with you, I must take as a reason for myself only a consideration that will be acceptable to you as a reason. That is why the same universalizability requirement is involved in both cases.

Guess, in the passages to which you refer, invoked Shlegel's vision of an agent who makes his own laws and then breaks them ironically or frivolously in order to prove that his identity cannot be captured by any universal law. As I said in the *Reply* in *Sources* (p. 232), I agree completely with Hegel's diagnosis of this character as someone who is engaged in a constant process of self-creation and self-destruction. In a sense the whole of *Self-Constitution* is an argument against this picture. But I address it most directly in an argument I make at the end of what is now Lecture Two, an argument that I call "the argument against particularistic willing." That argument is somewhat complex, but I can convey its main point by means of an example.

Suppose I decide to go to the dentist on a certain day in order to get a cavity filled. I think I have a reason to do this, or that this maxim – the maxim of going to the dentist in order to get a cavity filled – embodies a reason. When I make this my maxim, my commitment is universal in the following sense: I commit myself to acting as this maxim specifies in all relevantly similar circumstances, that is, unless there is a good reason why not. The universality holds over all relevantly similar circumstances; if there is good reason not to act on it, the circumstances are relevantly dissimilar. Now it may turn out through some extraordinary circumstances that in order to get to the dentist on time on the day in question I have to risk my life. Since there is good reason not to risk my life, I can give up the project of going to the dentist that day without abandoning my maxim, since my maxim says to act a certain way unless there is good reason why not. On the other

hand, it may be that I am really frightened of the dentist and therefore I am *always* tempted not to go when the day arrives. Let's say that that is not a good reason for not going to the dentist, so if I don't go because I am frightened I will have abandoned my maxim. Now if I am prepared to give up the project of going to the dentist in the face of any consideration that tempts me whatever – that is, if *anything whatever* counts as a good reason not to go – then clearly I have committed myself to *nothing*. In that case I am wholly wanton, and I lack any coherent identity as an agent, since I am basically just a random collection of impulses. So if I am to count as willing a maxim my volition has to rule something out – and to that extent, it has to universalize over all relevantly similar circumstances.

The claim that universalizing over *all relevantly similar circumstances* includes universalizing over *all rational agents* of course requires further argument. The argument I make in *Sources*, in Sections, 3.4.7-3.4.10, is addressed to that point. There I argue that anyone who accepts the force of any reasons whatever is committed to valuing rational nature itself as a form of practical identity, and so to making laws that universalize over all rational agents. The arguments I mentioned before from what is now Lecture Six of *Self-Constitution*, to the effect that the unity of the self is achieved in the same way as the unity of a group, are also relevant to this question.

**AMG.**— Until now you have been studying the notion of “practical reason”. I know you are now focusing more on the nature of action. In the last decades, analytical and continental philosophies have found a common ground in the study of action, which is crucial to the study of humanity. It seems as if practical philosophy were all the philosophy we can expect in our cultural situation. So much so, that when someone dares to speak again of metaphysics –such as Levinas- it is only by taking ethics as the point of departure. How do you think of the relationship between ethics and metaphysics?

**CMK.**— A lot depends here on what you mean by “metaphysics.” If metaphysics is taken to encompass questions of ontology and of the nature of space and time, then not many analytic philosophers share the attitude towards metaphysics that you describe; those questions don’t have to start from ethics. There is plenty of independent work being done on the nature of space, time, and matter, and related questions about material constitution and the general question of identity over time - work that is only indirectly connected to ethics. Of course, as a systematic philosopher, I think everything is connected in the end.

But there are certain questions that are sometimes placed in the realm of metaphysics that I think have to be treated in connection with ethics - specifically, questions about the nature of action and *personal* identity. Perhaps it would be helpful to draw a parallel here. Kant thought of speculative metaphysics as being specifically concerned with three big issues: God, Freedom, and Immortality. And he certainly thought that speculative metaphysics in this sense had to be approached through ethics: theoretical philosophy can tell us nothing on these subjects by itself. I agree with Kant that what there is to be said for and against the existence of God, or rather, for and against faith in God, falls within the realm of ethics, but I want to lay that aside and talk about the other two issues which are more germane to my own work. You could say, I think, that in contemporary philosophy, the question of action is the analogue of Kant’s question about freedom, and the question of personal identity is the analogue of Kant’s question about immortality. When talking about both action and freedom, we ask: what makes a certain set of movements my own, movements that are attributable to me, and for which I may be held responsible? When talking about both personal identity and immortality, we ask: what makes a person me, both now and over time, what is essential to me and under what conditions do I continue to exist?

I think there is an advantage to talking about these questions under their more modern rubrics. It is too easy simply to dismiss questions of free will and immortality as metaphysical and religious dreams, ideas that cannot be squared with the modern

scientific worldview. But people are not so ready to dismiss the ideas of action and personal identity, which seem like everyday realities to most of us. Some analytic philosophers treat these questions simply as questions of theoretical metaphysics, while others see them as essentially ethical or involving ethics. My own view, as I said, is that these questions have to be treated within the realm of ethics.

Perhaps I might put it this way. I follow Kant in thinking that we have different standpoints from which we view, and conceptualize, the world. We have one set of concepts that we must use when we view the world as a possible object of theoretical understanding, when our aim is to explain and predict. The mechanistic conception of cause serves this purpose, for instance. And we have another set of concepts we must use when we view the world as the scene of action and interaction, when our aim is to do things both by ourselves and with others. Now I think that the concepts of action and personal identity belong to this way of looking at the world. It is not because of the requirements of scientific explanation that we need to look at things in the ways given by these concepts, but rather in order to act and live, with ourselves and with others. So these are questions that belong to *the metaphysics of morals*, the way we conceptualize the world for purposes of ethical life.