

Published in Andrews Reath, Barbara Herman and Christine Korsgaard, Eds., *Reclaiming the History of Ethics: Essays for John Rawls* (New York: Cambridge University Press, 1997), available through <http://www.cambridge.org/uk/catalogue/catalogue.asp?isbn=0521472407>

Adrian M. S. Piper

Kant on the Objectivity of the Moral Law (1994)*

In 1951 John Rawls expressed these convictions about the fundamental issues in metaethics:

[T]he objectivity or the subjectivity of moral knowledge turns, not on the question whether ideal value entities exist or whether moral judgments are caused by emotions or whether there is a variety of moral codes the world over, but simply on the question: does there exist a reasonable method for validating and invalidating given or proposed moral rules and those decisions made on the basis of them? For to say of scientific knowledge that it is objective is to say that the propositions expressed therein may be evidenced to be true by a reasonable and reliable method, that is, by the rules and procedures of what we may call "inductive logic"; and, similarly, to establish the objectivity of moral rules, and the decisions based upon them, we must exhibit the decision procedure, which can be shown to be both reasonable and reliable, at least in some cases, for deciding between moral rules and lines of conduct consequent to them.¹

In this passage Rawls reconfigured the issue of moral objectivity and so reoriented the practice of metaethics from linguistic analysis to rational methodology. In so doing, his work has provided inspiration to philosophers as disparate in normative views as Thomas Nagel,² Richard Brandt³, Alan Gewirth⁴, and David Gauthier.⁵ Rawls replaced the Moorean question, Do moral terms refer? with the Rawlsian question, Can moral judgments be the outcome of a rational and reliable procedure? He later gave a resoundingly positive answer to this question⁶ and later still, a more tentative one.⁷ Rawls' considered qualification of his earlier enthusiasm about the extent to which moral philosophy could be "part of the theory of rational choice"⁸ is a tribute to the seriousness with which he took his critics' objections.

* © APRA Foundation Berlin

The above passage, and the article from which it is excerpted, make clear that Rawls took his original inspiration from a carefully worked out analogy with inductive logic in scientific procedure. But a fellow traveler among rational methodologists in metaethics – namely Kant – risked defending an even closer and more controversial relationship between scientific and moral objectivity. I shall argue that Kant's thesis that the moral law is objectively necessary⁹ relies on the same type and degree of objectivity he earlier claimed for scientific knowledge. Thus Rawls' youthful impatience with "speculative" metaethics, and the boldness of his claiming for ethics the same sort of procedural rigor to be found in the natural sciences, puts him in the best possible philosophical company: of those whose ambitions for moral philosophy – and the philosophical powers by which they serve it – are greatest.

In the *Groundwork* Kant characterizes the moral law as an objective principle that compels imperfectly rational human beings with objective necessity (Ak. 412-413). Kant establishes the metaethical foundations and technical terminology for his conception of moral objectivity in the Analytic and Dialectic of the *Critique of Pure Reason*.¹⁰ Although many of Kant's views undergo revision or development from the first *Critique* to the *Groundwork* to the second *Critique*,¹¹ we will see that the conceptual foundations Kant establishes early on for addressing the issue of moral objectivity remain firmly in place.

I. Understanding

A. Synthesis

A *representation*, for Kant, is any mental content. Representations can be either intuitional or nonintuitional. *Intuitional* representations "get directly to"¹² an object that is given to us in sensibility (A 19/B 34). An intuitional representation organizes the data of sense in space and time, which Kant calls the *forms of intuition*. So only intuitional representations are directly of given objects. All nonintuitional representations are themselves of representations rather than directly of objects (A 68/B 93). Since intuitional representations are by definition of objects and not of other representations, and since empirical objects are themselves representations, the objects intuitional representations represent cannot be, in turn, empirical objects. Rather, they are objects in themselves apart from their representations.¹³

A *concept* "orders various representations under one common representation" (A 68/B 93). We do this *spontaneously* in that this mental act is not a reaction to some

external cause, as sensation is (A 50/B 74, *passim*). Conceptualizing representations is something we initiate rather than something that is imposed upon us. This, Kant thinks, is what it means to be an intelligence (B 158, n.). And to conceive ourselves in this way as active, spontaneously reasoning and thinking agents is to conceive ourselves as *persons*.¹⁴

The representations we conceptualize can be either intuitional or nonintuitional. They have in common what is expressed by the representation that subsumes them. The common representation is the concept that unifies the various representations under it. A concept, then, is a rule of selection for collecting mental contents similar in a certain respect under the rubric of that concept; or, as Kant also puts it, a *function* (A 79/B 105). "Since," Kant tells us, "no representation other than intuition gets directly to the object, a concept is never in unmediated relation to an object, but rather to some other representation of it (be it an intuition, or itself a concept)" (A 68/B 93). Concepts, therefore, mediate and qualify our relation to externally given objects.

Synthesis is Kant's technical term for the process by which different, specifically intuitional representations are collected under one concept. Synthesis is collection when applied specifically to intuitional representations. It is, Kant tells us, "the act of adding different representations together, and of grasping their manifoldness in one cognition" (A 77/B 103). Synthesis supplies order and continuity to our moment-to-moment sense experience (A 99-102). It also unites intuitional representations into a particular, identifiable content (A 78/B 103). This content is what the concept that unites them expresses. So intuitional representations are synthesized according to a particular rule of selection, and this rule can be inferred from the content of the concept under which they are subsumed.

What determines the content of that rule of selection? That is, what determines which similarities among intuitional representations are relevant to their synthetic grouping under concepts? Kant uses the term "synthesis" specifically in connection with intuitional representations, and intuitional representations are direct and unmediated representations of externally given objects in themselves. So it would be tempting to think that the similarities represented were similarly given by those objects; and therefore that the basis for our grouping of representations was to be traced to attributes of those objects themselves. But Kant thinks that the similarities most salient to us, and therefore the best candidates for conceptual synthesis, are those which most

closely conform to the innate conceptual preconceptions we bring to the act of cognitive discrimination.

The concepts that select and group intuitional representations are what Kant calls the *pure transcendental concepts of the understanding, or categories*. These include substance and attribute, and cause and effect. These pairs are not *analytic*, i.e. conjoined by definition – as are the concepts of a bachelor and an unmarried man. Rather, they are conjoined in the spatiotemporal form of our intuitional representations. And since, according to Kant, they are innate, logically necessary¹⁵ and universal concepts, they are cognitive preconditions for experiencing something as a coherent and unified object at all.

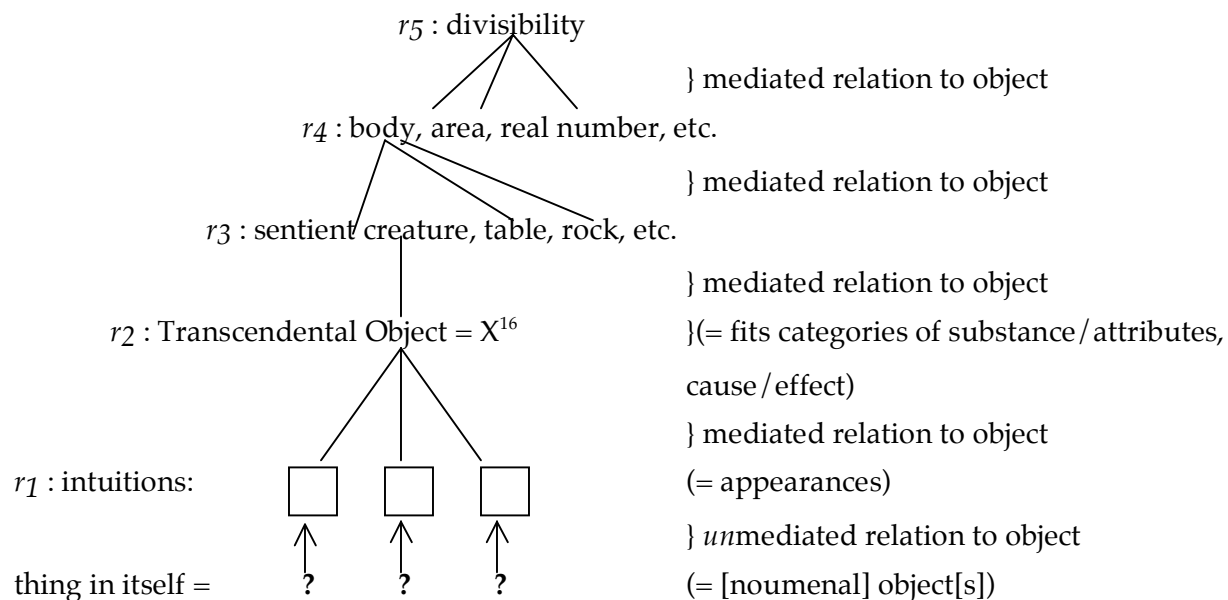
B. Judgment

Kant thinks that "the understanding can make no other use of these concepts other than to judge by means of them" (A 68/B 93). A *judgment* – more specifically, a categorical judgment that ascribes predicate to subject – does exactly what concepts do, only at a more abstract level: it collects relevantly similar representations that are already subsumed under a less abstract concept under a more abstract concept (A 79/B 105). A *synthetic a priori judgment*, then, is an innate, logically necessary and universal judgment that collects intuitional representations under the pure concepts of the understanding listed in the Table of Categories (A 80/B 106).

Kant tells us that by abstracting from the content of any such judgment and "attending only to the mere form of [our] understanding" (A 70/B 95), we can cull the formal rules of selection among representations. These are the *logical forms of judgment* set out by Kant in their entirety in the Table of Judgment (A 70/ B 95). So whereas the categories of the understanding listed in the Table of Categories combine the innate, logically necessary and universal rules for collecting representations with specifically intuitional representations, the logical forms of judgment listed in the Table of Judgments abstract from both intuitional and nonintuitional representations. Subtract representations themselves from the conceptual rules for collecting them and you get the logical forms of those rules.

Among the representations subsumed by a judgment under any concept, whether transcendental or empirical, there will be at least one which is intuitional, i.e. in

direct relation to a given object." [A]ll judgments," Kant says, "are functions of unity among our representations, since instead of an unmediated representation, a higher one, which comprises this and more, are used in knowing the object, and thereby many possible cognitions are gathered into one" (A 69/B 94). Thus the underlying structure of the categorical judgment, "All bodies are divisible," might look something like this:



In this judgment, relevantly similar intuitional representations of (a) given noumenal object(s) r_1 are synthesized under the transcendental concept of an object r_2 , to which certain categories apply, such as substance/accident and cause/effect. These situate the appearance of the object in space and time. They also define what it is for something to be an object of our experience. We must represent it as a discrete substance with attributes. We must also represent it as both having and being susceptible to causal force. Once we become aware of something as satisfying the cognitive criteria for being an object, we are in the realm of empirical objects of experience. Here we next try to ascertain what kind of object it is. We collect observational representations under higher-level empirical attributes. Representations of the relevantly similar empirical attributes of that object are in turn collected under a more abstract empirical concept r_3 , that of a sentient creature; that representation, in turn, along with representations of other relevantly similar three-dimensional entities under a yet more abstract concept r_4 , that of a body; and that representation, finally,

along with representations of other similarly individuated but non-three dimensional entities under the even more abstract concept *r5*, divisibility.

So on the one hand, all higher concepts or representations implicitly embed externally given objects in the judgments we make, no matter how tenuous or distanced our conceptual connection to those objects. On the other, any nonintuitional representation distances us from the object to some extent, including the representations "I", "I think," or "is my experience." All of our experience is theory-laden in this way. This is what Kant means when he says that "[j]udgment is ... the mediate knowledge of an object, that is, the representation of a representation of it" (A 68/B 93). Any cognition that I can recognize and represent as mine is thereby distanced from sensibility and mediated by that judgment itself. On Kant's view, being "detached" from my "feelings" is the necessary price of unified selfhood.

C. Categorical Indicatives

The two logical forms of judgment of interest for this discussion are the first two under "III. Relation," in the Table of Judgments of the first *Critique*, the categorical and the hypothetical. *Relational judgments* generally connect a priori concepts that have already synthesized intuitional representations. *Categorical judgments* – more specifically, categorical *indicative* judgments – have the form, "All (or some) A is B." Categorical indicatives relate the predicate concept to the subject concept by way of the representations subsumed under each. An example would be the judgment, "The human soul is immortal." A second example would be "All bodies are divisible." Both examples relate predicate to subject, but only the second relates attributes to substances because the first is true by definition of the concept of a soul, whereas the second connects intuitional (as well as nonintuitional) representations of objects.

Hypothetical judgments have the form, "If P then Q." Hypothetical judgments relate two categorical indicative judgments as antecedent (or, to use Kant's technical term, *ground*) to consequent. For example, the hypothetical judgment, "If there is a perfect justice, the obstinately wicked are punished," relates the two categorical indicatives, "There is a perfect justice," and "The obstinately wicked are punished" (A 73/B 98). Another example of a hypothetical judgment would be, "If one answers the telephone, it ceases ringing." This one relates the two categorical indicatives, "One

answers the telephone,"¹⁷ and "The telephone ceases ringing." Both examples relate antecedent to consequent, but only the second relates cause to effect, and again the reason is the same. The first is true by definition of the concept of perfect justice, whereas the second connects intuitional representations of objects.

At B 141 fn. and A 73/B 98 - A74/B 99 Kant inveighs against the reduction of hypothetical and disjunctive judgment forms to the categorical.¹⁸ Nevertheless he is wrong. Disjunctive judgments can be translated into hypothetical ones, and any hypothetical judgment can be replaced by a categorical one that ascribes to the subject a complex property, such as that of being an antecedent to the consequent or a cause of the effect. For example, the above two hypothetical judgments can instead be rendered as "Perfect justice requires punishing the wicked," and "Answering the telephone stops its ringing." By reducing the number of judgment forms Kant claims to be transcendently necessary to the categorical and extending its scope to cover antecedental and causal predicates, Kant's argument for the objectivity of empirical knowledge is strengthened, not weakened as he seems to think. The reason is that categorical indicatives have a special status in Kant's theory; and in the B Deduction Kant explains what it is:

[A categorical indicative] judgment is nothing other than the way to bring given cognitions to the objective unity of apperception. This is what the relational term "is" aims at: to distinguish the objective unity of given representations from the subjective. It denotes their relation to original apperception, and its necessary unity, even if the judgment itself is empirical, and therefore contingent, as for example, in the judgment, "Bodies are heavy." I .. here assert that these representations ... belong to one another in virtue of the necessary unity of apperception in the synthesis of intuitions Only thus does there arise from this relation a judgment, that is, a relation which is objectively valid, and so can be adequately distinguished from a relation of the same representations that would have only subjective validity, for example according to laws of association. In the latter case, I would only be able to say, "When I carry a body, I feel an impression of weight"; but not, "It, the body, is heavy;" which is as much as to say that these two representations are bound together in the object, i.e. without regard to the state of the subject, and not merely in perception (as often as that perception may be repeated) (B 142).

Kant's point is that the relational term that connects predicate with subject in a categorical indicative judgment – and it is a point that would apply as well to "has the attribute of" as it does to "is" – *objectifies* that connection as being in the object rather than in the subject. To recognize the connection of representations as being in the object is to recognize the representations that constitute that object as epistemically distinct from those that constitute the subject.

Now the logical function of judgment, Kant tells us, is to collect all such representations, whether intuitional or conceptual, in one consciousness, or *apperception*; that is, to finally subsume all of them under the concept "I think" (B 143). This means that necessarily, I must be able to conceive of any such representation, concept or judgment, no matter how abstract or removed from sensibility, as mine in order for it to be part of my consciousness. But the representations that collectively comprise the representation of a unified object presuppose a unified conscious subject – *the transcendental unity of apperception* – in which those representations are ordered and collected according to innate, logically necessary, and universal concepts. Therefore, a categorical indicative asserts a connection among representations that is as objective as anything can be for us.

Kant's doctrine of transcendental idealism clearly rejects any concept of empirical objectivity as consisting in the ontological independence of the object. And the Analogies demonstrate Kant's attempts to ground the objectivity of objects and causally related events in an analysis of temporally sequential representations. Objectivity, for Kant, is objectivity of knowledge, not the ontological independence of the object from the knowing subject. On Kant's view, objectivity just is the conceptual a priority, logical necessity, and universality *of judgment about* the object.¹⁹

Even if we concede Kant his insistence on the essential difference between categorical, hypothetical and disjunctive judgments, however, hypothetical and disjunctive judgments both still embed categorical indicatives that implicitly assert objective relations between subject and predicate. Therefore, all such relational judgments implicitly presuppose the objectivity of these relations. Whether the telephone ceases ringing because I answer it or not may be open to question; that I answer it, and that it ceases ringing, is not.

But all such empirical hypothetical judgments presuppose the transcendental forms of judgment that unify *a priori* intuitional representations. Kant describes all of

the synthetic *a priori* judgments that comprise the Table of Categories as universal rules of unity in the synthesis of appearances that have objective reality as necessary conditions of experience (A 157/B 196). What he means is that the relational judgment forms (inter alia) we find in the Table of Judgments, when applied to the synthesis of intuitional representations, further objectify those representations by linking them in the form of discriminable events and entities. These empirical events and entities are, from the point of view of my experience, fully objective and veridical. *A fortiori*, the hypothetical judgment form synthesizes intuitional representations already objectified by their categorical indicative form into a causal relation that, from the point of view of experience, objectifies them even further.²⁰

II. Reason

A. Ideas

In the Analytic of the first *Critique*, Kant has tried to show that the logical forms of relational judgment listed in the Table of Judgments, when combined with intuitional representations, yield objective knowledge. In the Dialectic he tries to show that these same three logical functions, when extended beyond intuitional representations, yield increasingly abstract, comprehensive concepts and theories that encompass all lower-level representations; and finally yield transcendent and transcendental *ideas of reason*. Kant defines an idea or concept of reason as "a concept formed from [pure concepts of the understanding] that transcends the possibility of experience" (A 320/B 377). Ideas are highly abstract concepts that unify less abstract concepts of a certain kind under a single, comprehensive concept. An idea can be a theoretical entity (such as a perfectly rational being or a quark), a theoretical law (such as that of freedom or relativity) or a concept strictly speaking (such as immortality or the unified field). Ideas deal only indirectly with experience (B 359), since "an object that would be adequate to the transcendental idea can never be found within experience" (A 327/B 384).

However, these most sweepingly universal concepts do not abstract from intuitional representations, but instead embed them in more abstract and general judgments. Thus when Kant characterizes reason as the faculty of principles, and knowledge from principles as "that knowledge in which I cognize the particular in the universal through concepts" (A 299/B 356-A 300/B 357), he is making the same point about the theory-ladenness of experience he made at A 68/B 93-A 69/B 94 discussed in

Section I.B, but at a higher level of abstraction. Reason, he says, is the faculty of judging mediately (A 330/B 386).

There are many ideas of reason.²¹ But Kant claims that the three most abstract, comprehensive and universal ones are engendered by the three relational judgment forms themselves. We have seen in Section I.A that when these judgment forms apply to the intuitional data of sense they yield the categories. But when their scope of application is extended sufficiently far beyond the data of sense they yield the higher-level concepts Kant calls ideas (A 323/B 380). First, there is the idea of immortality, which he argues is engendered by universalizing the categorical judgment form: The search for completeness and comprehensiveness of predication yields the concept of a subject which is never a predicate. Since ideas of reason must unify intuitional as well as nonintuitional representations and lower-level principles, this, in turn, according to Kant, becomes the idea of "the absolute unity of the thinking subject" (A 335/B 392), which itself undergoes no change.²²

Second, there is the idea of freedom, which is engendered by universalizing the hypothetical judgment form: In hypothetical reasoning, the search for completeness and comprehensiveness in the series of antecedents of a consequent yields the concept of the original antecedent without antecedents – the groundless ground. Because it, too, must unify intuitional as well as nonintuitional representations, it, in turn, yields the idea of freedom as the first or uncaused cause of everything, the "absolute totality of the series of conditions for any given appearance" (A 340/B 398). Of course a judgment characterizing a groundless ground or first cause could also be rendered as a categorical indicative that predicates the consequent as a property of the antecedent *qua* subject that cannot be predicated of any further subject. A judgment ascribing an action to a free agent would, for example, satisfy this criterion.

Finally, there is the idea of God, which is engendered by universalizing the disjunctive judgment form: Here the search for completeness and comprehensiveness in the set of disjuncts that jointly exhaust the scope of a concept (A 73/B 99) yields the concept of the aggregate of all disjuncts of the most complete concept; and this becomes the idea of God as the totality of all parts of the system, the "being of all beings" (A 336/B 393).

In the Dialectic Kant tries to show how all of these ideas of reason naturally develop as higher-level universal principles from the conceptual presuppositions or

Grund of objective empirical knowledge, namely the logically necessary functions of thought. He characterizes these ideas, and the reasoning by which we arrive at them, in the same indicative mood as he did the concepts of the understanding and the synthetic process of judgment by which we unify and objectify our experience. Kant argues that, being committed to the objectivity of our empirical experience, we then must be committed to the regulative authority of the abstract theories and universal principles it engenders. Since they transcend the empirically verifiable, we cannot experientially confirm their truth. But since they naturally arise out of it, we cannot easily reject them either.

B. *Vernunftschlüsse*²³

Kant's account of the way in which ideas of reason are engendered from objective knowledge is based on his conception of cognitive functions as fundamentally spontaneous – i.e. active rather than reactive – and synthetic – that is, unifying rather than particularizing. We saw in Section I.A that concepts collect representations, and in Section I.B that relational judgments collect concepts. Similarly, syllogisms, or inferences of reason – *Vernunftschlüsse* – collect relational judgments (A 301/B 358). Therefore inferences of reason indirectly collect representations and concepts.²⁴ Whereas the understanding collects intuitional representations under lawlike concepts, reason collects what we've conceptualized under yet more abstract and universal explanatory principles that organize and unify them in relation to each other (A 302/B 359). Kant thinks we are so constituted cognitively as to strive naturally to reduce, simplify, and extend the explanatory scope of theoretical principles to cover the greatest variety of intuitional and nonintuitional representations. Reason embeds the particulars of experience in universal judgments that define the broadest possible theoretical framework because to universalize over the particulars of experience is a natural and necessary expression of the self's striving for rational coherence.²⁵

According to Kant, the forms a *Vernunftschluß* can take depend on the logical form of judgment employed. Again the categorical and hypothetical forms are most important for our purposes, and the categorical form is the foundation from which the others are constructed:

	categorical (indicative):	hypothetical (indicative):
Major Premise:	All A are B.	If D is E then F is G.
<u>Minor Premise:</u>	<u>All/Some C are A.</u>	<u>D is E.</u>
Conclusion:	∴ All/Some C are B.	∴ F is G.

In each case the major premise is a given judgment stating a universal rule that ascribes a predicate to all subjects of a certain kind (A 322/ B 379). The minor premise is a judgment that subsumes a second subject, or condition, under the concept of the first. And the conclusion yields a further judgment by applying the universal rule of the major premise to the subsumed condition of the minor premise. A *condition*, for Kant, is a certain kind of state of affairs. In a categorical indicative judgment that functions as a premise in a categorical *Vernunftschluß*, it is the referent of the subject of the antecedent in the major premise (and, of course, of the subject of the minor premise).²⁶ Thus a condition functions as does the premise whose subject refers to it, as an antecedent from which some consequent or conclusion can be inferred (this is why Kant sometimes uses the terms “condition” and “premise” interchangeably). The conclusion then describes the *conditioned* (A 330/B 386 - A331/B 387), and the major premise *conditions* the conclusion: that is, it explains the condition or subject described in the conclusion by subsuming it under a general rule. Reason seeks to subsume that kind of subject under more and more abstract and universal kinds in a series of increasingly general judgments.

Take the categorical indicative *Vernunftschluß*,

(1)	Major Premise:	All humans are mortal.
	<u>Minor Premise:</u>	<u>Caius is human.</u>
	Conclusion:	∴ Caius is mortal.

The major premise of (1) may be in turn the conclusion of even more general and abstract ones, such as

(2)	Major Premise:	All sentient creatures are mortal.
	<u>Minor Premise:</u>	<u>All humans are sentient creatures.</u>
	Conclusion:	∴ All humans are mortal.

And the minor premise of (1) also may be conditioned by prior *Vernunftschlüsse*, such as

- (3) Major Premise: All featherless bipeds are human.
 Minor Premise: Caius is a featherless biped.
 Conclusion: ∴ Caius is human.

In both cases reason seeks more general and abstract principles under which a judgment can be subsumed, and from which it can be derived. For any *Vernunftschluß*, both of its premises can be regarded as conditioned by the more general *Vernunftschluß* from which they are derived as conclusions.

The principle of reason in general, according to Kant, is "to find for the conditioned knowledge of the understanding the unconditioned whereby its unity is completed" (A 307/B 364). This then becomes a regulative principle of pure reason on the presupposition that "if the conditioned is given, the whole series of conditions subordinated to one another, which itself is therefore unconditioned – is also given ...". A *regulative principle* is one that guides and directs our innate patterns of reasoning in a way that enables us to extend it past the empirical limits of experience (A 509/B 537). It contrasts with a *constitutive* principle that defines and determines objectively the existence of some object or state of affairs. The regulative principle of pure reason, then, leads us to seek that most abstract, universal and all-inclusive first principle or *Grund* from which everything else can be deduced.²⁷

Thus in order to generalize over such principles to increasing degrees of abstraction, Kant argues, we must assume a totality of such conditions (or premises) for any given conditioned (or subject to which those predicates are ascribed in the conclusion). The complete series of conditions contained in prior universal judgments that determine a conclusion about some object or state of affairs is what Kant calls *the unconditioned* (A 322/B 379). The unconditioned, then, is the idea of a first and most abstract, universal idea or descriptive principle that functions as a premise subsuming all lower-level descriptive principles or premises under itself. This idea of the unconditioned leads us to try to generate a series of *Vernunftschlüsse* whose major premises increase in generality and comprehensiveness relative to the particular facts

(or "empirical conditions") with which it begins – what Kant describes as the *ascending* or *regressive* series (A 331/B 388).

The transcendental ideas of God, freedom and immortality are unconditioned in that they express the most abstract and universal principles under which all lower-level disjunctive, hypothetical and categorical *Vernunftschlüsse* respectively must finally be subsumed (A 336/B 393 - A 337/B 395). According to Kant, we know that these are the most comprehensive and universal principles there are because, as we have already seen above, they are derived by universalizing the forms of principles in which reasoning itself occurs. Beyond these most fundamental explanatory principles, Kant thinks, there are no further principles that might subsume them. Just as experience defines the limits of understanding, the ideas of God, freedom and immortality define the limits of reason.

C. Hypothesis Construction

According to Kant, we embark on the ascending series of *Vernunftschlüsse* in our search for eventual theoretical completeness by formulating and testing general hypotheses that are intended to explain an increasing variety of particular events. This process of theory-building, the prototype of Hempel's covering-law theory of scientific explanation,²⁸ Kant describes as the *hypothetical employment* of reason:

If reason is a faculty of deducing the particular from the universal, then either the universal is already certain in itself and given, and so requires only judgment for subsumption, and the particular is thereby necessarily determined; or else the universal is given only problematically, and is a mere idea. Here the particular is certain, but the universality of the rule from which it follows is still a problem. Then more particular instances, each of which are certain, try on the rule for size to see whether they follow from it. In this case, if it appears that all particular instances assigned do follow from it, then we conclude to the universality of the rule, and thence to all instances, even those not themselves given (A 646/B 674).

The procedure is clear. We begin with a universalization we rationally suppose to be true, and a firmly established particular case we rationally suppose might instantiate it. We scrutinize several relevantly similar particular cases in order to ascertain whether they do, indeed, instantiate the universalization, such that the formulation of the universalization allows the derivation of the particular instances from it: Does

universalizing over the particular instances yield the universal from which they are supposed to be derived? Does the resulting rule apply not only to these cases, but also to all relevantly similar ones? If the relevant instances, and other citable ones, can be derived from the universalization, we are then justified in regarding it as a valid universal principle that will predict and subsume future instances of that kind.²⁹

The example Kant offers combines the ideas of God, freedom, and immortality in the idea of rational personhood as a substance with causal power. This is a particularly apt example in light of Kant's project of establishing metaethical foundations in the first *Critique* for the substantive moral theory he first invokes to illustrate moral reasoning in the *Groundwork*. There he argues that "because moral laws should be valid for every rational being as such, to derive them from the universal concept of a rational being in general is thereby to explicate all moral philosophy, which needs anthropology for its application to human beings, first independently from this as pure philosophy, i.e. entirely as metaphysics (which can very well be done in this kind of wholly abstract knowledge)" (Ak. 412). The concept of a rational being in general is thus both the highest-level idea of reason that unifies the ideas of God, freedom and immortality; and also the most fundamental concept of morality.

Beginning with various established mental dispositions – "sensation, consciousness, memory, wit, power of discrimination, inclination (*Lust*), desire, etc." (A 649/B 677) – Kant directs us to employ the "logical maxim through which we reduce, so far as possible, this seeming diversity, so that through comparison one might discover hidden identity." By sorting them into nonredundant groups and universalizing over them we collect them under higher-level principles such as imagination plus consciousness; then under the more general headings of understanding and reason; then under the yet more general explanatory principle of a fundamental power, i.e. a substance with causality, which each such disposition instantiates. Notice that this is one example of the reduction of a hypothetical judgment to a categorical one, i.e. through the ascription of causal properties to substances. But this is a very particular kind of causal substance, namely one that has mental powers with causal efficacy.

Finally, Kant says, we may subsume all such causally powerful substances under the principle of an absolutely fundamental power that underlies all relative ones (A 649/B 677). But he cautions us, "This unity of reason is merely hypothetical. It is not claimed that such³⁰ must in fact be encountered, but that we seek it for the benefit of

reason, that is, of establishing certain principles for the many rules which experience may offer us" (A 649/B 676 - A 650/B 678; see also A 682/B 710). The concept of the soul as a self-identical and unchanging entity with the causal powers of understanding and reason can be invoked to explain the diverse manifestations of consciousness. And the concept of God can be invoked to explain the existence of such causally powerful substances. But these are only *concepts* (or ideas), not empirical experiences. Therefore they are not themselves susceptible to empirical confirmation. They are merely regulative ideas of reason that unify the diversity of our particular experiences.

On the one hand, this employment of reason itself remains hypothetical because the ideas of God, freedom and immortality are merely hypotheses that explain our experience rather than statements of fact about our experience. As Kant later remarks about the idea of immortality, "one posits (*sich setzen*) an idea merely as the one and only *point of view* from which one can extend that unity which is so essential to reason and so beneficial to the understanding" (A 681/B 709; italics added). To say, however, that ideas of reason are regulative hypotheses is not thereby to underestimate their foundational necessity in human thought. Kant makes it quite clear that the validity of any such regulative hypothesis turns on its ability to unify our experience, our thought, and finally our selves in accordance with the demands of logical necessity.³¹ So to the formula for Kant's account of objectivity earlier adduced – conceptual a priority, logical necessity, and universality of judgment, reason adds something more, namely theoretical coherence. A higher-level theoretical principle of reason is objectively valid if it subsumes its lower-level principles, concepts, and representations, both intuitional and nonintuitional, under it, such that it allows the systematic inference of those lower-level principles, concepts, and representations as logically necessary syllogistic conclusions. Only then can the self be fully unified and the rules of understanding be true.

III. Action

In this section I apply the conclusions of the preceding account of Kant's models of understanding and reason to the special case of action, following Kant's claim in the Preface to the *Groundwork* that

the unity of practical with speculative reason simultaneously in a common principle must be able to be delineated, since *in the end there can be only one and the*

same reason, which must be differentiated solely in its application (Ak. 391; italics added).

A. Maxims

Earlier, in Section I.C, we saw that intuitional representations are synthesized by the categories of the understanding first and foremost into categorical indicatives that claim a certain degree of objectivity in virtue of their form. We also saw that these are then further synthesized into hypothetical indicatives, that is, descriptive causal judgments that are, from the perspective of one's own experience, even more objective. The objectivity of this causal relation holds as much for human actions and consequences as it does for other causally linked events. Action and passion, according to Kant, are *pure derivative concepts*, or *predicables* of the pure understanding, derived from the category of cause and effect (A 82/B 108). Since action and passion are transcendental concepts, they, too, synthesize intuitional representations into objects and events.³² This means that as agents and patients of action, we experience ourselves as genuine empirical causes and effects, and our actions as causally effective events in the world. To me, the causal link between my action and its consequences is as objective and lawlike as any other causal regularity I perceive.

The descriptions by which we encode our actions conceptually are no different in form than those by which we encode other causally linked events. They are reducible to categorical indicatives that ascribe causal properties, namely actions, to agents.³³ The "accordion effect"³⁴ of action-descriptions makes it possible to ascribe a range of such properties, depending on the causal scope of the action the description is intended to capture. So, for example, "I (will) answer the telephone" captures a more restricted causal scope than "I (will) stop the ringing of the telephone."³⁵ Each expresses a different intention but both may describe the same physical action. Kant's primary concern is with action-descriptions that capture a more restricted causal scope – and so a narrower intention, for two reasons. First, he is concerned primarily with our immediate intentions, regardless of their further actual causal consequences, because our intentions are for him the primary locus of moral value. On Kant's view, the rightness or wrongness of actions are derivative from the goodness or badness of our intentions.³⁶ And second, he thinks that only the immediate objects of our intentions,

and certain basic actions³⁷ (i.e. the "exerting of every means so far as they are in our power" (Ak. 394)), are under our direct control (Ak. 401).³⁸

A *maxim* is a first-person categorical indicative judgment that ascribes a causal property, namely an act, to the subject. It thus both expresses an intention and describes an action. For this reason it may function as a resolution or as a prediction respectively – or both simultaneously. It may also describe a mental or a physical act. Kant derives the term "maxim" from the notion of a *maximum* as the rational idea of what is "greatest and absolutely complete ... in the division and unification of the knowledge of the understanding under one principle" (A 665/B 693). We have already seen in Sections II.B and C that the role of reason is to unify the knowledge of understanding under fewer, increasingly abstract and comprehensive higher-level principles. We have also seen above that Kant includes among the objects of knowledge of the understanding knowledge of our actions and of ourselves as agents. Kant defines *maxims of reason* as

all subjective principles which are derived, not from the constitution of the object but from the [speculative] interest of reason in respect of a certain possible perfection of the knowledge of the object (A 666/B 694).

Maxims may include principles of thought (such as "I will gather as much relevant information as possible before drawing any conclusions") as well as of action. What all maxims have in common is that they are guided by reason's interest in theoretical completeness. In the case of action, then, we seek an idea of reason – a highest-level comprehensive explanatory principle – that enables us to understand all of our intentions and resolutions as theoretically coherent principles of action derivable from it – and therefore, as we have seen in Section II.C, as objectively valid in light of it.

When Kant moves from a formal analysis of how reason operates in general to the content of a particular idea of reason, he moves from rational procedure in general to its application to a substantive theory. In the Canon of Pure Reason Kant describes the idea of reason that guides the formulation of maxims. It is the idea of a world ruled by moral law – a necessary idea of reason (A 812/B 840) that has objective reality (A 808/B 836) "in the concept of which we abstract from all the hindrances to morality (the inclinations)" (A 809/B 837). Kant's idea of a moral world comprises two elements: an entire world as a theoretical entity; and a system of operative principles – moral law – that is universally binding on it. Since in this world moral law is fully operative in the

behavior of its inhabitants, it explains and describes their behavior. And since it consciously guides their actions as well, it also prescribes it. Finally, since the moral law is authored by those inhabitants themselves, it describes and prescribes action that is rationally self-determining.

The concept of rationally self-determining action, uncoerced by empirical antecedents, is an application of the principle of self-caused causation, i.e. of transcendental freedom, to the special case of rational agency.³⁹ Kant's conception of unconditioned moral freedom, that is, as freedom to act in accordance with moral self-determination, is not different or separate from the highest-level unconditioned idea of freedom, but instead an instantiation of it in a particular kind of cause.⁴⁰ Between the Resolution of the Third Antinomy and the *Groundwork* Kant preserves the consistency of his conception of unconditioned moral freedom with his accounts of God, freedom and immortality as furnishing highest-level regulative ideas that govern and unify our patterns of thought. All lower-level principles describing causally connected experience of any kind, whether action, desire, free association, or external events, must terminate in the unconditioned idea of freedom. So the criteria of objectivity remain the same for all of them.

The unconditioned speculative idea of freedom as autonomous and operative moral law in turn provides the practical foundation for moral conduct. For from the interest of reason in the idea of a perfect world governed by moral law (which for Kant is also "a mere idea" (A 813/B 841); also see *Groundwork*, Ak. 407, 409, 412, 433, 434, 436 n., 439), we then derive the subjective principles that actually govern our moral agency:

Practical laws, in so far as they become at the same time subjective grounds of actions, that is, subjective principles, are entitled maxims. The judgment of morality, regarding its purity and consequences, happens in accordance with ideas, the adherence to its laws in accordance with maxims (A 812/B 840).

So we evaluate moral law as a speculative idea of reason; and we follow moral law by deriving practicable maxims from it. To the extent that moral laws are antecedents of a subject's actions, they are maxims; and the entire course of our lives are necessarily subject to them (A 812/B 840). Maxims of action such as "I (will) answer the telephone" or "I (will) return borrowed books" are lower-level conceptualized intentions, derivable as conclusions from higher-level moral principles, and finally from the highest-level unconditioned idea of a world governed by autonomous moral law, i.e. from the idea of

freedom. The relationship between maxims and moral ideas, then, is the relation between particular intentional action-descriptions and the more universal and comprehensive principles that explain them.

B. Universalization

By now it should be clear how we get from one to the other. To move from practicable maxims to the highest-level theoretical principle of operative, universal moral law from which those maxims are derived, we enact exactly the same procedure we use in any rational inquiry, namely hypothesis-construction of the sort described in Section II.C and exemplified in the idea of a causally powerful substance. In the case of action, we begin with both a prereflective rational idea of moral law (Ak. 402-3), and also an established particular intention to act. In light of this prereflective conception, we consider several such relevantly similar intentions. These are *ex hypothesi* certain, whereas "the universal is admitted as ... a mere idea, ... [and] ... the universality of the rule of which [the intention's maxim] is a consequence is still a problem" (A 646/B 674).⁴¹ So we scrutinize these intentions in order to ascertain whether they do, indeed, instantiate the moral law, such that the formulation of the moral law allows the derivation of the particular intentions from it. Again we deploy the "logical maxim through which to reduce, so far as possible, this seeming diversity, so that through comparison one might discover hidden identity" (A 649/B 677). That is, *we ascertain whether universalizing over the maxims yields the rule from which they are to be derived, such that the resulting rule applies not only to these established cases, but to all relevantly similar ones*. If the relevant intentions, and other citable ones, can be derived from this formulation of the moral law, we are then justified in regarding it as a universal principle that may predict and subsume future instances of that kind. Kant's remarks about hypothesis-construction in the first *Critique* provides a more complete account of his universalization procedure for moral maxims than is to be found anywhere in the *Groundwork*.

As we have already seen in Section II.B, universalizing a principle that functions as a major premise in a *Vernunftschluß* results in a categorical indicative that ascribes a predicate to all subjects of a certain kind. And we have also seen in Section II.C that in the *Groundwork*, Kant tells us to ascend in the series to the most universal and comprehensive principle of action from which lower-level maxims can be derived,

namely to "the universal concept of a rational being in general" (Ak. 412). So, for example, I might begin by universalizing over the maxims,

- (1) (a) I (will) return borrowed books,
- (b) I (will) pay my bills, and
- (c) I (will) keep my appointments.

This results in the principle

- (2) (a) I (will) keep my promises.⁴²

because (1.a-c) are all instances of promise-keeping. In the next step up in the ascending series of *Vernunftschlüsse*, we universalize over (2.a) to get

- (3) (a) Rational beings (will) keep their promises, (Ak. 422, *passim*)

since I identify myself generically as (among other things) a rational being, and suppose trustworthiness to be a characteristic of rationality. Next we universalize over (3.a) plus other, relevantly similar *Vernunftschlüsse* such as

- (b) Rational beings refrain from acting on the opportunity and desire to commit suicide, (Ak. 422, *passim*)
- (c) Rational beings sometimes cultivate some of whatever their natural talents, (Ak. 423, *passim*) and
- (d) Rational beings sometimes help some of the individuals in need they encounter. (Ak. 423, *passim*)

Principles (3.a-d) have in common, first, that they result from having universalized over more localized *Vernunftschlüsse* and maxims; and second, that they themselves, according to Kant, can be further universalized without theoretical incoherence. Like the ideas of God and immortality, the unconditioned idea of operative moral law can be derived by universalizing the form of principles in which reasoning itself, specifically about action, occurs:

... nothing remains but the universal conformity to law of actions in general, which alone should serve the will as its principle (Ak. 402).

Since actions are simply a species of cause, the universal law to which actions should conform has exactly the same form as any other law of nature to which other events must conform, namely the universalized categorical indicative that predicates effects as properties of their causes – here, universalizable actions as properties of rational beings. And since the unconditioned idea of a moral world includes the idea of moral law as universally binding on and operative for all its inhabitants, we must regard such universal laws as objective laws of nature in such a world (A 808/B 836, A 815/B 843). So (3.a-d) are, according to Kant, derivable from the higher-level *Vernunftschluß*,

- (4) (a) Rational beings perform only those acts that can be universalized as laws of nature. (Ak. 402, *passim*)⁴³

C. Kant's Explanatory Moral Theory

We have seen in Section II.C that according to Kant, a higher-level principle of reason – any principle – is objectively valid if it subsumes its lower-level principles, concepts, and representations, both intuitional and nonintuitional, under it, such that it allows their systematic inference as logically necessary syllogistic conclusions; and this degree of rational coherence simultaneously secures both the unity of the self and the objectivity of experience. So if Kant's account of moral reasoning is as integral to his more general account of reason and understanding as I have argued, the same formal criteria he develops for the latter – of objectivity and subjective rational coherence – will apply without revision to the former. And indeed we have also seen in Section III.B that Kant's proto-Hempel procedure of rational hypothesis-construction developed for reasoning in general applies without revision to moral reasoning.

But is Kant's substantive moral theory entirely susceptible to these formal criteria? Is it a genuine theory, the hypotheses of which follow from its higher-level principles, can be objectively tested and at least in many cases confirmed? And is it internally rationally coherent to the extent of securing unity of thought and experience for the self that accepts it? In what follows I will address the first question only; I defer the second to another occasion, for reasons of space.⁴⁴ Given Kant's special interest in moral theory, it is not surprising that the resources he has developed for constructing

Hempelian explanatory theories – of the cosmos, or of the human psyche – are, in fact, sufficient for a genuinely explanatory theory of moral freedom as well.⁴⁵ Here, then, are some further explanatory principles to be found in the text of the *Groundwork*:

- (4) (b) Rational beings perform only those actions that treat humanity as an end in itself. (Ak. 427, *passim*)
- (5) (a) Rational beings are motivated by *Achtung*⁴⁶ for the moral law. (Ak. 400, *passim*)
 - (b) Rational beings will universal law through their actions. (Ak. 431, *passim*)
 - (c) Rational beings legislate autonomously for a kingdom of ends through their actions. (Ak. 433, *passim*)
 - (d) Rational beings are noumenally free and phenomenally determined in their actions. (Ak. 451, *passim*)⁴⁷
- (6) (a) The causality of the will of rational beings is expressed in action performed out of *Achtung* for the moral law. (Ak. 453, *passim*)
 - (b) The freedom of rational beings as noumenal subjects is expressed in such moral action. (Ak. 454, *passim*)

Principles (3.a) through (6.b) are categorical indicative judgments about the behavior of certain sorts of phenomena, namely rational beings. The concept of a rational being can be similarly rendered by ascribing to it theoretical and practical rationality. So (3.a) through (6.b) contain no prescriptive terms. They also satisfy the basic Hempelian criteria that identify a set of principles as a genuine theory.

A theory begins with *hypotheses* – that is, proposed lawlike explanations of phenomena that are accepted conditionally on confirmation of their experimental regularities, and from which we should be able to infer causal regularities that can be experimentally tested. The more confirmable predictions we can make, the more credibility accrues to the hypothesis. We have already seen in Section II.C that Kant's concept of the hypothetical use of reason enables us to do just this; and principles (1.a-c) and (3.a) suggest how this might work in the case of action. (3.a), together with the suppressed premise that you are a rational being, implies that you will return this borrowed book (1.a). If you do, then you have confirmed at least one experimental prediction of (3.a). A second prediction of (3.a) might be (1.b). If you do, indeed, pay

your bills, then you have further confirmation of (3.a). Notice that all the principles in group (3) are susceptible of the same sort of experimental testing, on oneself as easily as on others.

As Kant has already instructed us, the more confirming instances accrue to (3.a), the more we are entitled to regard (3.a) not just as a hypothesis but as a *law*, i.e. a true hypothesis stated in the form of a generalization that ascribes causal properties to subjects. Like all the principles in group (3), (3.a) satisfies the *nomological* requirement that that it support counterfactual conditionals: Rational being would keep any promises they made, and would have kept any promises they had made. (3.a) ranges over not only the actual past, present, and future, but over possible pasts, presents and futures as well. It has universal rather than merely spatiotemporally limited application. Thus (3.a) contrasts with a mere accidental generalization such as

(7) (a) Anyone who keeps her promises is a rational being,

since someone could conceivably keep her promises – say, because she had been hypnotized into doing so, without being a rational being.⁴⁸

Explanatory theories contain both *lower-* and *higher-level laws*. The latter are laws that satisfy the same criteria just discussed, but that generalize over lower-level laws with respect to more abstract features of the phenomena described. We have already seen in Section II.B that Kant's account of reason supplies us with plenty of those. (4.a) and (4.b) above are higher-level laws from which (3.a-d) can be deduced as experimental predictions: For example, because a rational being performs only those acts that treat humanity as an end in itself, she will keep her promises (because keeping one's promises treats humanity as an end in itself). Since Hempel's own covering law schema

(covering laws) L_1, L_2, \dots, L_n	}Explanans
<u>(particular circumstances) C_1, C_2, \dots, C_m</u>	
(phenomenon to be explained) E	}Explanandum

is a modern elaboration of Kant's ascending series of *Vernunftschlüsse*, it naturally organizes some of the principles to be found in Kant's moral theory quite well:

- (1) (L₁, L₂) Rational beings keep their promises (3.a); and rational beings sometimes help some of the individuals in need they encounter (3.d);
 ((C₁, C₂) Your colleague borrowed your textbook, and promised to return it in time for you to prepare your lecture;
 (E₁) ∴ Your colleague returns your textbook in time for you to prepare your lecture.
- (2) (L₃, L₄) Rational beings perform only those acts that can be universalized (4.a) and that treat humanity as an end in itself (4.b);
 (C₃, C₄) Keeping one's promises and sometimes helping some of the needy can be universalized (4.a), and also treat humanity as an end in itself (4.b);
 (E₂)(= (L₁, L₂)) ∴ Rational beings keep their promises (3.a); and sometimes help some of the individuals in need they encounter (3.d).

Just as Kant's higher-level *Vernunftschlüsse* finally terminate in unconditioned ideas of reason, Hempelian higher-level covering laws finally terminate in a theory's *theoretical constructs* and the principles governing it. The theory is a higher-level hypothesis that is accepted as true because it successfully explains lower-level, law-governed uniformities as manifestations of "deeper" unobservable entities and processes that are themselves governed by theoretical laws and principles. Examples of such constructs from Kant's moral theory appear with increasing frequency as the level of abstraction of the principles increases: "Reason," "will," "law," "humanity," and "end" are theoretical constructs in (4), according to this description, as are "kingdom of ends," "freedom," and "noumena" in (5). All are abstractions that combine to form an ideal type whose behavior explains the uniformities of behavior of rational beings as described in (3).

These theoretical constructs are, like scientific theoretical constructs, governed by two kinds of principles. First, there are *internal principles* that describe their behavior. Applied to maxims of action, Kant's account of the hypothetical use of reason describes the operation of the rational will as legislating moral law; his account of the ideas of reason explain how the concepts of humanity as an end in itself and of the kingdom of ends function for us; and in the Resolution of the Third Antinomy and third chapter of the *Groundwork* he explains in what freedom, autonomy, and the noumena-phenomena

distinction consist. These accounts provide internal principles that describe and explain the behavior of these theoretical constructs.

In addition to internal principles, Kant's explanatory moral theory also contains *bridge principles* that connect these constructs with the familiar empirical phenomena of moral action. (6.a) and (6.b) are bridge principles. Both contain what we might describe as "double connections". First, there is the *causal* double connection in (6.a), between

- (i) the causality of the will and the feeling of *Achtung*, and
- (ii) the feeling of *Achtung* and the resulting moral action:

Rational principles of action command *Achtung*, which in turn motivates moral action.

Second, there is the *evidential* double connection in (6.b), between

- (i) freedom and the noumenal subject, and
- (ii) the noumenal subject and the moral action:

Freedom is manifested by a subject whose behavior is not determined by empirical inclinations – that is, a noumenal subject, and noumenal subjecthood is evinced by moral action. In both cases, these principles link the moral actions we observe with the theoretical constructs that ultimately explain them. So Kant's moral theory does satisfy the basic requirements of a genuine theory, and so is fully congruent with his account of reason and theory-construction more generally.

I have argued that Kant's moral theory explicates substantive ethical principles in terms of the "universal concept of a rational being in general ... i.e. entirely as metaphysics" (Ak. 412), and so as categorical principles in the indicative mood. I have not mentioned Kant's famed categorical imperative at all. One reason for this is that, as a purely exegetical matter, Kant himself does not make much use of it. Out of thirty-two formulations of the fundamental principle of morality in the *Groundwork*, only four are in the imperative mood.⁴⁹ The remainder are in the form either of laws,⁵⁰ or of commands,⁵¹ neither of which express the imperative.⁵² This bias toward the categorical indicative is, as we have already seen, consistent with his metaethical conception of freedom as a highest-level explanatory idea of reason in the first *Critique*; and he reiterates this bias consistently throughout the *Groundwork* and second *Critique*.

But more importantly, Kant does not think imperatives apply to rational beings as such.⁵³ In the *Groundwork* Kant denies repeatedly that the "ought" (*sollen*) is to be found in the *intelligible world*, i.e. the viewpoint of reason and conceptualization that furnishes the cognitive foundation (or *Grund*) for the viewpoint of understanding and

empirical experience. If we were solely members of the intelligible world, he says, all our actions just *would* conform to moral law (Ak. 454). Because reason would be motivationally effective without any hindrance in such a being, the expression of moral intention would be not "I ought" but "I will" (Ak. 449). "The moral 'I ought'," Kant says, "is thus an 'I will' for us as members of the intelligible world..." (Ak. 455). A rational and perfectly good or holy will, he tells us, would be governed by objective moral law; but it would not be necessitated or compelled to conform to it, as we are. Instead, such a will would conform to the moral law naturally, in accordance with its subjective constitution (Ak. 414). Its maxims of action would necessarily conform to moral law, but unlike us, it would have no obligation or duty to do so (Ak. 439), since, as Kant reminds us in the second *Critique*, such a being would be incapable of any maxims that conflicted with the moral law – a model to which we as sensuous beings must (albeit in vain) aspire (Ak.V, 32; also cf. 84-85).

So as is true for Kant's metaphysics more generally, his moral theory is fashioned primarily with an eye to its application to rational beings in general.⁵⁴ The categorical imperative enters in only as a problem of the application of this theory to imperfectly rational instances. How this might affect human beings in particular is a different question for a different paper.

Notes

This discussion is excerpted from my *Rationality and the Structure of the Self, Volume II: Kant's Metaethics* (unpublished manuscript, 1994). It has benefited from conversations with Henry Allison, Gordon Brittan, Kenneth Winkler, Günter Zöller, and the Wellesley Philosophy Department Faculty Seminar.

Unfortunately, editorial restrictions on space necessitate deferring many issues raised in this essay to a fuller treatment in the larger project. Thus interpretations and arguments are often summarized rather than developed in depth, and extended exposition is replaced by the presupposition of a familiarity with all parts of Kant's *Kritik der Reinen Vernunft* (I use the edition herausg. von Raymund Schmidt [Hamburg: Felix Meiner Verlag, 1976]), the standard commentaries to that work (e.g. Allison, Bennette, Brittan, Ewing, Guyer, Kemp Smith, Melnick, Paton, Strawson, Vaihinger, Wolff), and German idiom. Nor is there any examination of relevant competing views,

such as Allison's *Kant's Theory of Freedom* (New York: Cambridge University Press, 1990), Susan Nieman's *The Unity of Reason* (New York: Oxford University Press, 1994), Onora O'Neill's *Acting on Principle* (New York: Columbia University Press, 1975) and her *Constructions of Reason* (New York: Cambridge University Press, 1989). These, too, are to be found in *Kant's Metaethics*.

¹John Rawls, "Outline of a Decision Procedure for Ethics," *Philosophical Review* 66 (1957), 177-197; reprinted in *Ethics*, Ed. Judith J. Thomson and Gerald Dworkin (New York: Harper and Row, 1968), 48-70.

²*The Possibility of Altruism* (Princeton: Princeton University Press, 1978).

³*A Theory of the Good and the Right* (New York: Oxford University Press, 1980).

⁴*Reason and Morality* (Chicago: University of Chicago Press, 1979).

⁵*Morals by Agreement* (New York: Cambridge University Press, 1984)

⁶*A Theory of Justice* (Cambridge, Mass.: Harvard University Press, 1971).

⁷"Justice as Fairness: Political not Metaphysical," *Philosophy and Public Affairs* 14, 3 (1985), 223-251.

⁸*A Theory of Justice*, 16, 47, 172.

⁹*Grundlegung zur Metaphysik der Sitten*, herausg. von Karl Vorländer (Hamburg: Felix Meiner Verlag, 1965), Ak. 413. Translations from the German texts are my own unless otherwise indicated. My renderings are generally more literal and attentive to Kant's fondness for colloquialisms – and so, I think, truer to Kant's thought – than the standard ones.

¹⁰*Op. cit.*

¹¹*Kritik der Praktischen Vernunft*, herausg. von Karl Vorländer (Hamburg: Felix Meiner Verlag, 1974).

¹²I reject Kemp Smith's translation of *unmittelbar* as "immediately" because of the latter's temporal connotations, which are inappropriate to Kant's meaning. Instead I substitute "direct" or "unmediated," depending on context.

¹³Of course this is not to claim that the intuitional representations of those objects are veridical, or even *about* those objects, but merely, as it were, *from* them. I discuss at length the textual evidence for the thesis that sensible representations are caused by things in themselves, and the case for this unorthodox use of the term "cause", in

footnote 17 of "Xenophobia and Kantian Rationalism," *The Philosophical Forum* XXIV, 1-3 (Fall-Spring 1992-93), 188-232. The term *appearance* (*Erscheinung*) raises more questions than it answers, so I'm going to disregard it for purposes of this discussion.

¹⁴See *ibid.*, Sections III, "The Concept of Personhood," and IV, "Self-Knowledge," for an extended discussion.

¹⁵That is, "logically necessary" in Kant's anachronistic and overly rich sense of comprising all of the constraints on thought imposed by the forms of judgment enumerated in the Table of Judgment. (A 70/B 95)

¹⁶I think Kant was wrong to drop this useful notion from the B Edition, since it captures the case of recognizing something as an object independently of knowing what kind of object it is.

¹⁷In German the categorical indicatives "You answer the phone," "You are answering the phone," and "You will answer the phone," ordinarily are all expressed by the sentence, "Sie antworten das Telefon" Compare, for example, "Wir fahren nach München," which may be translated as "We go to Munich," "We are going to Munich," or "We will go to Munich," depending on context. This should be borne in mind in what follows.

¹⁸Also see Kant's *Logic*, trans. Robert Hartmann and Wolfgang Schwartz (New York: Bobbs-Merrill, 1974), paragraphs 24-29, 60 fn. 2; and *Lectures on Logic*, trans. J. Michael Young (New York: Cambridge University Press, 1992), 374 and 601, Note.

¹⁹Even synthetic a posteriori judgments about particular empirical objects or events owe their objectivity to the synthetic a priori judgments they necessarily presuppose. I discuss the relation of transcendental to empirical concepts in "Xenophobia and Kantian Rationalism," Section II. Of course a judgment can be objective without being true.

²⁰If this is what objectivity is for Kant, how are mistaken judgments to be explained? Kant says that illusion can arise (a) by attributing to the object as an ontologically independent thing in itself a predicate that belongs only to its representation (B 70, fn.); (b) by falsely supposing some intuitional representations to indicate the existence of an object when in fact they are produced by the imagination from "previous outer perceptions, which, as has been shown, are possible only through the reality of outer things;" (B 278) and (c) by "tak[ing] the subjective necessity of a

particular connection of our concepts that is to the benefit of the understanding for an objective necessity in the determination of things in themselves" (A 297/B 353).

These passages indicate two ways in which erroneous judgments can arise: (1) by mistaking a representation or conjunction of representations for an ontologically independent entity, event, or thing in itself, i.e. by taking a predicate to be more ontologically fixed than it is; and (2) by mistaking a representation or conjunction of representations for an empirically external entity or event, i.e. by taking a product of imagination to be a product of transcendental synthesis. Kant later describes the first mistake as arising "by way of experience," and the second as arising "by mere play of imagination" (A 376).

The antidote for both, he says, is to "proceed according to the rule: Whatever connects with a perception according to empirical laws is actual." (A 376) This won't do, since of course our empirical laws may be wrong or imprecise or govern a particularly vivid dream or fantasy. What he should have said was that whatever is theoretically coherent with *all* of our law-governed experience is conditionally actual, and that whatever, in addition, is coherent with all fully informed law-governed experience is unconditionally actual (see Section II, below). That way it is possible for me to revise my judgment that the phone ceased ringing because I answered it in light of further information, if such information is forthcoming.

²¹See, for example, his discussion of the ideas of virtue, a just constitution, the order of nature, and humanity at A 312/B 369 - A 319/B 376.

²²As is so often true, Kant does not actually argue for these claims (or if he does, the arguments are not very good); he simply states them and relies on their intuitive philosophical plausibility. It is not impossible that they could be given a more rigorous, discursive form. But I will not attempt that here, since my primary concern is to establish what Kant thinks rather than whether he is justified in thinking it.

²³I use the term *Vernunftschluß* (inference of reason) instead of "syllogism" in order to emphasize its centrality to Kant's conception of how reason operates.

²⁴"All our knowledge begins with the senses, goes from there to understanding, and ends with reason, beyond which there is no higher faculty to be found in us for fashioning the matter of intuition and bringing it under the highest unity of thought. (A 299/B 355)...[I]n inference reason seeks to reduce the great manifoldness of knowledge

from the understanding to the smallest number of principles (universal conditions) and thereby to produce in it the highest possible unity" (A 305/B 361).

Compare A 86/B 118-9, where Kant acknowledges Locke's contribution to epistemology: "The impressions of the sense supplying the first occasion, the whole faculty of knowledge opens to them, and experience is brought about... Such an investigation of the first strivings of our faculty of knowledge in order to ascend from particular perceptions to universal concepts is no doubt of great use."

²⁵See "Xenophobia and Kantian Rationalism," Sections I-III.

²⁶In the *Lectures on Logic* Kant defines a condition as that in the subject "that makes the predicate be attributed to it" (497). Despite the category mistake I will assume we know what he means, sort of.

²⁷"[T]hat a certain systematic unity of all possible empirical concepts so far as they can be derived from higher and universal ones must be sought is elementary, a logical principle, without which no employment of reason would occur, since we can conclude from the universal to the particular only so far as universal properties of the thing are presupposed, under which the particulars stand" (A 652/B 680).

Kant goes to some length to distinguish the unity of understanding from the unity of reason (see, for example, A 302/B359, a 306-307/B 363, A 311/B 367, A 409/B 436, A 422/B 450), but only in order to establish how dependent the former is on the latter (see, for example, A 299/B 355, A 305/B 361 – A 306/B 363, A 326/B 383, A 329/B 385, and especially A 647/B 675 – A 651/B 679, quoted below in Note 31).

²⁸In the footnote to A 337/B 395 Kant declares that "[a]ll matters with which [the science of metaphysics] may otherwise concern itself serve merely as a means for reaching [the ideas of God, freedom and immortality] and their reality. It does not need [the ideas] for the purposes of natural science, but in order to pass beyond nature." In this passage Kant is saying that the main purpose of the science of metaphysics is to pass beyond nature and establish the reality of the ideas of God, freedom and immortality. He is also saying that the science of metaphysics does not need these three ideas for natural science itself, but instead for passing beyond it. He is *not* saying that the procedures of rational inference that characterize what he calls the science of metaphysics are not needed for natural science, nor that the impulse to seek increasingly inclusive and universal covering laws are not needed for natural science.

The passages already discussed make clear that he regards the procedures of natural science as continuous with and inevitably leading to metaphysics' three ideas of reason.

²⁹Although Kant does not explicitly state this, we can assume that this procedure also requires us to take care to preserve the integrity of the universalization on the one hand, and of the firmly established instances on the other. Although we may tinker with the formulations of each, we may neither distort the referential scope of the universalization through rationalization, nor deny nor dissociate any of the instances it is intended to explain on the other. (I discuss the concepts of rationalization, dissociation and denial at greater length in "Pseudorationality," in Amelie O. Rorty and Brian McLaughlin, Eds. *Perspectives on Self-Deception* (Los Angeles: University of California Press, 1988), 297-323. My account of these concepts is derived from Kant's accounts of reason and self-deception in the first *Critique* and *Groundwork*, although Kant has different uses for the term "pseudorationality." See especially A 311/B 368, A 339/B 397, A 406/B 433, A 421/B 449, A448/B 476, A 462/B 490, A 490/B 518, A 497/B 526, A 644/B 672, A 795/B 823.) Given the particular instances with which we began, the universalization is valid only if it subsumes all of them and all relevantly similar ones, and excludes those that are too dissimilar to have been grouped with them at the outset.

In this procedure two interconnected elements – the universalization as expressing a rational idea that is in some degree fixed by innate, a priori concepts of experience, and the certain particular empirical experiences it is intended to explain – are mutually determining. Criteria of validity for the scope of the universalization depend on the lower-level instances it subsumes, and criteria of salience, similarity and relevance of those instances depend on the rational idea the universalization attempts to express. It is because Kant presupposes the same innate conceptual structure both to the certain particular instances with which we start and to the inductive generalization we initially formulate to explain them that he is so sure this generalization will eventually prove to be a universal principle.

³⁰Kemp Smith gets this sentence wrong. <<*Eine solche*>> refers to the unity of reason. It is that and not the absolutely fundamental power that Kant wants to claim we must seek for the benefit of reason.

³¹"The hypothetical employment of reason therefore gets at the systematic unity of the knowledge of understanding, and this is the touchstone of the truth of its rules. (A 647/B 675)... [this] systematic or rational unity is a logical principle for assisting the understanding by means of ideas where the understanding alone does not reach rules; and at the same time for systematically giving to the diversity of its rules uniformity and consistency under one principle, and thus for providing coherence as far as possible. (A 648/B 676) ...the law of reason directing us to seek this unity is necessary, because without it we would have no reason at all; without this no coherent employment of the understanding; and in the absence of this no sufficient criterion of empirical truth" (A 651/B 679).

³²In this case the given immediate object I intuitively represent to myself is myself as I am in myself, i.e. as the *noumenal subject* to which I have no direct epistemic access. By acting I cause in myself intuitional representations that are, like other intuitional representations, passively received in sensibility. And as is true for other objects, the representations I as an acting noumenal subject cause in myself by acting do not have as their content the noumenal agent that causes them. Instead, these intuitional representations are, in turn, collected under higher nonintuitional representations – in this case, of the empirical self and finally of rational personhood – in the manner described in Section I.B and elsewhere. Whereas the spontaneous acts of understanding and reason cause us to become aware of ourselves as empirical and rational subjects, intentional actions – the "outward expression" of the will (A 798/A 826) – cause us to become aware of ourselves as empirical and rational agents. Both are necessary constituents of our concept of ourselves as persons.

³³In the *Groundwork* Kant often formulates his examples of maxims more complexly, in an "Out of ...[motive], I will ...[intention/resolution], in order to ...[purpose]" format. This is for the pedagogical purpose of contrasting morally valid with invalid motives, categorical with hypothetical imperatives, and identifying what sorts of results of action are irrelevant to the assessment of moral worth.

³⁴See Joel Feinberg, "Action and Responsibility," in *Doing and Deserving* (Princeton: Princeton University Press, 1970).

³⁵Compare the flexible temporal range of the indicative in German with its comparatively restricted temporal range in English. My analysis covers both. Also see Footnote 17.

³⁶For a discussion of the implications of this for Kant's purported deontology, see my "A Distinction Without a Difference," *Midwest Studies in Philosophy VII: Social and Political Philosophy* (1982), 403-435.

³⁷See Arthur Danto, "Basic Actions," in Norman S. Care and Charles Landesman, Eds. *Readings in the Theory of Action* (Bloomington: Indiana University Press, 1968).

³⁸"For all these [other] results (pleasant states and even the promotion of others' happiness) could have been brought about by other causes as well, and so did not need the will of a rational being... "(Ak. 401). Actually, even if we interpret the immediate objects of our intentions and resolutions as narrowly as possible, as involving only "attempt" verbs such as *trying* or *moving* to answer the phone rather than as "success" verbs such as *answering* the phone, the immediate objects of our intentions and resolutions could be brought about by causes other than our rational will, such as hypnosis or cortical stimulation.

³⁹"[S]ince the power of beginning a series in time entirely from itself is thereby proved ..., it is now also permissible for us to allow within the course of the world different series as capable in accordance with their causality of beginning of themselves, and so to ascribe to substances themselves a power of acting from freedom. ... For here we speak not of an absolutely first beginning in time, but rather in causality. If I now, for example, arise from my chair in full freedom and without the necessarily determining influence of natural causes, a new series thus begins simply in this event, together with all its natural consequences into infinity.... For this resolution and act do not lie in the succession of purely natural effects, and [are] not simply a continuation of them" (A 450/B 478).

⁴⁰Notice that self-determining action in the transcendently free sense that Kant allows at A 450/B 478 and requires at A 809/B 837 - A 810/B 838 is distinct from the practically free action that "can be proved through experience" (A 802/B 830) and about which Kant warns us that

[w]hether reason itself in the actions through which it prescribes laws is not again determined by further influences, and whether that which, in relation to

sensuous impulses, is called freedom, may not, in relation to higher and more remotely operating causes, be nature again, does not concern us in the practical field, as we are demanding of reason nothing but the rule of conduct ... (A 803/B 831).

Whereas practical freedom is a property of human beings in the empirical world, transcendental freedom is a property of rational inhabitants of the moral world.

⁴¹At A 480/B 508, Kant makes the brazen claim that "[i]n the universal principles of morals nothing can be uncertain, because the principles are either completely void and meaningless, or must emerge from our rational concepts. In natural science, on the other hand, there is endless conjecture, in regard to which certainty can never be expected. For the natural appearances are objects which are given to us independently of our concepts, to which the key lies not in us and our pure thought, but outside us; and therefore in many cases is not found; and so no assured explanation can be expected." He makes a similar point at Ak. 404. In these passages I take Kant to be bickering with Aristotle's claims, in Chapter One of the *Nicomachean Ethics*, about the degree of precision to be expected in ethics as opposed to the natural sciences. But even if he is not, the passages should be taken with a grain of salt because he has already provided ample evidence and discussion of the ways in which uncertainty can arise in the derivation of concepts from our reason, and in our attempts to ascertain our true motives without self-deception.

⁴²Of course this requires us to conceive of the subject of (1) as a universalization over discrete occurrences of subjective agency. But this is consistent with Kant's infamous claim in the A Paralogisms that "[t]he identity of the consciousness of myself at different times is therefore only a formal condition of my thoughts and their coherence but in no way proves the numerical identity of my subject" (A 363). See the rest of that paragraph for elaboration and also the Strawsonian fireworks in the footnote at A 364. In keeping with this doctrine we might indexicalize the "I" of act-token descriptions as follows: "I₁ (will) return this book," "I₂ (will) pay my bills," and so forth.

⁴³That Kant's account of moral reasoning is merely an application of his account of reasoning in general should not be surprising. It would be very odd if, given Kant's intellectual delight in architectonic and his preoccupation with it, his account of empirical human actions failed to conform to the category of causality; and odder still if

they did but yet failed to conform to the higher-level theoretical principles all events that fall under that category engender. Oddest of all would be if Kant's account of action satisfied all of these requisites for cognitive objectivity, yet required a completely separate and unconnected treatment of moral objectivity. Given Kant's emphasis on theoretical coherence as the foundational lynchpin of the a priority, logical necessity, and universality that he claims give *any* judgment objective validity in the first place, this would be no account of objectivity at all. What has not been sufficiently appreciated, despite Kant's repeated reminders, is how important and constant all of these requisites are.

⁴⁴I address the second question in "The Meaning of 'Ought' and the Loss of Innocence," Invited Paper delivered to the American Philosophical Association Eastern Division Convention, 1989; and give it a fuller treatment in *Kant's Metaethics*.

⁴⁵At Ak.V.137 in the *Critique of Practical Reason*, Kant reiterates his warning from the Dialectic of the first *Critique*, that the ideas of God, freedom and immortality are mere suppositions that reason must make for practical purposes, and not knowledge, in the following words:

[The predicates derived from our own nature we might be inclined to ascribe to God, the intelligible world, and immortality] can never be used in a theory of supersensuous beings and ... so on this side do not have the power to ground a speculative knowledge, but rather restrict their use solely to the practice of the moral law.

Astonishingly, Beck translates *spekulative* as "theoretical." But the original context makes clear that Kant means to warn us against confusing rational hypothesis-construction with knowledge in the technical, experiential sense he has defined in the first *Critique*: While we need to suppose the validity of these ideas of reason in order to act morally, we cannot thereby infer that we know what God, freedom, or immortality is. Kant here inveighs against the same sins of speculative metaphysics to which he devoted the Dialectic of the first *Critique*. He is not attacking rational theorizing as such (since, as we have already seen, this is precisely how he claims reason operates).

⁴⁶Again I use the German term because I find the English translations inadequate. I discuss the correct translation of *Achtung* in *Kant's Metaethics*. Also see "The

Obligations of Philosophical Performance," delivered to the Greater Philadelphia Philosophy Consortium, February 1994.

⁴⁷Here and below I make an unargued assumption about the semantic equivalence of Kant's use of the terms "noumenal" and "intelligible" on the one hand, and "phenomenal" and "sensible" on the other. To defend this assumption would require a paper of its own, but I have every faith that such a defense would succeed.

⁴⁸It might be objected that suppressed premises of the sort mentioned above, that the individual in question is a rational being, are themselves accidental generalizations over instances of behavior that happen to, but may not in all cases evince rationality. But this objection could be raised as well of any suppressed premise that identifies an event or state of affairs in its subject term: "This object is a paraffin candle" is similarly an accidental generalization over instances of object behavior that may not in all foreseeable cases evince paraffin candlehood (perhaps it will bob about when thrown into boiling water, like plastic, instead of melting). Hence these two kinds of suppressed premise must stand or fall together.

Similarly, it won't do to object that, unlike scientific laws, (3.a), and indeed all the judgments in (3), are true by stipulative definition of "rational being," since the same object could be raised about the status of the hypothesis, "Paraffin melts when put into boiling water": If it doesn't melt, then either we were wrong about what that substance is, or we were wrong about how paraffin behaves. Again both kinds of hypothesis must stand or fall together.

⁴⁹I.e. at Ak. 402, 433, 436, and 438.

⁵⁰I.e. at Ak. 424, 431, 432, 434, 436 (twice), 438, 439, 444, 447, 458 (twice), 461, 462.

⁵¹I.e. at Ak. 421 (twice), 429, 432, 434, 436-7, 437 (three times), 437-8, 438, 439, 440, 447. Note that what is widely considered the first, "universal law" formulation of the categorical imperative at Ak. 421 is mistranslated by Paton and in fact contains no "ought".

⁵²Many assume that imperatives and commands are interchangeable, and Kant himself sometimes speaks this way. But they are not, and Kant knows it (for example, compare his conflicting definitions of an imperative at Ak. 413 and Ak. 414). I discuss the distinctions among a law, a command, and an imperative at greater length in *Rationality and the Structure of the Self, Volume II*.

⁵³Although he says they *can* (Ak. 425).

⁵⁴See, for example, Ak. 408, 410 n., 413, 425, and 447.