Part 5: Practical reason and rational ethics: Kant After considering, in the last article of this series (Ulrich, 2009), the origin of practical philosophy in Ancient Greek virtue ethics, it is now time to turn to its modern counterpart, rational ethics. Rational ethics sees the basis of morality in reasoning rather than in convention; its concept of ethics is in this sense "postconventional" rather than "conventional," a central distinction in Kohlberg's (1981, 1984) seminal work on the development of moral consciousness. Like traditional virtue ethics, rational ethics assumes that humans have a natural sense of right and wrong and that this moral sense can and needs to be developed through education and learning. Unlike traditional virtue ethics, however, rational ethics does not assume that the standards of morally good action are given by traditioned customs and conventions (or rules) such as the biblical "ten commandments" or the Aristotelian virtues but rather, that they ought to be a matter of reasoned choice on the part of responsible agents.

Of course, we have seen that reasoned choice also plays an important role in Aristotelian virtue ethics; but the subject of deliberation changes. Aristotle limits the scope of ethical deliberation to the task of ensuring a "balanced" judgment of situations in the light of values that are given by custom, in the form of traditional virtues; whereas in rational ethics, the values themselves become the central issue, in the form of a search for the moral principles that should guide us in a certain situation. This is an important step in the history of practical philosophy; for as long as we rely on custom and tradition as its basis, morally good action remains subject to cultural relativism. If ethics is to provide us with a basis for intercultural understanding and cooperation, principles derived from reason must replace or at least complement cultural tradition as the basis of moral judgment. Rational ethics, then, can also be defined as an ethics of principles.

As we noted in the last Bimonthly (Ulrich, 2009, p. 2), "rationality" in this context is to be understood in a wide sense of reasoning that includes
substantive as much as procedural rationality; normative as much as instrumental considerations; practical as much as theoretical reason. Furthermore, we said that when we study a subject philosophically, we usually examine it with a special interest in the part that reason – careful and self-reflective judgment – plays in it. From the standpoint of rational ethics, the reverse holds equally true: when we study the part that reason plays in morally good action, we have to study it philosophically rather than say, theologically or psychologically. Rational ethics is necessarily philosophical ethics, that is, it cannot in the first place be grounded in tradition, theology, or empirical science (e.g., ethnology, sociology, economics, or psychology), although it may of course be informed by other disciplines. In Kantian terms, it aims at reason's self-tribunal in matters practical, that is, it examines the part that reason rightfully plays in ethically defendable practice; or in more contemporary terms, it aims at a critically-reflective effort of securing "good grounds" for the validity of practical-normative claims.

Rational ethics is therefore the core discipline of practical philosophy; it seeks to clarify the criteria and principles of philosophical reflection about practice. It thus responds to an understanding of philosophy that we have characterized as an effort of second-order reflection on reflective practice, an effort that "happens when a practice becomes self-conscious." (Blackburn, 1994, p. 286; Ulrich, 2009, p. 5). In analogy to our previous short formula for practical philosophy:

\[ \text{Rational ethics} = \{\text{reflective practice}\}^2 = \text{RP}^2 \]

While rational ethics accordingly places high demands on our reflective skills, it is interesting to note that in comparison to Aristotelian virtue ethics, it is rather less demanding with regard to its motivational presuppositions. Some readers who followed me into Aristotle's world may have felt that his attempt to ground practical reason in personal virtue, whatever its merits may be, is all too idealistic for our epoch. Although I believe we need ideals no less today than in Ancient Greece, I rather share this concern; but I would formulate it a bit differently. I locate a weak spot in Aristotle's strong moral voluntarism: in Aristotelian practical reason, the methodological virtue of "virtue" comes to bear only inasmuch as each of us is willing (and able) to cultivate our personal virtue. In comparison, putting rationality first appears to place less ideal demands on our motivation. Rational ethics need not assume that we are virtuous in the first place, but only that we want others to
listen to our ideas and arguments, as they recognize that we speak reasonably. It is not because we are (or want to be) particularly virtuous but because we want to be rational that we will act morally! Virtue may then perhaps be reserved for a few (Mother Teresa comes to mind), but rationality is for all. This shift of perspective might thus indeed furnish a powerful strategy for grounding ethics; for who would not want to be recognized as being reasonable?

Our guides now will be Kant and Habermas, two philosophers of reason par excellence. I will start with a few considerations about the relationship between Kant and Aristotle, before then entering in medias res and trying to summarize my understanding of Kantian rational ethics. My main focus will be on the central role that the principle of moral universalization plays in it. I will try to explain its importance and underlying moral intuition, as well as offer some "translations" into everyday language. That will prepare us for the subsequent step, in a next contribution to this series, in which we will consider how Habermas attempts to reconstruct Kantian rational ethics within a discursive framework – an attempt that in turn will prompt some further considerations, in the final part of the series, about what remains to be done, after Aristotle, Kant, and Habermas, to realize the unfulfilled promise of recovering practical philosophy for practice.

Meanwhile, please bear in mind that despite this practical aim, our topic is still philosophy of practice, rather than philosophy in practice (as distinguished in the last Bimonthly). Our aim, to be sure, remains to mobilize practical philosophy as the envisaged third pillar of reflective practice; but the way towards the aim leads us through some difficult theoretical territory. I believe it is worthwhile to try and find our way through this territory. The ideas of Kant and Habermas, despite not being of an immediately practical nature, are so fundamental and insightful that I suspect no approach to practical philosophy today can do without relating its ideas to theirs. Adopting a comment about Kant by Tugendhat (1993. p. 98), with works of this level of insight one learns even from their inevitable limitations. We will certainly need to simplify and "pragmatize" their ideas if they are to inform reflective professional practice; but well-understood simplification must come at the end, not at the beginning of an effort to gain new philosophical ground.
Kant's *good will and the principle of universalization* Along with Aristotle, Kant is without any doubt the classical giant of practical philosophy. Whatever notion of practical reason we may have today, we can hardly define and explain it without relating it to these two thinkers. As different as their philosophies are, they share a fundamental interest in the problem of practical reason, that is, the question of how reason can and should guide good practice in human affairs.

I emphasize this shared concern at the outset because there is a tendency in the ethical literature to oppose Kant and Aristotle in a rather schematic way, so as to highlight the paradigmatic differences of ancient and modern ethics. For instance, comments on Kant often emphasize that while Aristotle's view of ethics is functional and teleological (How can we be happy?), Kant's is categorical and deontological (What ought we to do?). Similarly, their notions of virtue are opposed by asserting that unlike Aristotle, who sees it as a positive force (a way to live life to the full), Kant sees it as a mainly negative force (a mere duty that stops us from living life in all its fullness). Or Aristotle is said to pursue a monistic conception of practical reason (all claims to reason are justified by reference to the single end of *eudaimonia*), whereas Kant's conception is supposedly dualistic (reason's claims are either to prudence or to morality but never to both at the same time). Aristotle's approach is portrayed as traditionalist, Kant's as universalist; and so on.

I do not think this kind of schematic opposition of Aristotle and Kant is particularly helpful to understand what their practical philosophies are all about. It hardly does justice to the richness and subtleties of their thought; it only makes us blind for the interesting relationships between ancient and modern ethics (for good critical discussion, see, e.g., Engstrom and Whiting, 1996). I find it more interesting to ask in what ways Aristotle and Kant complement one another. In particular, what is the essential connection that they both establish between good practice and careful reasoning? What is the great line of thought that leads from Aristotelian virtue ethics via Kant to...
contemporary conceptions of rational ethics, with its central idea of grounding morality in *reason* rather than, say, in authority, tradition, religion, or metaphysics of nature? To be sure, we cannot examine such connections in full here, the way a historian of philosophy might want to do it. My purpose is only to familiarize the reader with a few aspects of Kant's thought that I find essential for our attempt to ground reflective practice philosophically (rather than only scientifically and/or psychologically, as in today's "reflective practice" mainstream). Accordingly, I will not introduce Kant's larger epistemological framework in any detail; the interested reader can find elsewhere a full account of how I understand Kantian *a priori* science (i.e., his transcendental philosophy) and how it relates to my work on critical heuristics (see Ulrich, 1983, chapters 3-5, pp. 175-342).

Regarding Kant's relation to Aristotle and their shared interest in the role of reason for practice, some readers (without meaning to oppose Kant to Aristotle in any schematic way) might wonder at the outset where I locate the most essential difference in their views of that role. I would then say that Aristotle approaches the problem of practical reason primarily as a philosopher of *praxis*, whereas Kant is at heart a philosopher of *reason*. To appreciate Kant, we must first of all learn to appreciate him as a philosopher of theoretical (or speculative) and practical (moral) reason, that is, we must try to capture the critical spirit of his *two-dimensional account of reason* (for an easily accessible summary, see Williams, 2008; for a scholarly commentary on Kant's account of practical reason, see Beck, 1960).

*From the critique of theoretical reason to faith in practical reason:* The starting point of Kant's critical philosophy of reason was his discovery that both in its theoretical and its practical use, human reasoning faced questions that it could neither ignore nor answer. In the *Critique of Pure Reason*, Kant (1787, A341-A566, B399-595) discussed a number of famous examples, the so-called "paralogisms" and "antinomies" of pure reason. In its theoretical use, for example, we have no way to know whether the world has a beginning in time and is limited in space or is infinite; in its practical use, we cannot know how in a world determined by the laws of nature, free will and thus reasonable action are possible. To assure to reason its rightful claims in dealing with such questions – no more, no less – Kant invented a new philosophical method, to which he gave the outrageous name "transcendental philosophy" – a philosophy for dealing reasonably with issues that transcend
the limits of possible knowledge yet are unavoidable.

The answer this method furnished was no less outrageous: we can never know the answers to these questions, yet we can and need to think and act reasonably as if we could know the world as a whole and as if free action were possible. The point is, beyond the theoretical use of reason in science there is another use of reason, its practical use, in which such "as ifs" are perfectly rational, because unavoidable. What matters for practical purposes is not that we know they are true but only, that we make sure they are not in contradiction with what we know. It is sufficient that the theoretical and the practical uses of reason do not undermine one another, and that when it comes to human affairs, we have good grounds to trust in the ability of reason to promote morally good action. In Kant's famous words: "I have therefore found it necessary to deny knowledge, in order to make room for faith." (1787, Bxxx) Of course, what he meant to deny was only the reach of theoretical reason beyond the limits of phenomenal experience; and what he made room for was not just faith but practical reason. Practical reason is the realm of freedom, including the possibility of moral action! Without this denial of theoretical knowledge in matters practical, that is, imposing limits upon its valid use, no realm of freedom could exist – postulating such an "as if" would be inconsistent with the theoretical use of reason – and we would accordingly need to deny the possibility of moral action. Consequently, we would have no possibility to make sure we put theoretical reason to good use; no positive part would be left for reason to play in human affairs. As Kant explained in the preface to the second edition of the Critique:

On a cursory view of the present work it may seem that its results are merely negative, warning us that we must never venture with speculative reason beyond the limits of experience. Such is in fact its primary use. But such teaching at once acquires a positive value when we recognize that the principles with which speculative reason ventures out beyond its proper limits do not in effect extend the employment of reason but ... inevitably narrow it. These principles really only threaten to extend the use of theoretical reason beyond all limits of experience, and thereby to supplant reason in its pure (practical) employment. So far, therefore, as our Critique limits speculative reason, it is indeed negative; but since it thereby removes an obstacle which stands in the way of the employment of practical reason and indeed threatens to destroy it, it has in reality a positive and very important use.... [It means] there is an absolutely necessary practical employment of pure reason – the moral – in which it inevitably goes beyond the limits of sensibility. (Kant, 1787, Bxxivf; my edited version of N.K. Smith's transl., which for once does not match its usual level of accuracy and clarity here)

*From virtue ethics to rational ethics:* For both Aristotle and Kant, the human capacity of reasoning embodies essential aspects of our inner nature –
the soul's sensitivity and rational parts in Aristotle's terms, and the intrinsic
dignity and freedom of will of each person in Kant's terms. Both thinkers
also accept that the pursuit of happiness is an essential motive of all of
us. Since it is not usually obvious to us what makes us truly happy, both
thinkers deal extensively with the question of how we can understand the
pursuit of happiness on rational grounds (as reasonable agents, that is).
Aristotle finds the key in the quest for *eudaimonia*, Kant in the quest for a
*kingdom of ends*, that is, a moral community in which all people respect one
another as ends in themselves (Kant's highest good or *summum bonum*). In
both philosophies, true happiness is thus associated with the idea that well-
being and virtue supported by reason go hand in hand; conducting a good
and virtuous life is one and the same thing. Furthermore, we cannot separate
such a life from the good and just order of the larger communities of which
we are a part – the democratic order of the Greek city state or *polis* in
Aristotle's case and the civil constitution of the *res publica* at the level of the
nation state (a republican constitutional state) and ultimately also at the
international level (a constitution of world citizenship) in Kant's case (cf.
est. Kant, 1795; I have given a summary in Ulrich, 2005). Finally, there is
the decisive role of *reason* in both approaches: both Aristotle and Kant make
it clear that only through careful deliberation and judgment can we recognize
with some degree of certainty what is the right thing to do in concrete
circumstances. Both thinkers, albeit in different ways, boldly ground the
obliging nature of what we *ought to* do in a free decision of reason. Aristotle
refers to this obliging force as "virtue," Kant calls it "duty";¹ but the core
idea is the same, namely, that right action is grounded in right thinking.

But of course, the tasks that Aristotle and Kant assign to reason also differ in
important respects. Conforming to the ethical tradition of ancient Greece,
Aristotle assigns reason the main task of ensuring, through "good
deliberation," that agents learn to make virtuous choices of ends and means;
that is, reason is basically in the service of forming individual character.
Kant, on the other hand, responding to the Enlightenment thinking of the
18th century, assigns reason the main task of recognizing, through "critical
reasoning," the moral principle in each action (i.e., the criterion that ensures
us it is a morally good action); that is, reason is basically in the service of
promoting *universally right action*. With Kant's work, reason becomes the
great emancipating power that is constitutive not only of private but also of
public enlightenment; not only of individual maturity but also of a well-
functioning public domain.

With this shift of perspective, which gradually emerges in Kant's critical writings to culminate in his late political philosophy, a whole new world of Enlightenment ideas enter the discipline of practical philosophy. Particularly relevant to our present purpose are Kant's notions of personal autonomy, maturity, and responsibility; the central notion of good will and its embodiment in "pure" practical reason and in the moral law; the importance of the public realm in this conception of practical reason; and finally, Kant's fundamental – and fundamentally new – principle of practical reason, the principle of moral universalization. To understand its meaning and relevance, we need to begin with a brief clarification of some of the previously mentioned, underpinning notions.

*Autonomy, maturity, responsibility:* "What is enlightenment?" Kant asks us in one of his most famous essays, and his answer is so well known that I hardly need to recall it, except for the sheer pleasure of reading those forceful few lines again:

> Enlightenment is man's emergence from self-imposed immaturity. Immaturity is the inability to use one's reason without another's guidance. It is self-imposed if its cause is not a lack of reason but a lack of resolution and courage to use one's reason without another's guidance. *Sapere aude!* Have courage to make use of your own reason! is thus the motto of enlightenment. (Kant, 1784, A481, my transl.)

Kant does not assume this is an easy thing to accomplish:

> It is so easy to be immature. If I have a book that tells me what to think, a pastor who has a conscience for me, a physician who determines my diet, and so on, I need not trouble myself. I need not think for myself, if only I can pay others to think for me; they will readily undertake the effort for me....

> It is indeed difficult for anyone to overcome such accustomed immaturity.... There are few, therefore, who have succeeded by their own intellectual effort to grow out of immaturity.... (Kant, 1784, A482f, my transl.)

In an equally famous footnote to his later essay "What does it mean to orient oneself in thinking?" Kant offers a more encouraging account:

> Thinking for oneself means to seek the ultimate touchstone for truth in oneself, that is, in one's own reason; and the maxim of always thinking for oneself is [what I call] enlightenment.... To make use of one's own reason means no more than to ask oneself, whenever one is supposed to assume something: Would I find it adequate to make my reason for assuming it, or the rule that my assuming it may imply, a general principle of my thinking? This is a test that everyone can apply for oneself; one will then see unfounded assumptions (superstition) and wishful thinking (exuberance) soon disappear, even though one may not have sufficient knowledge to refute them on objective grounds. For [thereby] one only makes use of reason's own maxim of self-preservation. Individual enlightenment may thus very well be grounded in education; what matters is that we accustom young minds early to such reflection. It takes.
much time, however, to enlighten an entire age; for there are so many external obstacles that forbid such education or render it difficult. (Kant, 1786a, A330, my transl.)

A key phrase in this account is **reason's self-reservation**: proper reasoning is thinking that respects reason's autonomy, along with its intrinsic need for consistency. Since every human being is endowed with the faculty of reason, we are all called upon to make use of it and thus, to respect reason's autonomy and need for consistency in our thinking! The effort that Kant asks us to undertake, then, is that we form our views and decisions _independently_ of preconceived opinions and external pressures to conformity; or in more contemporary terms, that we develop the habit of thinking and arguing _authentically_.

Personal autonomy in this sense is a necessary condition of _reasonableness_; but is it also a sufficient condition? No, Kant tells us with the test he suggests for reasonableness: we must also be able to conceive of our specific reason for assuming something, or for accepting that assumption's implications, as a "general principle" of reasonable thought and action, a principle that everyone could make the basis of their actions. In today's language, the sufficient condition that must go along with autonomy is _accountability to others_, as the obvious similarity of the above test with the categorical imperative clearly suggests. This is so because the notion of a "general" principle implies precisely this: that everyone else might approve it on rational grounds. Authentic thinking thus embodies both, responsibility to oneself (i.e., I take my own thinking and concerns seriously) and to others (i.e., I take their ways of thinking and concerns seriously). Differently put, we are all called upon to respect reason's autonomy not only in our thinking but also in the thinking of others.

It follows that _maturity_ as Kant understands it aims at both _intellectual autonomy_ (thinking for oneself) and _moral autonomy_ (thinking in the place of everyone else so that they, too, are enabled to think for themselves). In this way autonomy, maturity, and responsibility imply one another and are all contained in Kant's concept of reason (and hence, of the individual's enlightenment).

**Pure reason, or reason's intrinsic requirements:** Let us be careful that we understand why exactly autonomy, maturity, and responsibility constitute indispensable aspects of reason and reasonableness. Kant is not talking about
these intellectual and moral virtues from an Aristotelian perspective, as character traits that one may or may not care to cultivate for the sake of becoming a good person, and thereby being worthy of happiness (eudaimonia); rather, he is talking about reason's intrinsic requirements, that is, the conditions of its own possibility. Reason must in all its uses preserve its own integrity, for otherwise it soon loses its argumentative force and credibility. Autonomy is the minimal condition of such integrity, for our reasoning cannot be authentic unless it is grounded in personal independence and in the freedom to "think for oneself." Maturity is the ability to think and act accordingly, so that the assumptions and implications of one's thought and actions withstand the test of elevating each of them to a general principle, or as we might say more simply: maturity is the ability, growing from the individual's quest for enlightenment, to "think and act on principle."

And finally, responsibility is the consequent need to respect and promote the autonomy and maturity of others, for without this consequence one's quest for autonomy and maturity could not serve as a general principle that others can follow for themselves.

I find this reading of Kant's call to reason confirmed in a neglected but remarkable passage on "common sense" in the Critique of Judgment (Kant, 1793, § 40, B158), where Kant sums up his intent as follows:

Under the sensus communis [i.e., well-understood common sense] we must include the idea of a sense common to all, that is, an ability of reflection that considers the ways all other humans may think, in an effort to compare one's own judgment to the collective reason of humanity, as it were, and thus to avoid the trap [orig.: illusion] of allowing one's private conditions of thought, which one might easily mistake for objective, to inform [orig.: affect in a harmful way] one's judgment.…

The following maxims of common human reasoning … may serve to elucidate the basic propositions [that I associate with well-understood common sense]. They are: (1) to think for oneself; (2) to think [as if one found oneself] in the place of everyone else; and (3) to always think consistently with oneself. The first is the maxim of unprejudiced thought; the second of enlarged thought; the third of consequent thought. (Kant 1793, B157f, my simplified transl.; similar formulations can be found in Kant, 1798, § 43, and 1800, end of Sec. VII)

The reflective capacity of reason as it is understood in common sense thus unfolds into a community sense. This is an aspect of Kant's concepts of reason and of maturity (and the link he establishes between them) that is rarely considered in the huge body of secondary literature; it gives a precise and relevant meaning to what Kant somewhat nebulously calls enlarged thought. Taken together, reason's constitutive conditions thus are self-legislation (autonomy) combined with maturity (seeking orientation in self-
chosen principles) and responsibility (accountability to others).

*From good will to pure practical reason:* Kant's (1786b) *Groundwork of the Metaphysics of Morals*, despite its forbidding name, is perhaps the most eloquent and thought-provoking book on the foundations of practical reason ever published, and (along with the *Critique of Pure Reason*) certainly the most influential. At the same time, it is probably also one of the most difficult texts of moral philosophy ever written and for this reason lends itself to different interpretations and translations. I rely mainly on the classical translation by H.J. Paton (1964), along with the original German text (*Grundlegung der Metaphysik der Sitten*) in W. Weischedel's (1968) complete edition of Kant's works; in addition, I find B.E.A. Liddell's (1970) modern version of the *Grundlegung* quite helpful. Kant begins his argument with these famous words, which immediately get us to the heart of the matter:

> It is impossible to conceive anything at all in the world, or even out of it, which can be taken as good without qualification, except a good will.... A good will is not good because of what it effects or accomplishes – because of its fitness for attaining some proposed end; it is good through its willing alone – that is, good in itself. (Kant, 1786, B1-3)

The crux of the problem of grounding ethics consists in the question of how reason can identify and justify an action as "good" (i.e., as the right thing to do). There are only two ways in which this is conceivable, Kant tells us: either, because the action serves to accomplish some other good that is presupposed to be good, or else because this way of acting is good in itself, that is, it has an unconditional quality of being right, in the sense that it may be said to be good under all circumstances. Only this second way can furnish a sufficient foundation for ethics; for anything else would mean to try to ground ethics in mere expediency, that is, in an action's usefulness with a view to some other good. That would not only beg the question of what constitutes good action; it would indeed make ethics redundant. Expediency – instrumental efficacy – serves whatever ends and means we choose, regardless of whether we are guided by a good will. Against such plain relativism, Kant maintains that there must be some less subjective and self-serving form of reasoning about the ends and means of justified action. "Drop the ego!" is perhaps the most basic intuition underlying all ethics, including rational ethics and its quest for grounding – and orienting – good will in reason. This is how I would basically translate Kant's central
This intuition of holding back the ego (which is not the same as denying it) makes it understandable why classical ethics was preoccupied with psychological and educational questions of character and thus was conceived primarily in terms of virtue ethics. To this preoccupation with character, the Medieval scholastics later added religion (i.e., faith) as a basis for explaining the binding character of the moral idea, which in effect moved ethics further away from a grounding in reason than it had been in Aristotle's work. But just as a theological grounding of ethics is helpful only for those who believe, as it presupposes faith without being able to create it, virtue ethics as a methodological (though not as an educational) approach, too, tends to presuppose what it aims to produce: moral character and good will. As a theory of good practice, it ultimately relies on an appeal to the good will (or in Aristotle's framework, the good character) of agents; for whether or not an agent will in a given situation act virtuously depends on his being goodwilled – an act of faith that cannot be grounded any further but at best be encouraged through education and custom.

Kant's solution to this difficulty is ingenious: he places the origin of a good will within the nature (or concept) of reason itself. To avoid a possible misunderstanding, by reason's nature he does not mean its empirical state or development in the individual but rather, those general structures and requirements which characterize it by inner necessity (a priori), in the sense that reason cannot operate without them (e.g., reason cannot help but regard itself as free, 1786b, B101). In other words, he refers to the nature of pure reason – the sheer idea of reasonableness regardless of its empirical occurrence in individuals. Unlike all previous ethics, including Aristotle, Kant does not assume that the binding force of the moral idea needs to reside in some external psychological or religious condition such as character, faith, or virtuous action. Rather, he understands it as residing in one of the most fundamental conditions of reason itself, conditions that reason cannot ignore except at the price of becoming inconsistent with itself and thereby losing its authority or credibility. The condition at stake is the third of the three maxims of reason quoted above (Kant 1793, B157f), the requirement of consequent (or coherent) thought. Living up to its own intrinsic requirements is what Kant calls a "pure" interest of reason or "interest of pure reason" – an interest that has no other aim than preserving the possibility of reason itself,
its "self-preservation" (1786a, A330).

The implication of this new concept of pure reason is powerful indeed: in its practical no less than in its theoretical employment, reason is itself in charge of the conditions of its successful operation. We can only recognize as true, both in an empirical and in a moral sense, what our mind creates itself; or in Kant's (1787, Bxiii) words, "reason has insight only into that which it produces after a plan of its own." In the Preface to the second edition of the Critique of Pure Reason, Kant (1787, Bxvi and Bxxii) famously referred to this shift of perspective as the Copernican revolution of speculative philosophy, or as we would rather say today, of epistemology. Copernicus was the first astronomer to recognize very clearly that counter to the observed (phenomenal) movements of the planets around the earth, their true (noumenal) movements were ellipses around the sun. He was able to achieve this revolution of our worldview because he "dared, in a manner contradictory of the senses, but yet true, to seek the observed movements, not in the heavenly bodies, but in the spectator." (1787, Bxxii note). A similar shift of perspective now is to inform Kant's revolution of practical philosophy. Kant himself does not say so – he refers to the Copernican revolution only in his critique of theoretical reason – but I would argue that the notion of a Copernican revolution of ethics provides a very immediate and helpful key to the core of Kant's concept of practical reason: the reason why we ought to act morally is not because some external authority obliges us but simply because we recognize such action to be reasonable. The moral force resides in our will to be reasonable! It is, in the language of the first Critique, "reason's own plan." In a short formula (again "translating" Kant into the language of our epoch):

\[
\text{We ought to act morally because we want to be reasonable.}
\]

Can there be any more unconditional and stronger source of motivation than the mere wish to be reasonable rather than unreasonable, and thus also to be taken seriously and listened to by others? No-one ever claims to be unreasonable; everyone wants to have reason on their side. This is the basic innovation in Kant's conception of ethics as I see it: with Kant, ethics first becomes practical reason in a methodologically compelling sense.

Pure practical reason as a mere limiting concept, or the vain search for an absolute beginning: One may wonder whether Kant somehow lost sight of
the above-mentioned consequence of his own "Copernican" approach when in the last chapter of the *Groundwork*, we find him searching for some mysterious absolute source of the binding force of the moral idea, a source that would explain why pure reason, before and beyond all empirical motives, is compelled to be moral. Such a force, if it really existed, would need to be independent of all human willing and reasoning and thus external to our mind, if not external to all nature (transcendent) – an implication that runs counter to Kant's core idea of grounding ethics in reason. I therefore tend to understand "pure" reason as a mere *limiting concept*, an admittedly unreal (nonempirical) ideal-type of reason that serves Kant to undertake his great experiment of thought, the experiment of submitting reason to its own tribunal (cf. 1787, Axif). Thus seen, Kant's ultimate and vain effort of finding an absolute source of universal moral obligation is a remarkable testimony to his relentless selfcritical determination to push his inquiry to its utmost limits, even if such an effort is ultimately bound to fail:

But how pure reason can be practical in itself without further motives drawn from some other source; that is, how the bare principle of the universal validity of all its maxims ... can by itself ... supply a motive and create an interest which could be called purely moral; or in other words, *how pure reason can be practical* – all human reason is totally incapable of explaining this, and all the effort and labour to seek such an explanation is wasted. (Kant, 1768b, B124f)

But Kant's effort is far from wasted. Without apparently being fully aware of it, he actually uncovered that there is no need at all for such an explanation. The fact that a reasonable agent *wants* to act morally (i.e., to act out of good will) is quite sufficient for saying he *ought* to do so; for anything else would undermine the integrity of reason. It belongs to the peculiar force of reason in its "pure" form, that whatever it makes us want, we ought to do. Hence, if as a reasonable being I want to act morally, I ought to do it; and conversely, if reason tells me I ought to do it, as a reasonable being I want it. This, then, is the core idea of a *rational ethics* as Kant conceives it: the force of the moral idea resides at bottom in the power of reason, and that must be quite good enough for us as reasonable beings; for being reasonable is the same as being *willing to follow the call of reason*. As Kant makes us understand (but temporarily disregards in his search for an absolute source of moral obligation), the moral idea is an immanent rather than a transcendent idea of reason.

*Practical reason and the public realm:* There is another important
implication of reason's intrinsic need for self-preservation: its necessarily public nature. As is the case with the Copernican shift of perspective, Kant argued the point for theoretical reason but it applies equally to practical reason, despite being rarely mentioned in connection with it:

Reason must in all its undertakings subject itself to criticism; should it limit freedom of criticism by any prohibition, it must harm itself, drawing upon itself a damaging suspicion. Nothing is so important for its usefulness, nothing so sacred, that it may be exempted from this searching examination, which knows no respect for persons. Reason depends on this freedom for its very existence. For reason has no dictatorial authority; its verdict is always simply the agreement of free citizens, of whom each one must be permitted to express, without let or hindrance, his objections or even his veto. (Kant, 1787, B766f)

I would argue that this requirement characterizes reason both in its theoretical and in its practical employment; but in practical reason, it gains particular significance. It now implies that accountability to others is constitutive of practical reason. Being reasonable and accepting accountability are inseparable! When Kant (1784), in his earlier-quoted essay on the meaning of enlightenment, urges us to "have courage to make use of your own reason," he really asks us to think as members of a society of world citizens (Weltbürgergesellschaft). There is no natural end to the scope of our accountability, in that everyone concerned belongs to the moral community to which we are accountable. Excluding people concerned from this community would mean to disregard their autonomy and dignity as reasonable beings, and thus to treat them in a way that we would not want them to treat us. We begin to understand why Kant, in the previously mentioned passage of the Critique of Judgment on the meaning of well-understood "common sense" (Kant, 1793, § 40, B158), points out that of his three maxims of reason, the third, of always thinking consistently with oneself, is the most difficult to put into practice. "It can only be achieved by jointly achieving the other two" (1793, B159f), that is, by thinking for oneself and putting oneself in the place of every other reasonable being that might be concerned. It means that our thinking as reasonable agents must address the citizens of the whole world, or in one word, become universal.

The principle of moral universalization: We arrive, then, at the most fundamental contribution that Kant has made to practical philosophy, I mean his formalization of the moral idea in terms of the principle of moral universalization or, as he calls it, the categorical imperative. It is so important that I will dedicate more space to it than to the previous points.
There are different versions of the imperative, but the underlying principle remains the same and is best (in the sense of most generally) expressed in the imperative's basic form, the so-called

*Formula of Universal Law:*

"Act so that the maxim of your will could at all times hold as a principle of general legislation."

(Kant, 1788, A54, my transl.)

or simpler:

"Act according to a maxim that you could want to become a universal law."

(Kant, 1786b, B52 and B81, my simplified transl.)

or still simpler:

"Act only on a premise that can be everyone's premise."

(Kant, 1786b, B52 and B81, my free transl.)

The underlying intuition is not difficult to grasp, nor is it Kant's invention. It is popularly known since ancient times as the *golden rule,* for example in the biblical command "Do unto others as you would have them do unto you" (e.g., Matthew 7:12, Luke 6:39) and in similar form in virtually all world religions. It asks us to treat others the way we would want them to treat us, or negatively put, not to do to others what we do not want to be done to ourselves. The core idea is *reciprocity* of mutual respect and consideration, an expectation grounded in our shared human nature; our equal claim to dignity and free will; our equal sensitivity and vulnerability, in short: our human condition. Reciprocity is the common normative core of notions such as fairness and security in human relationships, cooperation and peace, solidarity and compassion, barter and contract, international law and human rights, and so on. It is one of the few moral concepts of which we can say that it appears to correspond to a truly universal moral intuition, one that is present in all epochs and cultures.
Contemporary research in social psychology, sociology, ethnology, moral education, and other fields has shown that the idea of an *ethics of reciprocity* is indeed a universal concept that can be found in all cultures and epochs. Before entering into a discussion of Kant's intent, it may be useful to consider what this research can tell us about the universality of the idea of moral universalization. A helpful place to begin with is George Herbert Mead's (e.g., 1913, 1925, 1934) work on "symbolic interactionism," with its central question of how we form our personal identity as members of society, our "social self." The main source is *Mind, Self and Society*, a book that some of his students published in 1934, three years after his death, based mainly on lecture notes and other unpublished material, as Mead himself published numerous articles but never summarized his ideas in a book; for a selection of some of his published papers, see Reck (1981).

Some of Mead's (1934, pp. 152-226) basic work was on the importance that playing and exchanging of roles in games (along with communication and activities in family life) have in the development of the child's self-concept. Through games and other forms of symbolic interaction (gestures, words, etc.), children learn that they are expected to meet certain norms of behavior, the "rules of the game," not only in playing but also in their relation to parents and other family members, in kindergarten, school, and so on. In a children's game, for instance, every participant must respect the roles of *all others* playing the game, along with the rules that shape these roles. *One has to play the game:*

> What goes on in the game goes on in the life of the child all the time. He is continually taking the attitudes of those about him, especially the roles of those who in some sense control him and on whom he depends.... He has to play the game. (Mead, 1934, p. 160)

Gradually, in the course of this process of socialization, children learn to distinguish between "me" and "I." The "me" is the emerging notion of all the different expectations that others have towards the child, and the "I" forms in response to it. "The self appearing as 'I' is the *memory image self* who acted toward himself [the 'me'] and is the same self who acts toward other selves." (1913, p. 375, italics added). A sense of identity, as the confluence of "me" and "I," thus develops in reaction to the child's self-image in the eyes of *all* others. Mead describes this internalized attitude of all others with his famous notion of the *generalized other* – the sum-total of all the experienced roles and role expectations of others that gradually become an...
internalized part of the individual's self-consciousness, an internalized attitude (e.g., 1925, pp. 268f and 272; 1934, pp. 154-156 and 196). It is this cumulative notion of the generalized other that later enables the child, in adult life, to anticipate roles and role expectations in new situations, a process that Mead describes as an "internal conversation of the individual with himself … from the standpoint of the 'generalized other.'" (1934, p. 155n)

In this capability of role taking, of taking the roles and attitudes of others in all situations of communication, cooperation or competition with others, Mead recognizes the socio-psychological and micro-sociological kernel of morality. It is constitutive of the sense of reciprocity that he finds embodied in Kant's categorical imperative. In it he finds a plausible explanation for the universality that Kant associates with moral judgments: because they originate in our social self, they are never a merely private affair of the "I" but always also speak to the generalized other of the "me." Thus we can understand why the sense of reciprocity, despite its origin in the role expectations and conventions of the social microcosmos of early childhood, develops in the mature individual into a universalizing kind of attitude, an orientation that frees itself from the conventions and pressures of the "here and now" in favor of an attitude of acting on principle.

Maturity in a socio-psychological perspective such as Mead's and in an ethical perspective such as Kant's converge in an orientation towards the generalized other and consequently, towards universal principles of action.

The universalizing thrust of the categorical imperative thus finds a socio-psychological and sociological foundation. Not only is man a social being because he is a rational being, as Kant assumes, but conversely "man is a rational being because he is a social being," namely, because both his mental faculties and his identity are socially constituted:

Man is a rational being because he is a social being. The universality of our judgments, upon which Kant places so much stress, is a universality that arises from the fact that we take the attitude of the entire community, of all rational beings. We are what we are through our relationship to others…. We can agree with Kant that the "ought" does involve universality…. Wherever the element of the "ought" comes in, wherever one's conscience speaks, it always takes on this universal form. (Mead 1934, p. 379f)

And further:

Only a rational being would be able so to generalize his act and the maxim of his act, and the human being has such rationality. When he acts in a certain way he is willing that everyone should act in the same way, under the same conditions. Is not that the statement we generally make in justifying
ourselves? When a person has done something that is questionable, is not the statement that is first made, “That is what anyone would have done in my place”? Such is the way in which one does justify his conduct if it is brought into question at all; that it should be a universal law is the justifiable support that one gives to a questioned act.’ (Mead, 1934, p. 380, italics added)

The subsequent history of thought continued in two directions: empirical corroboration and conceptual refinement. Inspired not only by Mead but also by Jean Piaget's (1932, 1970) research on the intellectual and moral development of children, with its central notions of "stages of cognitive development," "sociogenesis of moral judgment," and "genetic epistemology" (an early form of constructivist and evolutionary epistemology), Lawrence Kohlberg (1976, 1981, 1984) studied the development and structure of moral conscience empirically, through qualitative interviews on moral issues with children and adolescents, and then analyzed his findings both psychologically and philosophically. He analyzed them psychologically by expanding Mead's and Piaget's theories of cognitive development in the light of his findings, and philosophically by relating the resulting stage theory of moral development to philosophical ethics, especially to Kant's (1786b, 1788) principle of moral universalization and to Rawls' (1971) theory of justice. Kohlberg summarized his conclusions in terms of four levels of moral development: the preconventional level of the small child, the conventional level of the socialized child, a transitional level of the adolescent, and the postconventional level of the mature adult (Table 1; some of the terms in the right-hand column are my own).

<table>
<thead>
<tr>
<th>Level one: Preconventional morality</th>
<th>Stage 1: Punishment-obedience orientation (&quot;How can I avoid punishment?&quot;)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level two: Conventional morality</td>
<td>Stage 2: Self-interest orientation (&quot;What's in it for me?&quot;)</td>
</tr>
<tr>
<td></td>
<td>Stage 3: Conformity orientation (&quot;Good boy&quot; / &quot;nice girl&quot;)</td>
</tr>
<tr>
<td></td>
<td>Stage 4: Social norms orientation (&quot;Law and order&quot;)</td>
</tr>
<tr>
<td>Transitional level:</td>
<td>Stage 4½: Subjectivist-skeptical orientation (&quot;I do my own thing&quot;)</td>
</tr>
<tr>
<td>Level three: Postconventional morality</td>
<td>Stage 5: Social contract orientation (&quot;I agreed to it&quot;)</td>
</tr>
<tr>
<td></td>
<td>Stage 6: Universal ethical principles orientation (&quot;Can this be a general principle of action?&quot;)</td>
</tr>
</tbody>
</table>

Table 1: Kohlberg's stages of moral development
At the preconventional level of the small child, behavior is regulated by the anticipation of pleasure and pain; at the conventional level of the well-socialized older child, by conformity and loyalty to norms of the family and the social groups to which one belongs (peers, teachers, etc.); and at the postconventional level of the mature adult, by principles of behavior agreed to and/or recognized as valid autonomously. Between levels two and three there is a transitional level of young adults who waver between these two levels, as their behavior is shaped by growing disrespect for merely conventional morality imposed by others but not yet supported by clear ethical principles of their own.

The first two and the fourth levels are each structured into two stages, so that the development of moral consciousness can on principle be characterized and assessed in terms of six (or seven) stages of development. In line with Piaget's understanding of stages of cognitive development, Kohlberg describes his stages of moral development as qualitatively different ways of thinking (cognition) about moral issues that must be learned by each individual (i.e., they are not an automatic biological development but grow through reflection about cognitive dissonances experienced with previously reached stages). Furthermore, like Piaget's stages, they unfold in an invariant, consecutive, and irreversible sequence (i.e., children do not skip stages or regress backward in stages) and are cross-culturally universal (i.e., moral development follows this same order in all societies). Finally, they represent a hierarchy in the sense of embodying increasingly broader ways of seeing moral issues, that is, the lower levels are preserved in the higher ones but become subordinated to wider considerations – a cognitive development from egocentric to social, from intuitive to rational, and from heteronomous to autonomous moral judgment.

Despite continuing and partly controversial discussion of Kohlberg's work, particularly with respect to its cross-cultural universality claim and its possible neglect of gender differences, its main findings are widely accepted today. In essence they confirm the importance of Mead's notion of role taking in the moral development of the individual, as well as the universal thrust which this ability of taking the role of others assumes with growing maturity. Mature morality, then, is autonomous moral reasoning informed by universal role taking, a finding that comes remarkably close to Kant's understanding of morality, despite the fact that Kant does not mean to
ground it empirically in the "nature" (or socio-psychological constitution, as we might now say) of human reason.

I would not say though that Piaget and Kohlberg's studies "confirm" Kant empirically, as both apply a Kantian framework in the first place and then interpret their empirical results in its light. Likewise, the central methodological idea with which Kohlberg associates his stage six, Rawls' (1971, §§ 3, 24, and 40) "veil of ignorance," is based in Kant. While it is problematic, then, to see in Piaget and Kohlberg's developmental moral psychology an empirical corroboration of Kantian ethics, we might perhaps say with Popper (1959/2000) that Kohlberg's empirical findings at least do not "falsify" or run counter to Kant's conception. In addition, as Habermas (1990, p. 117) suggests, "Kohlberg's theory of moral development offers the possibility of (a) reducing the empirical diversity of moral views to variations in the content, in contrast to universal forms, of moral judgment and (b) explaining the remaining structural differences between moralities as differences in the stage of development of the capacity for moral judgment." I may not be thoroughly convinced, but in any case the point of this brief excursion into moral psychology is not to validate Kant's approach (I do not believe philosophical conceptions can be "proven" empirically) but merely to elucidate the central intuition of reciprocity in human interaction that underlies the categorical imperative, and perhaps to give some plausibility to it.

Another, related intuition underlying the categorical imperative is certainly the idea of impartiality. We find it most prominently in one of Kant's sources of inspiration, the writings of the Scottish philosopher and economist Adam Smith (1795). One of Smith's central concepts is the notion that ethics is an effort to examine our actions in the eyes of an imagined impartial spectator (1795, e.g., pp. 16-19, 33, 66, 98, 130, and passim). That is, ethical justification is the observance of "what every impartial spectator would approve" (1795, p. 78, cf. pp. 100 and 343). Smith emphasizes that taking the role of an impartial spectator involves both moral sentiments (esp. sympathy, the core of which is placing ourselves in the other's situation; cf., e.g., 1795, pp. 1f, 72, and passim) and moral reasoning (1795, pp. 329-331), an aspect that makes Smith's ethical philosophy a bridge between classical virtue ethics and modern rational ethics. This may explain why Kant explicitly (but without mentioning Adam Smith) refers to the notion of a "rational and
impartial spectator” right at the beginning of the *Groundwork*, immediately after introducing his core concept of a "good will" (1786b, B2), although he hardly ever uses it again later on in his writings on practical philosophy.

More recently, the same intuition of impartiality has been in the center of John Rawls' (1971) *Theory of Justice*. He observes that universalization as Kant conceived it may not be enough to ensure moral impartiality; for even when in our minds we put ourselves in the places of others concerned, we still do so as "ourselves," that is, we see the world as conditioned through our own position in life and are accordingly biased. Hence, Rawls argues, strict universalization requires that we free ourselves from this sort of bias; we need to universalize our maxims from the point of view of an imagined (and idealized) *original position* in which we do not yet know what particular position life will offer us. A *veil of ignorance* is to stop us from seeing the world (and the situation of others) through the lens of our own more or less privileged situation as we acquire it through birth and efforts of our own making. If principles of justice are to be fair, they must be fixed from behind such a veil of ignorance (1971, § 4 and 40).

From a Kantian perspective, Rawls' veil functions as a procedural and institutional corrective to the usual dominance of subjective motives in moral reasoning, which is what the categorical imperative is all about. Kant approaches the same issue from the other side, as it were, by concentrating on the limiting case of "pure" reason as the arbiter of moral decision making. Pure reason by definition is autonomous and thus not conditioned by "heteronomous" (empirically conditioned, varying) individual ends, desires, inclinations, and (as Rawls would add) "positions" or life-situations. Autonomy *qua* pure reason, then, is Kant's way of formulating the crucial motivational presupposition of impartiality, or with Rawls, *fairness*: moral reasoning needs to judge from a standpoint of mutual disinterest and unconditional mutual respect – the *moral point of view*. Rawls' work thus adds to our understanding of the central intuitions embodied in the categorical imperative. Despite all the differences and tensions between Kant's and Rawls' conceptions of ethics, they share the basic intuition of an impartial spectator or agent inherited from Adam Smith, and can thus be understood as different attempts to reconstruct this intuition methodologically.

A small *caveat* is in order before we leave the topic of the central intuition.
that informs the categorical imperative. Kant may have good reasons for hardly explaining this intuition and instead focusing on his abstract, rational construction of "pure" practical reason. As I have already noted with reference to Piaget and Kohlberg's work, I do not believe psychological research (or empirical science in general) can validate a philosophical conception such as Kant's. It can at best give it some plausibility and elucidate some of its implications (e.g., for moral education). It should be equally clear that reference to underpinning intuitions does not replace methodological argument. Rather the contrary: the attempt to ground ethics in moral intuition, if stretched beyond its limits, risks undermining the constructive methodological effort that is called for to show how practical reason works and can be practiced systematically. The reader may indeed have noted a certain irony in Rawls' interpretation of Kant: it renders the categorical imperative even more abstract than we encounter it in Kant's work; for the "original position," despite its procedural and institutional implications, is no less an idealized limiting concept than Kant's "purely" rational agent. That is, we now need to idealize our quest for moral competence even more than before: as perfectly rational agents we not only need to imagine ourselves in the place of all others possibly concerned (the generalized other) but moreover should now do so from behind a fictitious veil of ignorance – a tall order. Insofar as (with Mead and Kohlberg) knowing the generalized other and knowing ourselves are inseparable and constitutive elements of our mental faculties, the veil of ignorance risks undermining rather than supporting Kant's intent of "enlarging" our thinking (the second of the three maxims of enlightened thought, as quoted above from Kant, 1793, B158); for how can I know myself in the eyes of others (or with Mead; my "social self") through a veil of ignorance? Moral reasoning thus, if we are not careful, risks becoming a bloodless affair with no relation to who I am and to the social contexts of action and responsibility in which I find myself.

**The variations of the categorical imperative:** Let us now return to Kant and see how he deal with the somewhat bloodless, because purely formal, character of the universalization principle. How does he try to convey its intent to his readers? And why does he insist on a purely formal principle in the first place? I'll begin with the first question and then consider the second question in the subsequent section.
One way in which Kant tries to add some substance to the basic formulation of the categorical imperative, so as to render it easier to understand and to apply, is by reformulating it in terms of three variations (Table 2).

<table>
<thead>
<tr>
<th>Formula (traditional name)</th>
<th>Imperative (Kant's formulation)</th>
<th>Suggested reformulation (or short imperative in contemporary language)</th>
<th>Suggested short name (or imperative)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formula of universal law (basic formula)</td>
<td>&quot;Act so that the maxim of your will could at all times hold as a principle of general legislation.&quot; (1788, A54)</td>
<td>&quot;Act according to a maxim that you could want to become a universal law.&quot; (cf. 1786b, B52 and B81) or short imperative: &quot;Act according to a principle that can hold generally!&quot;</td>
<td>Universalization formula (the imperative of acting on principle)</td>
</tr>
<tr>
<td>Formula of the law of nature (first main variation)</td>
<td>&quot;Act as if you wanted the maxim of your action to become a law of nature.&quot; (1986b, B52)</td>
<td>&quot;Act so that your maxim could be a law of nature.&quot; (cf. 1786b, B81f) or short imperative: &quot;Design with nature!&quot;</td>
<td>Natural law formula (the ecological imperative)</td>
</tr>
<tr>
<td>Formula of humanity as an end in itself (second main variation)</td>
<td>&quot;Act so that you always treat humanity, whether in your person or in another, as an end and never merely as a means.&quot; (1986b, B66f and B74f)</td>
<td>&quot;Act so that your maxim considers every rational being as an end it itself, never merely as a means.&quot; (cf. 1786b, B82) or short imperative: &quot;Respect people!&quot;</td>
<td>Humanity formula (the imperative of human dignity)</td>
</tr>
<tr>
<td>Formula of autonomy in a kingdom of ends (third main variation/combined formula)</td>
<td>&quot;Always act as a self-legislating member of a kingdom of ends, that is, a community of rational and self-legislating beings.&quot; (1986b, B75)</td>
<td>&quot;Act so that your maxim could serve as a general principle for all rational beings.&quot; (cf. 1786b, B84) or short imperative: &quot;Act as a world citizen!&quot;</td>
<td>Community formula (the imperative of societal responsibility)</td>
</tr>
</tbody>
</table>

It is striking to see that these three variations closely parallel a previous attempt, about 1800 years earlier, to formulate some principles or rules of conduct that might help us to resolve the eternal conflict between personal interest and duty, between the useful and the honorable. The attempt was undertaken by the Roman statesmen and philosopher Cicero (44 B.C., Book III, Sections 13, 26, and 28) in his work De Officiis (= On Obligations), and it led him to suggest three such rules:

1. *Convenienter naturae vivere:* "Properly live according to nature."
2. *Omnino hominem ex homine tollat:* "Honor every person as a human being."
3. *Communis humani generis societas:* "See yourself as a member of the universal human community."

The obvious parallel between Cicero's and Kant's imperatives is not widely...
known, as Kant does not refer to Cicero at all. He may well have formulated his three variations independently, due to common sources in the works of the Stoics, in whose ethics Kant was interested and with which Cicero was of course familiar. I think it is more likely though that Kant consciously adopted Cicero's rules as variations of the categorical imperative. He may have chosen not to refer to Cicero lest he made it look as if the universalization principle could be derived from such rules of thumb; he may nevertheless have adopted them as he realized they served to illustrate the categorical imperative. In any case, Cicero's rules have helped (and encouraged) me to translate Kant's somewhat tedious formulas into more contemporary language, and that is why I quote them above. They yield three short imperatives that roughly correspond to Kant's three variations yet should make it easier for the reader to grasp the idea of "universalization," as well as to remember the different forms Kant gave to the universalization principle. Conversely, Kant's work allows us to appreciate the internal link between Cicero's three rules: they express the same underlying principle of universalization, in a way that looks amazingly modern and as if Cicero had anticipated Kant's work.

The fundamental concern of all the formulations suggested in Table 2 is indeed the same. To judge the moral quality of an action or action proposal, we should first ask ourselves what is the underpinning maxim. In Kant's language, a maxim is a subjective rule or norm of action (i.e., a personal premise), while a principle is an objective, because generally binding or "necessitating" rule or norm of action (i.e., a premise that everyone may be expected to make the basis of their actions). Kant's point in distinguishing the two is that he does not want us to presuppose that our individual premises are naturally in harmony with principles that everyone could hold; quite the contrary, the problem of practical reason emerges from the divergence of the two perspectives (cf. 1786b, B37f and B102f). It is because individual and collective premises do not usually converge that Kant asks us to consider what our personal premises are and to what extent they might be the premises of all others concerned. Insofar, the categorical imperative, counter to what is often assumed, cannot be said to be idealistic.

Once we are clear in our mind as to what is the underpinning maxim of an action, the consequent next issue is whether this maxim could hold as a principle of action, in the precise sense just defined. To answer this question,
we may assess it against the categorical imperative or any of its variations, as summarized in Table 2. If our maxim runs counter to any of these imperatives, it is not an adequate principle, for it cannot be properly universalized. This makes it understandable why Kant calls his three variations "equivalent," despite their different contents. Like Cicero's rules, they address different object-domains (of nature, humanity, and society); but they are equivalent in the sense that they yield the same result, that is, help us to examine the extent to which any specific maxim (or norm) of action is morally generalizable.

In conclusion, for all those who find Kant's formulations of the universalization principle in terms of the categorical imperative and its variations hard to remember and to use, it may help to employ these three short imperatives instead, as a kind of shorthand for the idea of moral universalization:

1. The imperative of ecological sustainability: "Design with Nature!"
2. The imperative of human dignity: "Respect people!"
3. The imperative of global societal responsibility: "Act as a world citizen!"

To be sure, none of these three imperatives should be taken for an accurate reformulation of Cicero and Kant's intent; for example, when Cicero and Kant use the term "nature," they may both have thought primarily of doing justice to human nature rather than of ecological issues. But we live in a different epoch and have to ask ourselves what the principle of universalization means to us, here and now. Taken together, the three imperatives may indeed constitute something like a nucleus of moral thinking for our time; an ethos of acting on principle that remains faithful to the idea of moral universalization but renders it easy to use. Moral reasoning must not remain the privilege of philosophers.

Why a purely formal moral principle? We have seen that Kant's basic idea consists in requiring rational agents, as individuals endowed with reason and free will, to act in ways that respect these same qualities in others as well as in themselves. This is a normative idea that we cannot substantiate through empirical research but only through an appeal to reason's need for being consistent with itself. But why does Kant not openly declare this normative, humanistic intent? Why does he insist on a purely formal principle, although it then causes him so much effort to explain its meaning? We are now
Kant aims to discover a single, truly universal principle of moral reasoning. Any more modest ambition would imply that ethical practice could be based on alternative fundamental principles, so that moral judgment would remain empirically and normatively contingent, that is, depend on choices conditioned by our opinions and interests of the moment. A secure grounding of ethical practice would then remain out of reach. As a rigorous thinker, Kant has no choice but to search for a single, “categorical” (i.e., unconditional, absolute) principle of morality. Consequently, he needs to eliminate empirical and normative presuppositions as much as possible. He first eliminates all empirical presuppositions except the “rational” character of moral agents, by focusing on the limiting case of pure reason. All other, empirical considerations he leaves to what he calls “pragmatic” rather than moral reasoning (1787, B828). Next, he needs to minimize normative presuppositions, which however is not possible in the same radical way. Although Kant does not say it in these words, the reason is that there is no action without end; we speak of action as distinguished from mere behavior precisely when we consciously pursue ends. We act morally when we do so with a moral purpose in mind, that is, for the sake of some unconditional good. How, then, can Kant define the unconditionally good without presupposing too much? He must catch two flies at once, as it were: he needs to give a strong normative thrust to moral action yet keep its normative presuppositions to an absolute minimum, lest it become a conditional sort of action.

This is where the earlier discussed intuitions of reciprocity and impartiality become important. They embody the minimal normative core that Kant can associate with the ideal of an unconditionally good will without endangering his project. Kant resolves the dilemma by associating the minimal normative core of a morally good maxim – of a "good will," that is – exclusively with its form, rather than with its content. That is, we reason and act morally well not because we pursue some specific ends or rely on some specific virtues or methods, but simply because we think consistently: we are to grant others the same degree of free will, the same claim to personal autonomy and dignity, that we take granted for ourselves (and which, in fact, we cannot help but presuppose in deciding about our maxim of action). In this unconditional reciprocity and impartiality, which obliges us to recognize in other human
beings the same needs for dignity, respect, and fairness of treatment that we claim for ourselves, Kant finds the constitutive, but minimal, normative core of the categorical imperative. In an ingenious way, Kant manages to give normative substance to his universal principle without compromising its universality.

Kant's formal principle of moral reasoning, then, does not really propose a new basic intuition as to what ethics is all about. The intuition remains the same as humanity has known it for millennia in the form of the golden rule, the biblical ten commands and other religious testimonies, Cicero's three rules of conduct, Adam Smith's impartial spectator, and so on. Rather, what is new is the way Kant uses this intuition to ground ethical practice in practical reason. He is the first and thus far (as far as I can see) only philosopher who has reconstructed the basic intuition of reciprocity in terms of a single, strictly universal principle of moral reasoning; a principle that is as close to being independent of all empirical and normative presuppositions (except the will to reason) as is conceivable. But Kant's methodological reconstruction of the intuition is novel in other ways, too. On the one hand, Kant specifies and narrows down the issue to one of judging and justifying the basic principles informing our actions, rather than directly the actions themselves (e.g., the specific ends we pursue, the means we rely on, and/or the consequences our actions may have). On the other hand, he simultaneously extends the reach of reciprocity to all of humanity, here and there, now and in future. Furthermore, he adds a new degree of rigor and precision, by transforming the intuition into a methodologically ingenious "test" that every agent of good will can apply before acting.

**Personal appreciation: the idea of moral universalization**

Kant's preferred way of describing the universal (or "objective," as he likes to say) yet normatively significant and obligating ("necessitating") nature of the universalization principle is by analogy with "the law." He wavers a bit between the law of the state (legal norms) and the law of nature (natural laws); but for our purpose it suffices to take legal norms as the basic metaphor and natural laws as a derived, more illustrative metaphor (for metaphors they are both, just like the concept of "duty" that he derives from it and which similarly stands for an unconditional obligation legislated by our own will). As Kant sees it, the principle of moral universalization obligates us not unlike the way a legal norm obligates everyone under its
jurisdiction. The difference is that a legal norm obliges us only conditionally, namely, to the extent that we belong to the community of individuals that have given themselves such legislation and that there is no applicable legislation of superior authority that poses different demands. By contrast, a moral norm (or "moral law") applies unconditionally or categorically – it is the ultimate source of obligation beyond which we cannot refer to any other, supposedly superior source of obligation.2)

My own preferred way of thinking of the unconditional moral thrust of the universalization principle is in terms of never treating others merely as a means for one's own ends, or in short, of an imperative of non-instrumentalization:

"Do not instrumentalize other people!"

This reading draws on Kant's second variation of the categorical imperative, in combination with Cicero's second rule of conduct cited above. Expressed in its simplest form, it amounts to the short formula suggested in Table 2: "Respect other people!" To be sure, only in combination with the other two other imperatives that we derived from Kant's variations of the categorical imperative (along with Cicero's rules) does this imperative of non-instrumentalization fully capture the moral thrust of "universalization"; but I find in this notion of non-instrumentalization a humanistic core that for me comes closest to a truly universal norm of human conduct. Furthermore, to overcome any anthropocentric bias one might object against, we may apply it not only to people but equally to animals and all living nature: "Do not instrumentalize other living beings!"

Finally, readers who prefer the basic formula but expressed in contemporary language, might opt for the reformulations offered in Table 2:

"Act according to a maxim that you could want to become a universal law."

(cf. 1786b, B52 and B81; read "law" as "principle")

or simply:

"Act according to a principle that can hold generally!"

But of course, no kind of simpler formulation can alter the fact that the universalization principle makes strong demands on us. We do not usually act in ways that do justice to everyone. Universalization is an ideal rather than a practical premise. Nor does the universalization principle tell us
anything about what our premises should be; necessarily so, as these depend on the situation. Hence, while the suggested reformulations make Kant's intent fairly easy to grasp, they do not tell us how to act accordingly. How can we make sure our actions conform to the Kantian imperative?

The answer may well be: we cannot! Moral perfection is not usually achievable by human agents, not even through inaction. Thus seen, the categorical imperative, unlike what Kant assumed, is not really a guideline for determining (much less for justifying) moral action. It cannot assure us of acting morally, only help us identify moral shortcomings of the way we act. Universalization thus becomes a conceptual test – a tool of reflective practice – rather than (as Kant thought) a vehicle of justification, that is, a tool that would allow us routinely to claim morality for our actions. All it allows us is to subject the maxim or our action to an experiment, as it were, namely, by asking how that maxim might look in the eyes of others. More exactly: in the eyes of all others that might conceivably find themselves in the same situation or be concerned by the same kind of situation, here and there, now and in future. To speak with Mead: How would the premise of my action look in the eyes of the generalized other? The test, then, consists very much in the "universal role-taking" that we encountered in Mead's work.

Through universal role-taking, we can recognize a maxim's limited degree of ethical universality, and can thus qualify our claims accordingly. Remember what we said about Kant's non-idealistic intent: he does not assume that our individual premises can always be in harmony with those of all others. The problem of practical reason emerges from the fact that such harmony is the exception rather than the rule. It is because individual and collective premises do not usually converge that Kant asks us to consider what our personal premises are and to what extent they may be inconsistent with the premises of all others concerned. It is a simple, but imperative demand of reason that we do not claim more for our premises than what we can justify – therein consists the basic connection that Kant establishes between morality and rationality. Through Kant's work, morality becomes a matter of practical reason rather than individual values and virtues. On this "formalization" rests its claim to universality.

The universalization principle is universal – a principle that no reasonable person can deny – because what it demands in essence is that as rational agents (i.e., agents guided by reason) we do not claim too much! This also
means that we can claim no exception from what we expect from others: I cannot expect others to act according to principles that my own actions violate. I cannot claim it is all right for me to eat your cake and then complain when you eat mine, that is, claim it is wrong for you (but not for me) to do so. The imperative of not claiming too much thus urges us to orient our thinking away from our usual egocentric perspective towards the perspective of all others concerned, and in this way to review our premises accordingly. In the forceful words of Herbert Mead (1934, p. 381): "We cannot demand from others what we refuse to respect. It is a practical impossibility."

The "test" to which the categorical imperative amounts might then be stated by a simple question:

"Do I claim an exception for myself?"

Still, the difficulty remains: How can we ever assert with certainty that our premises of action amount to generalizable maxims, that is, do under no circumstances amount to claiming an exception for ourselves? The difficulty resides in the holistic implications of Kant's formulation of the universalization principle. As it is a purely formal principle, it has no way of qualifying and limiting the scope of situations and situation aspects to be taken into account; it thus amounts to an endless process of "sweeping in" ever more aspects of the real world, until we have considered the totality of all conceivable circumstances – or suffered mental breakdown (on the concept of "sweeping in," see Singer, 1957; Churchman, 1982; Ulrich, 1994).

I would argue that unless we are very careful, this decontextualizing thrust of the universalization principle (cf. Ulrich, 2006, p. 55) works against its usefulness. That we understand it merely as a negative test (a tool of reflective practice) rather than (with Kant) as a method of affirmative moral reasoning makes little difference in this respect. The principle's apparent disregard for the context of action cannot reasonably mean that as rational agents, we should or may disregard the concrete situations in which we try to act morally. Rather, it means we need to be careful about the way we formulate the maxim of action that we submit to the test of moral universalization! It is through the specification of the maxim that the relevant context is to be brought back into the picture, as it were. In fact, we cannot meaningfully "universalize" a maxim without making some assumptions
about the kind of contexts within which the universalized maxim is to hold unconditionally. In the terms of my work on critical systems heuristics (CSH) and critical pragmatism, the universalization principle (U) requires a counter-principle that reads: "Consider the context (C)!

The formulation of (U) then calls for a small but essential qualification:

"Act according to a principle that everyone can hold in the same situation!"

Accordingly, the "exception test" suggested above will read:

_Could I want all others in the same situation to make an exception?"

The point is, we cannot apply the universalization principle without somehow or other (whether consciously or not) delimiting the kind of situations to which the maxim in question is to be applied. The way we do this needs to be open and transparent, and may itself be subject to moral testing. Critical systems heuristics has proposed a tool to this end: _boundary critique_ – a systematic effort of unfolding the unavoidable selectivity and partiality of all our claims and actions (see, e.g., Ulrich, 2000 and 2006). That which is to be universalized is then not our maxims of action as such but rather, maxims of action considered with a view to certain contexts of action. I will return to this issue later on in this series of reflections, when we will look at some of the ideas we can draw from Habermas' work, as well as from my work on critical systems heuristics and critical pragmatism, all with a view to grounding reflective professional practice in practical philosophy. At this point, I would like briefly to consider an example Kant offers for applying the categorical imperative as he imagines it, but which in my view perfectly illustrates the danger involved in too absolute an understanding of the universalization principle, an understanding that ignores the need for delimiting the context of application. The example is the so-called case of the inquiring murderer.

**The Case of the Inquiring Murderer**

Source: Kant (1797, A302, with reference to B. Constant, 1797, p. 123)

Suppose you have allowed a person fleeing from a murderer to hide in your home. Then the murderer knocks at your door and asks you whether that person stays in your house. Should you tell him the truth or lie?

Kant (1797) asks this question in a little essay titled "On a supposed right to
lie from altruistic motives.” Does such an extraordinary situation, in which we face a conflict between truthfulness and helping someone in need of help, supersede the duty not to lie? Counter to what one might expect, Kant's answer is negative. There must be no such exception, he maintains, and thus uses the example to demonstrate the unconditional character of the prohibition of lying. If we lie, he argues, however well intentioned we may be we must answer for the consequences this may have, however unforeseeable they may be. Depending on how the circumstances in this case develop, it might be that the murder happens even though we lie, for instance, because the person we want to protect does not want to rely on our willingness to lie and therefore quickly leaves the house through the backdoor when he realizes the murderer is at the front door, and thus both find themselves face to face in the street. Whoever lies might justly be accused later on of having acted wrongly, and thus being accountable for the consequences. While this argument appears to beg the issue of moral conflict – why shouldn't we be just as responsible for not helping a person whose life is in danger as we are for not telling the truth? – Kant's underlying reasoning is of course that any exception we claim for ourselves has self-defeating consequences; for if that exception were adopted universally, we could never again be certain that others are telling us the truth. Even the act of lying would become meaningless; for its effectiveness, too, depends on the universal validity of the principle of not lying. If the prohibition of lying is not an absolute principle, we would soon no longer be able to rely on what others assert, and consequently would not even believe the lies others use for altruistic reasons. This reveals for Kant how self-defeating any exception to the principle may be.

Is such an absolute understanding of the universalization principle (or of the moral norms it is to support) really sound? I do not think so. Ethical problems often arise through a conflict among different moral goods, such as in the inquiring murderer case, and in such situations of conflict it must be allowed to balance the conflicting goods carefully with a view to minimizing the harm that each option of acting might cause. I cannot see why such careful reasoning should violate the universalization principle; but careful reasoning certainly demands that we thoroughly consider the situation of conflict before we choose a way to respond to it. I would argue that what is at issue is not the universalization principle as such but rather, the maxim (or rule of action) to which we apply it. Whether a maxim withstands the test of
the categorical imperative or not depends entirely on how we formulate the maxim. The decisive issue is the way our maxim specifies the kind of situation in which we expect all others to act in the same way! Remember that earlier we reformulated the basic form of the categorical imperative so that it asks us to act according to a principle that everyone can hold \textit{in the same situation}. Applied to the inquiring murderer case, the Kantian exception test then reads:

"Could I want all others who find themselves \textit{in the same situation} to make an exception from the prohibition of lying?"

Thus reformulated, my answer is clear, and different from Kant's: yes. Acting according to a thus-qualified maxim does not run counter to the universalization principle but only to its unreflecting, decontextualized, employment: we cannot meaningfully grasp and assess human practice free of any contextual assumptions and limitations.

As soon as we specify the relevant context, Kant's warning about the self-defeating character of any exception constitutes no longer a compelling argument against helping. The crucial point is no longer whether we could want everyone to feel free to lie – of course not – but rather, how we delimit the type of situation for which we claim an exception. We can will without any self-defeating implications that all others who find themselves \textit{in the same situation}, facing a conflict between telling the truth and saving another person's life, should give preference to saving the endangered life. I, you, all of us might be the person needing help. The only premise I can want to hold in such a situation \textit{consistently} (i.e., without claiming an exception for myself when \textit{my} life is in danger) is that when anyone is threatened by a murderer and another person can help by misleading the murderer, it is adequate for that other person to do so. It is always morally right to help save another person's life if one is able to do.

In conclusion, as soon as we properly frame the situation to which the maxim in question applies, the universalization principle does indeed furnish an adequate standard. Unless we care about the way we frame specific contexts of actions, it makes little sense to apply the principle of moral universalization. Moral universalization brought down to earth means caring about situations, according to the motto: \textit{that which may concern each of us should concern us all}. 

\text{http://geocities.com/csh_home/bimonthly_march2009.html}
Kant, then, did indeed formulate a truly universal standard of moral reasoning – the only truly unconditional principle of practical reason there is – although he may have underestimated the difficulties of application. The difficulties arise because, as Kant was the first to recognize, such a principle is necessarily merely formal and thus empty of content. Would it consider empirical circumstances, its validity would forever be subject to the changing nature of the circumstances of action. But in his inimitable determination to push things philosophically to their limits, Kant may have gone one step too far. While it is correct to say that a generic moral principle must hold independently of specific circumstances, it is not correct (because impossible) that we should apply it without considering the specific circumstances. Abstracting from subjective premises – dropping the ego – is not the same as altogether ignoring an action's contextual circumstances and consequences.

Counter to what is often asserted about Kantian ethics, I do not believe moral universalization precludes that we care about the consequences of our actions. It only precludes that we claim unqualified exceptions for ourselves, exceptions that we would not grant all others facing the same kind of situation. We must not violate the rules or principles that we expect others to respect. We cannot expect others to bear consequences we are not prepared to bear ourselves. It is, as Mead (1934, p. 381) said, a practical impossibility; not in the sense that we might not act against the universalization principle – moral action depends on free decision, yet freedom is for evil as for good – but rather, as we have learned from Kant, that to the extent we do act against it, we are bound to reason inconsistently and our argumentative position will be accordingly weak.3 There exists a deep, inextricable link between ethically tenable action and consistent reasoning or argumentation. This is the great lesson that Kant's practical philosophy can teach us all – a lesson that certainly is as pertinent today as it has ever been.

Notes
1) The notion of "duty" makes many people think of the feudal Prussian society of which Kant was a part, and perhaps also of the spirit of Protestant Puritanism (if not Stoic attitudes) with which his parents may have educated him; such associations then lead people to doubt the universal validity of the Kantian principle of moral universalization. However, I do not think we need to follow this widespread tendency of overinterpreting Kant's legalistic language of moral "law" and "duty." Methodologically speaking, its point is to highlight the unique force that moral considerations acquire when they take as their standard the universalization principle; for what has universal validity cannot easily be ignored or put aside by reference to particular circumstances. [BACK]
2) I should mention that there is a second, better known sense in which the universalization principle represents a categorical imperative, namely, as distinguished from hypothetical imperatives, that is, rules of action which only hold inasmuch as some subjective purpose is presupposed (1786b, B39f). Hypothetical imperatives aim at action that is good as a means to an end, with the ultimate end being personal happiness; the categorical imperative aims at action that is good in itself or unconditionally good, and the ultimate end is doing one's duty and in this sense being worthy of happiness (a notion that is bound to recall the Stoics' and Aristotle's *eudaimonia*, although Kant now links the idea of being worthy of happiness with the idea of duty as captured in the universalization principle). [BACK]

3) As an example, I cannot help but think of the current controversy around all those "tax havens" around the globe which facilitate and protect tax evasion and tax fraud, by allowing their financial institutions to have a competitive advantage in the world-wide market for financial consulting and asset management based on an ethically and legally problematic use of the so-called "banking secret." This business model violates the principle of moral universalization, as the countries that serve as tax heavens for the citizens of other countries cannot reasonably want those other countries to serve as tax heavens for their citizens. Supporting tax evasion clearly is not a universalizable business model, much less a universalizable national policy. The recent controversies have made it only too obvious that the argumentative position of the tax havens is accordingly weak; it cannot surprise us, therefore, that their policies to defend the status quo have been increasingly inconsistent and confuse, with the result that this immoral business model is now about to collapse, despite the strong economic interests behind it. [BACK]

References


---

**March 2009**

<table>
<thead>
<tr>
<th>Su</th>
<th>Mo</th>
<th>Tu</th>
<th>We</th>
<th>Th</th>
<th>Fr</th>
<th>Sa</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
</tr>
<tr>
<td>29</td>
<td>30</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**April 2009**

<table>
<thead>
<tr>
<th>Su</th>
<th>Mo</th>
<th>Tu</th>
<th>We</th>
<th>Th</th>
<th>Fr</th>
<th>Sa</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Picture data** Digital photograph taken on 24 June 2008, around 10:30 a.m., near Gruyères, Switzerland. ISO 100, exposure mode manual, exposure time 1/30 seconds, aperture f/5.6, exposure bias -0.30, focal length 14 mm (equivalent to 28 mm with a conventional 35 mm camera). Original resolution 3648 x 2736 pixels; current resolution 700 x 525 pixels, compressed to 120 KB.

---

"Dropping the ego"
or how to act according to Kant's principle of moral universalization

---

*The spirit of Kantian ethics: ‘Drop the ego!’*

„Act according to a principle that can hold generally!”


26.03.2009
“Wherever the element of the 'ought' comes in … it always takes on this universal form. … We cannot demand from others what we refuse to respect. It is a practical impossibility.”

(Immanuel Kant, *Groundwork of the Metaphysic of Morals*, 1786b, B52 and B81, simplified transl.)

(Herbert Mead, *Mind, Self, and Society*, 1934, p. 380f)

Write down your thoughts before you forget them!
Just be sure to copy them elsewhere before leaving this page.

Last updated 26 Mar 2009 (first published 19 Mar 2009)