

*Kant's Incorporation Requirement: Freedom and Character in the Empirical World**

RICHARD McCARTY
East Carolina University
Greenville, NC 27858
USA

In *Religion within the Boundaries of Mere Reason* Kant wrote that 'freedom of the power of choice has the characteristic, entirely peculiar to it, that it cannot be determined to action through any incentive *except insofar as the human being has incorporated it into his maxim.*'¹ This is an obscure statement, in both meaning and provenance. Yet almost all recent interpreters of Kant's practical philosophy find it crucial for understanding his theories of freedom and motivation, since it seems to indicate what we are required to do in order to act by our own free choice. Here I refer to Kant's statement expressing the requirement that incentives be incorporated into maxims as his 'incorporation requirement.' How that requirement is best understood will be the leading question in what follows: a question I shall answer by showing why the incorporation requirement, and Kant's theories of freedom and mo-

* I am grateful for Elizabeth Radcliffe's help in clarifying several of the arguments that follow.

1 Immanuel Kant, *Religion within the Boundaries of Mere Reason and Other Writings*, Allen W. Wood and George di Giovanni, trans. and ed. (Cambridge: Cambridge University Press 1998), 49/6:23-24. Subsequent references to *Religion* refer to this edition, unless otherwise indicated.

tivation, should be understood differently from the way they are now usually understood.

The Kantian perspective in ethics is perennially subject to criticism for the obscurity of its explanations of freedom and moral responsibility, and for the rigidity of its criterion for the moral worth of actions. Although gifted scholars have written in Kant's defense on these points, I believe present interpretive strategies for responding to these stock criticisms have made matters worse; especially from the perspectives of critics. Recent work on Kant's theory of freedom struggles to come to grips with the determining influence of character that is acknowledged by almost everyone, including Kant himself.² Recent accounts of Kant's theory of moral appraisal propose *ad hoc*, psychologically implausible models of human motivation designed to allow actions on admirable emotions to qualify as 'acting from duty,' in order to merit 'moral worth.' By calling attention to an arcane requirement for free choice now considered essential for understanding Kant's practical philosophy, my overall aim here is to push prevailing views of his theories of motivation and moral appraisal toward positions more conciliatory to those of critics. To this end, in what follows, I shall be emphasizing moral responsibility for *character*, not just for action, and moral admiration for *virtue*, not just for moral worth.

John Silber may have initiated recent interest in the incorporation requirement, almost fifty years ago.³ Not long afterward, and for the next couple of decades, John Rawls began emphasizing it in his influential lectures on Kant's practical philosophy.⁴ Gerold Prauss found the requirement interpretively fruitful in the early 1980s; and by 1990 Henry Allison had labeled the passage from *Religion* containing the incorporation requirement the 'Incorporation Thesis,' hailing it as 'the centerpiece of Kant's conception of rational agency.'⁵ Recent interpret-

2 See Immanuel Kant, *Critique of Pure Reason*, Paul Guyer and Allen W. Wood, trans. and ed. (Cambridge: Cambridge University Press 1998), A554-5/B582-3. Subsequent references to *Pure Reason*, or to the first *Critique*, refer to this edition, unless otherwise indicated.

3 John R. Silber, 'The Ethical Significance of Kant's *Religion*,' in Immanuel Kant, *Religion Within the Limits of Reason Alone*, Theodore M. Greene and Hoyt H. Hudson, trans. (New York: Harper & Row 1960), lxxix-cxxxiv, esp. xcv-xcvi. Prior to Silber's introductory article the incorporation requirement had been discussed briefly in H. J. Paton, *The Categorical Imperative* (Chicago: University of Chicago Press 1948), 275.

4 See John Rawls, *Lectures in the History of Moral Philosophy*, Barbara Herman, ed. (Cambridge, MA: Harvard University Press 2000), 294.

5 See Gerold Prauss, *Kant über Freiheit als Autonomie* (Frankfurt am Main: Vittorio

ers still rely on the so-called the Incorporation Thesis, as can be seen, for example, in the work of Christine Korsgaard, Paul Guyer, and the latest work of Allen Wood.⁶ To my knowledge, no one has expressed any doubt about the interpretive adequacy of that Thesis, although Rüdiger Bittner has challenged it on conceptual grounds.⁷ With the possible exception of Silber, all of these interpreters, and many others, have read the incorporation requirement as expressing the view that freely choosing to act is incompatible with being causally determined to act by the impelling forces of desires or incentives.⁸ Two advantages of reading the requirement this way are the following.

First, this reading helps to sharpen the contrast between Kant's rationalist theory of motivation and Hume's empiricist, 'belief-desire' model. As Hume's theory has it, we are caused to act always by desires for objects we believe our actions will produce; and when our desires compete, the strongest among them causally determines our choice to act. But Kant's interpreters see the incorporation requirement as telling us, on the contrary, that whenever we act by our own free choice we 'incorporate' the desires or incentives on which we choose to act into maxims, making them the reasons for our actions. We therefore act on freely chosen reasons, and we are never caused to act by the strengths of our desires, which makes us free to act contrary to our strongest desires.

Kolstermann 1983), 93-4; and Henry E. Allison, *Kant's Theory of Freedom* (Cambridge: Cambridge University Press 1990), 189. Allison gave the name 'Incorporation Thesis' to what Rawls had called the 'Principle of Election.' See also Andrews Reath, 'Kant's Theory of Moral Sensibility,' *Kant-Studien* 80 (1989) 284-302, 290n, and Allison, (1990), 126.

- 6 See Christine M. Korsgaard, *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press 1996), 162, 165; See also Christine M. Korsgaard, *The Sources of Normativity*, Onora O'Neill, ed. (Cambridge: Cambridge University Press 1996), 94; Paul Guyer, *Kant on Freedom, Law and Happiness* (Cambridge: Cambridge University Press 2000), 293-4; and Allen W. Wood, *Kant's Ethical Thought* (Cambridge: Cambridge University Press 2000), 51-3. Earlier, Wood defended an interpretation of Kant's theory of freedom seemingly antithetical to the Incorporation Thesis. See 'Kant's Compatibilism' in Allen W. Wood, ed., *Self and Nature in Kant's Philosophy* (Ithaca, NY: Cornell University Press 1984), 73-101.
- 7 Bittner quipped: 'An incentive that cannot determine the will unless the agent has [incorporated it into] his maxim is in fact an incentive that cannot determine the will, period'; see Rüdiger Bittner, *Doing Things for Reasons* (Oxford: Oxford University Press 2001), 47. For some other conceptual problems related to the Incorporation Thesis see my 'The Maxims Problem,' *Journal of Philosophy* 99 (January 2002) 29-44.
- 8 Silber's interpretation was that we are always caused to act by the impulsive strength of our incentives, though it is we who freely choose their relative strengths (see xcv).

Second, the requirement has been invoked in response to a familiar objection to the Kantian approach to moral appraisal: the approach that denies moral worth to actions motivated by admirable concerns of friendship, love, felt gratitude, and other emotions. In defense of Kant's rigoristic criterion for 'moral worth,' recent interpreters remind us that psychologically forceful desires and emotions do not cause our actions, nor are we caused to act by the moral incentive of duty.⁹ Rather, we freely choose by an act of incorporation. So whenever the incentive of duty prompts someone to act, how could it be good for her to pass over that moral motive for her action and choose to act on some motive of inclination instead?¹⁰ When we understand the psychological procedure for acting by our own free choice, Kant's defenders say, we are in a better position than his critics to understand why actions are appraised as morally worthy only if they are motivated by the moral incentive of duty.

Following Henry Allison's lead, in what follows I shall refer to the way recent interpreters have understood Kant's incorporation requirement as the 'Incorporation Thesis.' Here I shall be arguing against that Thesis as an interpretation of the text in *Religion* where Kant expressed the incorporation requirement. In Section I just below I begin by attempting to clarify the terms of that requirement, explaining what *maxims* and *incentives* are, and what it means to *incorporate* incentives into maxims. Section II focuses next on the context in which Kant expressed the incorporation requirement. Interpreting the wider argument for which that requirement serves as a premise will help us see that it has no direct application to the way we make everyday free choices, contrary to what proponents of the Incorporation Thesis have assumed. Sections III and IV will then show that the free choice to which the requirement was meant to apply belongs to a noumenal world, and that this choice makes its appearance in the phenomenal world in the form of a human being's empirical character. Section V, finally, introduces a new response to critics' complaints over moral worth and acting from duty: a quick, simple solution to this perennial problem for Kantian ethics that becomes plausible once the Incorporation Thesis is given up.

9 Cf. Allison (1990), 51, 189; and Guyer (2000), 293.

10 There is a possibility that it is sometimes one's duty to incorporate certain inclinational incentives. On this point see Guyer (2000), 298-303.

I

Considering the crucial role that maxims are supposed to play in Kant's moral theory, he seems to have said much too little about them. His published texts contain only cryptic definitions of maxims, like 'subjective principles of volition,' or 'of action,' and these are not very illuminating for readers in the twenty-first century. But Kant did not need to explain maxims any more clearly to his eighteenth-century audience. They already understood what maxims are, especially those familiar with the widely read ethical theory of Christian Wolff. Wolff had laid out a theory of practical reasoning that, like much of his system, became a source of textbook material for philosophy courses in German universities — including the courses Kant taught throughout his career. So it should not be surprising to find that in his lectures Kant told his students that maxims are, as Wolff had said, major premises of the practical syllogism.¹¹ In standard form, the Wolffian practical syllogism goes: 1. X is good (maxim); 2. Doing Y would cause X; 3. Therefore, doing Y is good.¹²

Wolff was a mind-body dualist who held that practical thinking in the soul is coordinated with movements of the body through pre-established harmony. He assumed that upon drawing the final conclusion of a practical syllogism the mind acquires a motive-state, or has a moving reason (*Bewegungs-grund*) capable of explaining bodily action.¹³ So Wolffian maxims, which are thoughts or perceptions like 'X is good,' explain action in the material world through motives derived from them by syllogistic inferences. It is worth noting here that desires

11 Immanuel Kant, *Lectures on Metaphysics*, Karl Ameriks and Steve Naragon, trans. and ed. (Cambridge: Cambridge University Press 1997), 380/28:678. Maxims were called '*maioribus propositionibus syllogisorum practitorum*' in the Latin textbook Kant used in his ethics courses: Alexander Gottlieb Baumgarten, *Ethica Philosophica* (1751), §§246, 449; see Immanuel Kant, *Kant's gesammelte Schriften*, vol. 27, pt. 2, no. 1 (Berlin: Walter De Gruyter 1975), 800, 857. See also my 'Maxims in Kant's Practical Philosophy,' *Journal of the History of Philosophy* 44 (2006) 65-83.

12 Christian Wolff, *Vernünfftige Gedancken von der Menschen Thun und Lassen, zu beförderung Ihrer Glückseligkeit* (Frankfurt, 1733; facsimile, Hildesheim: Georg Olms 1976), §§190, 400.

13 Christian Wolff, *Vernünfftige Gedancken von Gott, der Welt und der Seele des Menschen, Auch Allen Dingen Überhaupt* (Halle, 1751; facsimile, Hildesheim: Georg Olms 1983). 'Das erste Register' (glossary) of this text equates German *Bewegungs-grund* with Latin *Motivum*.

and even emotions can serve as maxims in Wolff's system; for these, he held, are just different ways of representing objects as good.¹⁴

Kant's several examples of maxims do not seem to fit the Wolffian paradigm, however.¹⁵ They never state that something 'is good'; nor are they simple statements, as in Wolff's examples. Yet if the formal predicate of every maxim is the same, then for an informed audience it need not be included in examples. When Kant asked his readers to imagine someone whose maxim is 'to increase my wealth by every safe means,' he could have expected them to imagine someone thinking this policy of action 'is good.' Regarding the relative simplicity of Wolffian maxims, it need only be observed that practical thinking is often a complicated process, which means that if maxims are principles of practical thinking they may be complicated as well. Kant's most complicated example of a maxim, found in *Groundwork of the Metaphysics of Morals*, appears to be this one: 'From self-love I make it my principle to shorten my life when its longer duration threatens more troubles than agreeableness.'¹⁶ This is a complicated example because it is actually a maxim of maxim-making. Expressed in standard syllogistic form, Kant's example here seems to be the conclusion of a sorites containing several maxims:

1. Self-love is good. (Maxim I)¹⁷
2. Minimizing life's troubles is required by self-love.
3. So, [from self-love I make it my principle that] minimizing life's troubles is good. (Maxim II)

14 *Ibid.*, §492; emotions (*Affecten*), according to Wolff, are just exceptionally strong, sense-based desires; see also §439.

15 Kant at least once provided an example in which a principle he called a maxim seems to function like the major premise of a syllogism. 'I have, for example, made it my maxim to increase my wealth by every safe means. Now I have a *deposit* in my hands, the owner of which has died and left no record of it. This is, naturally, a case for my maxim.' Immanuel Kant, *Critique of Practical Reason*, Mary Gregor, trans. and ed. (Cambridge: Cambridge University Press 1997), 25/5:27. Subsequent references to *Practical Reason* refer to this edition.

16 Immanuel Kant, *Groundwork of the Metaphysics of Morals*, Mary Gregor, trans. and ed. (Cambridge: Cambridge University Press 1998), 32/4:422

17 It would be more accurate to say here, 'My happiness is good.' But because Kant expressed his example of the suicide maxim in terms of 'self-love,' I have followed suit. It will become evident below that the subject term of a maxim can be expressed for certain purposes *objectively*, referring to an object or state of affairs like 'my happiness,' and for other purposes *subjectively*, referring to a motivational state or disposition like 'self-love.'

4. Shortening my life when its longer duration threatens more troubles than agreeableness will minimize my life's troubles.
5. So, [from self-love I make it my principle that] shortening my life when its longer duration threatens more troubles than agreeableness is good. (Maxim III)¹⁸

It should be easy to see how this line of reasoning could continue, with the addition of a few more premises, to terminate in a motivating practical conclusion. It would state that some particular suicidal act, say, jumping off this bridge, now, 'is good'; and then the action could be expected to follow.¹⁹ So even the most complicated of Kant's examples of maxims can be adapted to fit the Wolffian model; and there is an additional reason for thinking that Kant's examples of maxims are supposed to serve in practical thinking and motivation in roughly the way Wolff explained. It is that Kant was not shy about expressing his disagreements with Wolff's philosophy whenever the opportunity arose. But he never took the trouble to disagree with Wolff's widely accepted views on the structure of practical reasoning. Nor, most importantly, did he ever provide his audience with an explicit, alternative explanation of the form and function of maxims.

So if maxims can be understood to serve in this way as major premises of practical reasoning, then what are *incentives*? And how would a maxim of the form 'X is good' incorporate an incentive? An incentive, as Kant's word '*Triebfeder*' suggests, is something driving, something powered with psychological force. But it has been hard to see how something like an inciting power with strength or weakness can be 'incorporated' into a principle of reasoning with logical properties like generality or consistency. To understand incentive-incorporation it is necessary to recognize that Kant used 'incentive' equivocally. Sometimes he used it to refer to an object of desire, like 'happiness,' and other times to a motive, like 'self-love.' Within the short space of a few pages in his *Critique of Practical Reason*, for example, Kant referred to

18 Kant probably compressed everything into just the one maxim of his example in order to demonstrate the supposed practical contradiction of suicide from self-love; see *Groundwork*, (1998), 31-32/4:421-23.

19 It is important to bear in mind that the final, motivating conclusion of practical reasoning is not a maxim. Maxims, as major premises, can generate motivation only by our subsuming particular information under them: in this case, information about the effects of performing a particular action, here and now. Wolff and Kant used terms like '*Bewegungs-Gründe*' or '*Beweg-Ursachen*' for the motivating conclusions derived from maxims in practical reasoning.

both the objective moral law, and the subjective attitude of respect for the law, as *the sole moral incentive* (see 62/5:72 and 67/5:78). So there he either carelessly contradicted himself, or he presumed that one and the same 'incentive' can be both an object of thought or desire, as well as a subjective state.²⁰ Incentives seem therefore to play two important, and importantly different roles in Kant's theory of action. Considered objectively, they serve as objects or ends toward which our actions are directed. Considered subjectively, they are relatively stronger or weaker motive forces to actions as means to achieve those ends. Since happiness is an end of every human being it is an objective incentive of action incorporated into our most fundamental maxim; but when we pursue happiness we moved to act by our subjective incentive of self-love.

To 'incorporate' an incentive into a maxim then, is just to think of some object as good, including it as a subject-term in a principle of practical reasoning. This is actually what we do whenever we desire any object, Kant believed.²¹ For to desire an object is to represent it as good, in a maxim. When we next recognize some action as capable of causing the desired object's existence, we will be led, through an inference in practical reasoning, to desire to do that action. And that action, if chosen, will be motivated by the subjective incentive (or desire) for the objective incentive (or object) incorporated into the maxim. But it should not be supposed that practical reasoning like this goes on against a psychological background of forceful desires or incentives lobbying to be incorporated into maxims. For all the incentives we have, objectively considered, are objects of desire already incorporated into our maxims. And considered subjectively, as sources of motivation, they all serve as driving forces behind our choices and actions.

Someone deriving motivation for a suicidal act, as in the example above, might at the last minute shrink from an opportunity to do away with himself. Animal fear might block his reasoned determination to end it all, and save his life. But not as an affective gate-crasher; not as a disruptive emotion derailing a deliberate train of practical reason-

20 Kant compared objective and subjective aspects of the concept of 'incentive' in at least two other places: one in a passage quoted at the end of this section, where the moral 'good' is said to provide an objectively 'irresistible' incentive that is also subjectively 'weaker' in comparison with inclination; and another in a lecture example, where a person's choice between two incentive objects (two ducats) is determined by different subjective incentives (*Lectures on Metaphysics*, 268/29:902).

21 Kant agreed with scholastic thinkers, and with Wolff, that we desire nothing without thinking it good. He endorsed the slogan, '*Nihil appetimus nisi sub ratione boni*'; but in doing so he stressed that '*boni*' is ambiguous, between natural 'well-being' and rational, 'moral good' (*Practical Reason*, 51-52/5:59-60).

ing. Fear can be represented as incorporated into a maxim just as easily as self-love. The moment we recognize some object or condition as threatening and feel afraid, we have in just that recognition and feeling a maxim representing avoiding it as good. Then in judging what to do to avoid that fearful object we recognize an action as falling under our maxim of fear, and derive a motive to do it. Fear's incorporation into a maxim of the agent on the verge of suicide can be expressed in a practical syllogism as follows: 1. Avoiding injury and death is good (maxim of fear); 2. By stepping back from this ledge I will avoid injury and death (judgment); 3. So, stepping back now is good (motive). This line of practical reasoning can run through someone's mind roughly concurrent with the reasoning supporting the suicidal act, setting up a momentous and mortal choice. And what will decide the matter, in the end, are the relative subjective forces of each maxim's incorporated incentive.

Supporters of the Incorporation Thesis often tell us that Kant did not hold an empiricist, 'conflict-of-forces' theory of agency, where our choices are supposed to be causally determined by the relative strengths of competing incentives like self-love or duty. But the model of practical reasoning outlined here has it, on the contrary, that practical conflicts arising from maxims incorporating incentives of various degrees of strength are resolved when the subjectively strongest incentive prevails, through its strength, causing action on the maxim incorporating it. Reasoning with the practical syllogism guides us in acting by connecting possible actions with our ends, as means. But which of the concurrently recognized means to our various ends we finally choose is a function of the dynamic, subjective dimension of the incentives incorporated into our maxims.²² Here, for example, is how Kant once described what happens in a case of failure to act morally, through weakness of will, where one incentive incorporated into a moral maxim loses a contest of strength with a competing maxim of inclination:

[I]n the complaint of an Apostle, 'What I would, that I do not!' i.e. I incorporate the good (the law) into the maxim of my power of choice; but this good, which is an irresistible incentive objectively or ideally (*in thesi*), is subjectively (*in hypothesi*) the weaker (in comparison with inclination) whenever the maxim is to be followed. (*Religion*, 53/6:29)

22 Maxims and the conclusions that follow from them through practical reasoning evaluate objects (ends) and ways of acting (means) simply as 'good.' But comparative maxims expressing evaluations of 'better' and 'best' are possible also. An example would be: 'Confronting students I suspect of having cheated is better than avoiding it because I don't like it'; or, simply, 'Confronting students I suspect of having cheated is always best.'

Kant indicates here that weakness in the subjective force of the moral incentive explains why action on a competing maxim of inclination was chosen instead. But for supporters of the Incorporation Thesis, who take that Thesis as implying that stronger incentive forces never causally determine a free agent's choice to act, this counter-example provided by Kant himself raises a difficult interpretive problem.²³ It is worth noting here also that Kant wrote this example fewer than five pages beyond the sentence in which he expressed the incorporation requirement.

II

If the preceding interpretation of Kant's theory of practical reasoning with maxims incorporating incentives has been clearly explained, then it should seem like little more than a translation of the Humean, 'belief-desire' theory into a more clumsy vocabulary. For this reason, many will see it also as having ignored virtually all that is thought to be interesting and important about Kant's theory of rational agency. It has been explained that we incorporate an incentive into a maxim merely by desiring something, and that incorporated incentives move us to act one way or another by their own strength, without our first having to choose to be moved by them, or without our first having to endorse them as our reasons for action. In the eyes of many, no doubt, this model of human 'agency' presents us as living puppets animated by the determining forces of natural desires or inclinations, in competition with whatever incentive force the moral law of practical reason might muster against them. Any observer well enough acquainted with the causal forces by which these incentives operate could therefore predict our every step. Yet if this result seems antithetical to Kant's theory of free agency, then consider that it was Kant himself who wrote that 'all of the actions of the human being in appearance are determined

23 Allison's response to this evident counterexample to the Incorporation Thesis has been to interpret it as illustrating a case of self-deception. But he has not adequately explained what deception, what mis-believed proposition, is supposed to be illustrated by Kant's use of the example of the Apostle's lament. One explanation he gives has it that the Apostle is deceiving himself into thinking that he is not responsible for his moral transgression, owing to moral weakness (*Kant's Theory of Freedom*, 159). But this reading is entirely contrary to the spirit of Romans 7, where, as Kant presumably knew, St. Paul calls himself 'wretched' in a fit of self-reproach. He is not there attempting to exonerate himself on account of his weakness, as Allison seems to suggest. For more on this point see my 'Moral Weakness as Self-Deception,' in Hoke Robinson, ed., *Proceedings of the Eighth International Kant Congress*, (Milwaukee, WI: Marquette University Press 1995), 587-93.

in accord with the order of nature by his empirical character and the other cooperating causes;’ so that if we could know enough about these causes, ‘there would be no human action that we could not predict with certainty, and recognize as necessary given its preceding conditions.’²⁴ Bearing this statement in mind, let us turn now to consider the context where Kant expressed his incorporation requirement. We see in this section how the wider argument in which the requirement appears can instruct us in its meaning.

No sentence of Kant's *Religion* is quoted more frequently than the sentence expressing the incorporation requirement. It enjoys this distinction because it is cited so often as the textual basis for the Incorporation Thesis. Yet the beginning of the sentence is almost never quoted. The complete sentence containing Kant's incorporation requirement is:

On the rigorist's criteria, the answer to the question just posed is based on the morally important observation that [quotations usually begin here:] freedom of the power of choice has the characteristic, entirely peculiar to it, that it cannot be determined to action through any incentive *except so far as the human being has incorporated it into his maxim* (has made it into a universal rule for himself, according to which he wills to conduct himself); only in this way can an incentive, whatever it may be, coexist with the absolute spontaneity of the power of choice (freedom). (48-49/6:23-24)

From the complete sentence Kant wrote we can see that understanding the ‘rigorist's criteria’ mentioned at the beginning, and the just-posed question referred to, should be crucial for understanding his incorporation requirement. At this point in the text of *Religion* Kant was attempting to demonstrate why human beings are evil by nature, or tainted with what he called ‘radical evil.’ A central premise offered for this conclusion is a doctrine he called ‘rigorism,’ which holds that there is no intermediate position between moral good and evil. The rigorist judges by the criterion that a moral agent who is not wholly good must be wholly evil. But to this criterion, as Kant realized, some would object that it poses a false dilemma, ignoring the possibility of a person's being partly good and partly evil. So the question posed in the text just preceding the incorporation requirement is, roughly, why is there no alternative to being by nature either wholly good or wholly evil? This is the question Kant's incorporation requirement was intended to provide a basis for answering.

24 *Pure Reason*, A549/B577. See also *Practical Reason*, 83/5:99, where Kant compares human behavior to predictable lunar and solar eclipses.

But before showing how the requirement answers this question we must concentrate a bit more on its context. Kant's doctrine of radical evil is his version of the Christian doctrine of original sin. It accounts for the fact that each of us has already failed at living a morally blameless life, and the fact that hardly any of us is likely to become more than mildly interested in the idea. For this reason none of us can achieve moral perfection or holiness; so the best any of us ever can do is to continue struggling to become morally better.²⁵ What is important to understand about our original or radical evil is that since it is evil we must be culpable for it. So unlike Christianity's tracing the original sin of humanity to the primordial sin of Adam, Kant indicted each human being for his or her own moral shortcomings. Each of us is our own Adam, as Kant saw it, having corrupted our own nature by our own primordial misdeed. By 'nature' here Kant meant: 'the subjective ground — wherever it may lie — of the exercise of the human being's freedom in general (...) antecedent to every deed that falls within the scope of the senses' (46/6:21, emphasis added; cf. also 64/6:42-43).

So from the context in which Kant expressed the incorporation requirement we can see that what he meant by it is this: that whether we are good or evil by nature depends upon a free choice antecedent to our moral experience. In that primordial act that would determine once and for all the moral quality of a human being's nature, someone who would incorporate the moral incentive into her maxim would acquire a nature that is morally good. But anyone who would not do this would be evil by nature. There is no possibility of having a nature that is partly good and partly evil, because in order to have a good nature at all a person must incorporate the moral incentive into what the incorporation requirement calls 'a *universal* rule for himself, according to which he wills to conduct himself' throughout his life, which would preclude any evil in his nature. That, therefore, is Kant's answer to the just-posed question about the rigorist's criterion. And from this rigoristic doctrine it follows easily that we are all evil by nature as soon as we add the simple, empirical observation that human beings are morally imperfect.²⁶ Any transgression of the moral law within the course of a person's moral experience shows that in his primordial free choice he has not incorporated the moral law as the universal rule for his nature.

25 See Kant's comment on the 'road of endless progress toward holiness' (*Religion*, 67/6:47).

26 '[A]ccording to the cognition we have of the human being from experience, he cannot be judged otherwise' (*ibid.*, 56/6:32).

What its context shows, therefore, is that the incorporation requirement is about the incorporation of incentives in a unique free choice, antecedent to experience, that determines the empirical nature or moral character of each human being.

III

At this point it will help us better understand the nature and levels of choosing in Kant's overall theory of action if we can develop, for purposes of illustration, an admittedly incredible, 'hydraulic' model of practical life. Revising Plato and anticipating Freud, Kant traced all human motivation up to its headwaters in three pure springs he called 'predispositions.' The first predisposition is *animality*, the drive for self-preservation and species propagation. The second is *humanity*, or rational self-love, seeking at least equality in comparison with others, but preferring superiority. Third is *personality*, or the incentive of the rational moral law (*Religion*, 50-52/6:26-28). We may imagine that in each human being the waters of these ever-replenished springs acquire their variable forces as motives by cascading down a mountainside of discursive channels and switchbacks of practical reasoning; and that each person's actions in experience are just so many manifestations of the pressures of these winding currents. Differences in human characters are therefore due in great measure to the variations in the incline of the terrain down which these waters fall. For any chosen action it could, in principle, be shown how its motivation has descended though a zig-zagging sequence of maxims and syllogisms from one or another of the originally incorporated springs. If one knew enough about 'hydro-psychology' and the inclinations of the terrain forming anyone's character, it would be possible to predict how that person would choose to act in any circumstance.

The predictability of human choices in this model of agency may seem to imply that they are never free, because here actions are determined by sheer force, and not by free choices incorporating incentives into maxims. But Kant's incorporation requirement is not about choices in experience he thought of as always predictable in principle. It is instead about an original choice incorporating the fundamental springs of human motivation as incentives into a single, most fundamental maxim governing the whole of a person's empirical life. If both the headwaters of motivation and the inclines down which they tumble to determine empirical choices through contests of conflicting forces are themselves freely chosen by each human being, then the predictability of the choices they necessitate in experience remains compatible with human freedom. The three original sources from which all

human motivation stems can be reduced to two basic incentives, since the predispositions of animality and humanity are different forms of self-love. The radical evil of human nature that Kant attempted to explain through rigorism and the incorporation requirement is traced to an original act of incorporation by each human being in the following passage: 'He indeed incorporates the moral law into those [original, highest] maxims, together with the law of self-love; since, however, he realizes that ... one must be subordinated to the other as its supreme condition, he makes the incentives of self-love and their inclinations the condition of compliance with the moral law' (59/6:36). The result is that, in moral experience, whether the moral incentive will prevail over inclinations of self-love in any particular choice becomes a matter for conflicting psychological forces to determine. The degree or extent to which the motive force of the moral incentive prevails over self-love throughout the course of a person's life can be traced also to the degree of the original, evil subordination. In this choice from the original position of human agency we are free, moreover, because as a so-called 'noumenal act' our choice is timeless, and therefore unprecedented by any prior event that could contain a determining cause.²⁷

So as Kant saw it, human beings with freedom of choice can be presumed to have incorporated one or another incentive into their highest maxims, as the incorporation requirement indicates. But the 'antecedent-to-experience' free choice in which we do this takes place outside of time, in another world. The everyday choices we make in this world, by contrast, are just so many consequences (appearances) of that otherworldly free choice, together with the natural conditions in which we find ourselves. Each choice we make in this world is causally determined by some incentive of sense-based desire or inclination, or by the moral incentive, depending upon the disposition of their respective forces at the time. But this does not mean that practical reasoning has no role to play in our causally determined empirical conduct. For we always act on major-premise maxims incorporating incentives of various degrees of motive force; and these incorporated forces determine our conduct only in conjunction with minor-premise judgments about actions.

27 Kant anticipated that some would ask about the possible 'reasons' or maxims behind the original choice. His response was that in the regressive series of maxims behind our particular empirical choices we must come ultimately to a noumenal end. Though we cannot avoid asking about the antecedents of the final choice in the regress, the fact is that we can have no insight into this. Kant would say that the same is true about, for example, why human beings experience the world in three-dimensional space. See *Religion*, 50/6:25, and *Pure Reason*, A556-7/B584-5.

Today, however, there are a number of well known, critical responses to this line of thinking about Kant's view of human freedom. Even his most sympathetic commentators express misgivings about interpreting his views in this way, though some agree that it may provide the most plausible reading of some of his more obscure comments on human freedom.²⁸ Few interpreters today seem satisfied with the prospect that a timeless noumenal action, in another world, is somehow supposed to make a difference for our actions in this world. Yet this is a prospect I have explained as essential for understanding Kant's incorporation requirement. We have space here for briefly considering only three of the main objections usually raised against this interpretation of Kant's practical philosophy, where the 'two worlds' of noumena and phenomena are understood as presenting an ontological or metaphysical distinction.²⁹ Two standard objections will be considered in the remainder of this section, and the third occupies our attention later, in Section IV.

The first type of objection is voiced by Jonathan Bennett, who objected to the intelligibility of Kant's idea of timeless noumenal action when he wrote: 'When Kant says of a noumenon that 'nothing happens in it' and yet that [a free noumenon] 'of itself begins its effects in the sensible world' (B569), he implies that there is a making-to-begin which is not a happening; and I cannot understand that as anything but a contradiction.'³⁰ A related problem worried Ralf Meerbote. 'According

28 Jonathan Bennett, for example, called a version of this interpretation presented by Allen Wood 'flawless.' See Jonathan Bennett, 'Kant's Theory of Freedom,' in Allen W. Wood, ed., *Self and Nature in Kant's Philosophy*, 102-12, 102. See also Eric Watkins, *Kant and the Metaphysics of Causality* (Cambridge: Cambridge University Press 2005), 301-61.

29 The more popular alternative these days interprets Kant's noumenon-phenomenon distinction as merely methodological or epistemological: as a 'two-aspect' theory, rather than a 'two-worlds' theory. Cf. Henry E. Allison, 'Transcendental Idealism: The "Two Aspect" View,' in Bernard den Ouden and Marcia Moen, eds., *New Essays on Kant* (New York: Peter Lang 1987), 155-78.

30 Bennett, 102. Cf. also Meerbote: 'The notion of [a timeless free choice] is intrinsically puzzling, since it is the notion of something coming about in an agent, but timelessly so. We normally view any coming-about as temporal, and we normally take any reason from which an agent acts to obtain temporally before or during his (temporal) performance.' Ralf Meerbote, 'Which Freedom?' in Predrag Cicvacki, ed., *Kant's Legacy: Essays in Honor of Lewis White Beck* (Rochester, NY: University of Rochester Press 2001), 197-225, 207. Kemp Smith raised a similar objection to Kant's resolution of the Third Antinomy: 'A solution is rendered impossible by the very terms in which he formulates the problem. If the spiritual and the natural be opposed to one another as the timeless and the temporal, and if the natural be further viewed as a unitary system, individual moral freedom is no longer defensible.' Norman Kemp Smith, *A Commentary to 'Kant's Critique of Pure Reason'* (New

to the Critical Kant,' he wrote, 'it is a necessary condition of the empirical knowability of temporal, natural events that whatever causes them be itself temporal. Since some of the effects of [any exercise of human free agency] are temporal and empirically knowable, this agency must then itself be temporal.'³¹ Each of these critics would deny that it makes sense to think of a free choice apart from space and time as making a difference for human conduct in the empirical world.

But what Meerbote's argument overlooks is the prospect that the empirically knowable *causal relation* between human actions and their psychological antecedents can be an effect of timeless, noumenal agency. In Kant's view of causation, every effect-object is altered by a cause-object according to a rule of the former's empirical character: the way water is affected by very cold air depends on what water, as a rule, is like. Causal rules make it necessary for the effect-object to respond always in the same way to the same action of the cause-object. So it can be a natural fact about myself as an appearance in the temporal world that, owing to a rule of my empirical character, my desire for Φ causes me to Ψ . But it could also be that the rule for this relation between my desire for Φ and my action of Ψ -ing is an effect of my timeless free agency. It could be, in other words, that something I do timelessly in the noumenal world causes the appearance or applicability in the phenomenal world of the rule governing my Ψ -ing as an effect of my prior desire for Φ . Hence, as required by the 'Critical' Kant, my empirical act of Ψ -ing depends always upon a knowable, temporal cause, which in this case it is my desire for Φ . But it also depends upon my free, noumenal act that supplies the rule for the way I am affected by this cause.

What Bennett found incomprehensible is how, by acting in the noumenal world, I would establish this causal rule partly constitutive of my character in the phenomenal world. His question would be: How can I, as a noumenon, begin this temporal 'appearing' in the phenomenal world if nothing can 'happen' in or to a noumenon? The answer depends on Kant's concept of a 'world.' If there is a noumenal world, and Kant's practical philosophy gives us ground for thinking there is, then that world must be a community of active substances, with each one making a difference for all. This interactivity is the only condition under which a plurality of entities can be united into a whole called a

York: Humanities Press 1962), 518. Cf. also H.J. Paton, *The Categorical Imperative*, 274.

31 Meerbote, 208. Here Meerbote credits and cites Hud Hudson, *Kant's Compatibilism* (Ithaca, NY: Cornell University Press 1994), 25.

'world.'³² So according to Kant's rational cosmology, if I am a member of a noumenal world, then everything else in that world must be in some way different because I am a part of that whole. In other words, I must have world-wide effects just in virtue of my membership in the noumenal world community; and all other members of that community must likewise affect me in some way.³³ So a timeless noumenon acts by making a difference in its world, without beginning or happening to act in time. And since its effects in that world can be ordered as appearances in the time series of the empirical world, it is possible to understand how as a noumenon in which nothing 'happens' I can begin effects in time. The beginning of my Ψ -ing in this world, as an effect of my desire for Φ , is a temporal appearance of some difference I make through my membership in a timeless community of noumenal beings.

The second critical strain giving interpreters doubts about this 'noumenalist' reading of the incorporation requirement has been pointedly expressed by Ralph Walker, and seconded by Bennett, among others. It is that understanding one's character in the phenomenal world as chosen from a position outside of time should make one morally responsible for all of the causes of one's character in the empirical world, and even all of *their* causes. '[M]y noumenal self must have freely chosen the entire causal series that makes up the phenomenal world,' Walker concludes. As a consequence, 'my responsibility extends far beyond my own character: I can be blamed for the First World War, and for the Lisbon earthquake that so appalled Voltaire.'³⁴ But Allen Wood once responded to this line of objection by reminding us that in his theory of freedom Kant intended to establish nothing more than that we *could be* free, even while every event has a cause. He did not intend to establish that we can know we are free; nor did he attempt to determine exactly what we do as free agents that results in our having the character we have in the empirical world. Accordingly, Wood argues, *it remains at*

32 'The form of the world is a real connection because it is a real whole. For if we have a multitude of substances, then these must also stand together in connection, otherwise they would be isolated. Isolated substances, however, never constitute a whole' (*Lectures on Metaphysics*, 208/29:851).

33 '[B]ecause the substances in the world stand in interaction [*commercio*] and find themselves in action and reaction, then each, or rather its state, is dependent upon the action of another' (*ibid.*, 210/19: 852, cf. also 20/28:196). Kant used the ideas of interaction, community, or reciprocal determination in *Pure Reason's* third analogy of experience to account for co-existence of objects in the phenomenal world (A211-15/B256-262).

34 Ralph C.S. Walker, *Kant: The Arguments of the Philosophers* (London: Routledge & Kegan Paul 1978), 149; cf. also Bennett, 102-4.

least possible that because of whatever we do, noumenally, we are each morally responsible for just those phenomenal deeds for which we normally hold ourselves responsible, and that is all Kant would have wanted to prove.³⁵

A better reply to Walker's objection is available, however. It is that the objection's crucial error lies in equating the empirical character of a moral agent with an *event* requiring an empirical cause. Kant's notion of empirical character is different; it is the notion of a law, or set of laws, of empirical causality: 'But every effective cause [like each of us] must have a *character*, i.e., a law of its causality ... through which its actions, as appearances, would stand through and through in connection with other appearances in accordance with constant natural laws' (A539/B567). In Kant's understanding of the phenomenal world there would be no empirical cause of any empirical character; neither in psychology nor in physics.³⁶ The laws of nature cannot have natural causes; at least not all of them. For this reason each of us can be morally responsible, through a noumenal choice, for the set of laws that constitutes his or her empirical character. And this need not make us responsible for any prenatal, empirical causes of our characters.

Another, related objection can now be anticipated. It will likely be pointed out that there certainly are character-shaping events in our lives — education, good or bad; adversity; fortune — and these life-determining events have prior causes, which reach back to even more ancient causes. So Kant surely must admit, it will be said, that events antecedent to our births are causally responsible for shaping our empirical characters. Hence, we cannot be morally responsible for our characters without being morally responsible for the events that shape them, and so for all prior causes of those events as well. But the simple rejoinder to this objection is a point already made: that an effect-object always responds to its cause 'in character.' How life-changing events

35 See Wood, 'Kant's Compatibilism,' 92.

36 Kant may seem to have said otherwise at one point in the first *Critique*, where empirical character seems to be called an effect of other appearances: 'Because this empirical character itself must be drawn from appearances *as effect*, and from the rule which experience provides... .' (A549/B577, emphasis added). But it makes no sense that a law of empirical causality, which is how empirical character is defined, should be derived or inferred *as an effect* of appearances. Kant wrote here: '*dieser empirische Charakter selbst aus den Erscheinungen als Wirkung und aus der Regel derselben, welche Erfahrung an die Hand giebt, gezogen werden muß...*' Kemp Smith's translation reads: 'this empirical character must itself be discovered from *the appearances which are its effect* and from the rule to which experience shows them to conform... .' (emphasis added). Immanuel Kant, *Critique of Pure Reason*, trans. Norman Kemp Smith (New York: St. Martin's Press 1965).

shape our characters always depends upon the shape of our characters at the time of their occurrence. The causes of the events that shape my character are not my responsibility. What I am responsible for is the way those events shape my character.

IV

A third line of resistance to interpreting Kant's incorporation requirement as referring to a timeless noumenal choice will emphasize the slogan: "'Ought' implies 'can'." If all my empirical actions are causally necessitated though prior events and motive forces, then whenever I do not act as I ought, I could not have done otherwise. So how does this noumenalist interpretation of Kant's incorporation requirement preserve his fundamental doctrine that I always can do what I ought to do?

It will probably not satisfy many to say, in response to such a question, that by a timeless free act I ought to have, and so could have, chosen a different empirical character. This would be much of what Kant's theory would say, however. This answer may not be fully satisfying because it seems to imply that I am burdened today, as well as yesterday and tomorrow, with the consequences of a bad noumenal choice. It seems to say that I am not *now* always free to do what I recognize I ought morally to do, even if I am the one who has freely determined how I shall respond to that recognition. Kant's theory of noumenal freedom may tell us how we can be morally responsible for what we are causally determined to do. But it cannot, the complaint will go, give us control over our moral destiny. It implies a fatalism as depressing as any theory of natural determinism; actually, worse. In the empirical world where we live and act, we are only conscripts of our self-determined fate, not its captains.³⁷ Yet we do not even have the determinist's advantage of attributing our faults to alien causes!

Wood attempted to defend Kant's theory of freedom from criticism along these lines also, pointing out the fallacies it involves. One fallacy is to think of one's character as something chosen far in the past, perhaps at the beginning of time, thereby predetermining one's entire life in the empirical world as a *fait accompli*. The mistake here lies in

37 Cf., *Lectures on Metaphysics*, 364/28:663, where Kant admits that human beings are subject to 'fate' at least in the empirical world, although not to blind chance.

not recognizing that a timeless choice cannot *predetermine* anything.³⁸ According to Wood, 'Kant's theory says that our timeless choice does not predetermine our actions but has its influence immediately on each of them and should be considered simultaneous with each action as it occurs in the temporal order.'³⁹ Wood recognized that some may balk at the notion of one timeless choice that is simultaneous with many choices made at different times, but he pointed out that this problematic prospect 'is built into the very notion of a timeless being exerting influence on temporal events.' If we accept Kant's theory of timeless free agency, then the simultaneity of one timeless choice with many choices in a temporal series introduces no additional problem, as Wood saw it.

But here again it seems that we can give a better explanation. It will help if we abandon all expressions suggesting that temporal events are in some way *influenced* by noumenal reality, in order to avoid implying that noumenal agents can act *at the time* of temporal events. It is not that a noumenal being can periodically exert influence sufficient to alter the course of natural history. Noumenal action is not a phenomenal miracle. The point is rather that the totality of the actions of members of the timeless noumenal community unrolls as a series of empirical events throughout the course of time. My going to the post office at three-thirty this afternoon was not something scheduled as long ago as the beginning of the phenomenal world. It is rather that who I am in time, who I have always been, and will always be, is the character going to my local post office at three-thirty on the twenty-eighth of the month in which I wrote this, just as Hamlet is always the one in the play who thinks aloud, 'To be, or not to be.' If Nietzsche is right about recurrence, then somewhat like the sequence of language and action that is a production of *Hamlet*, the play of the phenomenal world unrolls over and over again. My role on the empirical stage is to perform the sequence of actions I am responsible for having generated through my contribution to the noumenal world. I could have determined a different sequence of actions for my phenomenal role, since by the terms of my noumenal freedom nothing prior could have causally necessitated the contribution to the noumenal world that I make. Yet to wish that my role in the empirical world-play were different seems as senseless as Hamlet's considering aloud whether 'To be Horatio, or not to be Horatio.' To lament empirical fatalism, or to wish that my actions in the phenomenal world were bound to no temporal script is as senseless as Hamlet's

38 Kant is careful to embrace determinism and reject *predeterminism*; see *ibid.*, 488-89/29:1020-21.

39 Wood, 'Kant's Compatibilism,' 96

wishing he could walk about onstage saying whatever he likes. Just as the character of Hamlet has no idea that he is acting on a stage, much less what he might like to say out of character, so I cannot understand what it would be like to act independently of the temporal forces that move me to act, much less how I might choose to act if I were free to deviate from the determining laws of my character.

But Kant's moral theory tells us that we ought to be better than we have been, and so we can be. Doesn't this mean that we can alter at least the moral dimension of the character we play on the phenomenal stage? Wood's response to a question like this was: 'Presumably, [Kant's] theory is that for every imaginable course of conduct in the phenomenal world, there is a timeless choice ... that would yield that course of conduct. Hence there are some such choices whose results in the world of appearance involve *changes in empirical character*, drastic conversions from evil to good, or sudden degenerations from good to evil' (94, emphasis added). Wood is correct here, but we must be careful to avoid the fallacy of thinking that owing to some dissatisfaction with the way things are unfolding in time, the noumenal person outside of time can choose to effect an instantaneous change in the character of the phenomenal person. Personal moral progress in Kant's theory implies a change of character from bad to better, but this can be consistent with the idea that empirical character is the appearance of a timeless, invariant noumenon. We need only stop to consider the perplexities to be encountered in the very idea of empirical character. To illustrate, let us adapt our earlier, falling waters analogy. A cup of water dropped from a height of one meter will splash. A cup of solid ice dropped in the same way will bounce. Here, apparently, are two physical objects whose behaviors differ in the same circumstances because of the invariant laws of their different characters. But the cup of water can undergo a change at 0°C and take on the character of ice; and we know that water's changing to ice at that temperature is, so to speak, in character. So does the water's change from being something whose character determines it to splash when dropped, to something whose character determines it to bounce, constitute a *change of character*? In one sense of character, yes; in another sense, no.

The same equivocation applies in human psychology, only the field is much more complicated. What must always be acknowledged here, unlike in the lifeless physical world, is that the human being's reflection upon of the laws of her character can make a difference in the laws of her character. Self-consciousness alters the conscious self.⁴⁰ So if some-

40 Kant suggested this in part of his famous argument against the possibility of a

one should come to recognize the moral shortcomings of her character, that recognition itself may change her character. And whether it changes her character for the better or makes it worse must depend, of course, on her character: on what she is like at that time. It could therefore be said that someone's change of character for the better was wholly in character, provided an equivocal use of 'character' is allowed. But if we speak about character in this way it becomes impossible for any finite temporal observer to encompass the complete character of a person. The elusive totality of appearances of someone's character is what Kant called, in the *Critique of Pure Reason*, the 'intelligible character' of the human being. It is also what he thought we are right to consider as identical with an original, noumenal act, which can be represented for practical purposes as a timeless choice of how one's empirical character develops in time.⁴¹ Kant assumed we have sufficient empirical evidence for saying that the moral character of each human being is evil, or morally imperfect. The moral command that we *ought* to become good, or to become at least better, is therefore addressed to becoming, empirical agents who *can*. But whether and how far moral agents considered empirically can be expected to improve in response to that command is entirely up to them as agents considered noumenally. Nothing empirical can ever justify the judgment that a moral agent recognizing that she ought to become better cannot do so.

V

We now turn finally, and briefly, to the second problem mentioned in the beginning: the problem related to 'moral worth.' Recent interpreters sympathetic to Kant have responded to that problem by introducing new perspectives on his theory of motivation. But the incorporation requirement has been misapplied in this arena also, in the form of the Incorporation Thesis.⁴² Part of what I want to show here, before closing, is how Kant's wider ethical theory provides resources for addressing

science of empirical psychology: 'even observation by itself already changes and displaces the state of the observed object.' See Immanuel Kant, *Metaphysical Foundations of Natural Science*, in *Theoretical Philosophy after 1781*, Henry Allison and Peter Heath, eds., Michael Friedman, trans. (Cambridge: Cambridge University Press 2002), 186/4:471; and *Lectures on Metaphysics*, 381/28:679.

41 Cf. *Religion*, 55/6:31, where Kant explains the equivocal use of 'deed.'

42 See Guyer, 293-4; Wood, *Kant's Ethical Thought*, 51-53; and Marcia Baron, *Kantian Ethics, Almost without Apology* (Ithaca, NY: Cornell University Press 1995), 134-5, 188-93.

critics' complaints over the value of acting from duty. But I want to show also how the Incorporation Thesis actually poses an insurmountable obstacle to responding adequately to those complaints.

Critics have objected to Kant's theory of moral appraisal, with a great deal of plausibility, that it denies there are any morally good actions not done 'from duty alone.' They point out, as counter-examples, that it is virtuous to act as a good friend, or out of admirable emotions like love, kindness or compassion. Critics point out also that Kant's theory appraises some actions as morally worthy that are not virtuous. A husband's acting from duty alone in rescuing his wife is morally worthy in Kant's estimate. But this is not how a good husband acts, the critics say. So Kant's theory of moral appraisal seems to get it wrong in two ways: it says of some morally good actions that they lack moral worth; and it says of some actions not morally good that they have moral worth.

Today the most well known responses to these criticisms present complex motivational structures allowing for acting from both duty and inclination in the same action. Different models of motivational 'over-determination' have been developed for this purpose, often emphasizing a distinction between so-called 'primary' and 'secondary' motives. The Incorporation Thesis tells us that we are always free in choosing the incentives or motives for our actions. So some suggest that a free agent can choose to act on an admirable inclination as a primary motive, just so long as her action would somehow also exhibit a commitment to duty as a secondary motive. But others have taken a wholly different approach to the problem. In responding to the critics' complaints they have attempted instead to clarify the meaning of Kant's appraisal concept of 'moral worth.' They have found it significant that Kant introduced the idea of moral worth in the first place just in order to explicate his concept of a good will or character. To them this suggests that appraising actions as 'morally worthy' is not to praise them for their motives, but in some way to signal the good character of their agent. The first set of interpreters responding to the problem of moral worth take what can be called a motivational approach, showing how actions can be morally worthy even if not motivated by duty, or not primarily by duty.⁴³ The others approach the problem by clarifying what it means for an action to be appraised as 'morally worthy,' relating this

43 See Richard Henson, 'What Kant Might Have Said: Moral Worth and the Overdetermination of Dutiful Action,' *Philosophical Review* 88 (1979) 39-54; Barbara Herman, 'On the Value of Acting from the Motive of Duty,' in *The Practice of Moral Judgment* (Cambridge, MA: Harvard University Press 1993) 1-22; and Paul Benson, 'Moral Worth,' *Philosophical Studies* 51 (1987) 365-82.

value in some way to goodness of character.⁴⁴ Some who address the problem have also mixed these two approaches.⁴⁵

The strategy that will be recommended here falls under the second of these two approaches. It calls attention to the difference between 'moral worth' and 'virtue.' Kant's critics evidently identify the two, because they fault him for denying the moral worth of actions they consider virtuous, and for finding moral worth in actions they see as lacking virtue. But are the concepts of 'moral worth' and 'virtue' supposed to be the same? Kant believed we cannot be certain that there ever has been an action meeting the criterion for moral worth. Yet he did not take the same attitude in regard to virtue.⁴⁶ This gives us reason to think that in denying an action's moral worth a Kantian appraiser need not also deny that it could be virtuous, and that in affirming an action's moral worth she need not also affirm its virtue.⁴⁷ It helps to recall here that, as made clearest in *The Metaphysics of Morals*, Kant's moral theory divides our duties generally into two classifications: juridical duties, or 'duties of right,' and ethical duties, which include a set of 'duties of virtue.'⁴⁸ What is distinctive about the latter set of duties is that they are duties to have certain objects as our ends. That is the same as to say that they are duties to regard certain objects as good, or, as has been explained above, to desire them, or have maxims incorporating these objects as incentives.

The two ultimate ends that Kant tells us we ought to have incorporated into our maxims are our own perfection and the happiness of others. We ought, that is, to care about bettering ourselves, and about making others happy. We ought to regard these objects as goods because our do-

44 See Nelson Potter, 'Kant and the Moral Worth of Actions,' *The Southern Journal of Philosophy* 34 (1996) 225-41; Walter E. Schaller, 'Should Kantians Care about Moral Worth?' *Dialogue* 32 (1993) 25-40; and Allison, (1990), 120.

45 See Baron, (1995), 129-45; Philip Stratton-Lake, *Kant, Duty and Moral Worth* (London: Routledge 2000), 60-77; and Guyer, (2000), 287-329.

46 Kant talks about virtue as a 'phenomenon' in *Religion*, (1998), 67/6:47.

47 Kant seems often to equate 'virtue' and 'moral worth.' But he used 'virtue' in two senses, one formal and one material: one corresponding to what he called the single 'obligation of virtue,' and one corresponding to the multiple 'duties of virtue.' So acting dutifully but not *from duty* would show a shortcoming of virtue in the formal sense; but only in this sense. In the material sense of 'virtue,' it would often be appropriate to affirm or deny the virtue of an action without reference to its meeting the criterion for moral worth, just as it is appropriate to affirm or deny the rightness of an action without referring to its moral worth.

48 Roughly the same distinction is implicit in the *Groundwork's* distinction between 'perfect' and 'imperfect duties' (31n/4:421n).

ing so means that we will have subjectively strong desires/maxims that can move us to act whenever we judge that we have opportunities to pursue them. Duties of virtue are not duties *to act* in ways that improve ourselves, or in ways that make others happy. They are instead duties *to care about* what can be achieved by acting in these ways, so that the forces of these concerns can cause us to act. In Kantian ethics it is not expected that someone possessing these virtues will act *from duty* in, for example, brushing up his Spanish in order speak to new immigrants, or in visiting her sick friend in the hospital. Actions like these will be virtuous and admirable when motivated by virtuous cares, as subjective incentives incorporated into maxims. So they will not, just for that reason, have moral worth; but there is no reason to expect them to.

The expectation of Kantian morality is that we fulfill our various duties, both juridical and ethical. We are not expected to aim at acting from duty as often as possible, or at maximizing moral worth. The virtuous person's concern with the motives of her actions is for the most part only indirect. The matters of direct concern are typically what Kant called the various *material* ends that ought to be achieved. Cultivating or strengthening our concern for those ends within the economy of our interests leads to moral improvement in empirical character, and so to more frequent virtuous action. It is true of Kant's ethical theory nevertheless that our attachment to what he called the *formal* end of complying with duty for duty's sake ought to be cultivated as well. For this one strives to strengthen the incentive of respect for the moral law. Kant wrote that 'the capacity (*facultas*) to overcome all opposing sensible impulses ... as *strength* (*robur*) is something [we] must acquire; and the way to acquire it is to enhance the moral *incentive* (...), both by contemplating the dignity of the pure rational law in us (*contemplatione*) and by practicing virtue (*exercitio*)' (*Metaphysics of Morals*, 200/6:397; cf. also 201/6:399). So both material ends of morality, like self-improvement and the happiness of others, as well as the formal ends of morality, like respect for persons or for the moral law, are matters of concern for the virtuous person. But it is not easy to understand why this should be so, if we follow the Incorporation Thesis. For the absolute freedom to choose our motive for any action, as guaranteed by that Thesis, undermines the very point of enhancing or strengthening incentives. It is hard to see why we ought to strengthen our motivating concerns for moral ends if, as supporters of that Thesis so often point out, our freedom of choice means that the relative strengths of our desires do not causally determine our conduct.

VI Conclusion

When the Incorporation Thesis is set aside we are in a better position to understand Kant's theory of action, and to address familiar complaints about the Kantian theory of moral appraisal. Two principal aims of this essay have been to provide an alternative interpretation of the text widely seen as supporting the Incorporation Thesis, and to call attention to advantages to be gained by rejecting that Thesis.

Some of the interpretive strategies offered in the preceding arguments are admittedly regressive, considering recent trends in Kant interpretation. Since at least the 1970s the 'two-worlds,' ontological reading of Kant's noumenon-phenomenon distinction has been losing ground to alternative interpretations emphasizing merely perspectival or methodological versions of that distinction.⁴⁹ I have attempted here to address some of the more common critiques of the ontological version of the distinction as it applies to human freedom, but there are others that could not be discussed.⁵⁰ The 'two-worlds' distinction seems at least to provide the clearest way of understanding the compatibility of human freedom and causal determinism that Kant always insisted upon. But the main obstacle to its wider acceptance, apart from concerns over ontological dualism, seems to be the apparently intractable attachment of most interpreters to a libertarian, anything-here-and-now-goes conception of human freedom that Kant's view of moral responsibility is believed to require. Partly in order to overcome this obstacle I have attempted here to show how the ontological version of the noumenon-phenomenon distinction shifts the Kantian view of moral responsibility to one primarily about responsibility for empirical character, and only secondarily about responsibility for individual actions and their motives.

I have tried to show here also how this shift can help us better understand the Kantian theory of virtue. It is helpful to give up the idea that human freedom is an exemption from psychological determination in selecting one's motives or reasons for particular actions. For once that idea is given up, actions can be morally appraised on the basis of their motive causes, as far as these can be discerned; and character can be

49 See Karl Ameriks, 'Recent Work on Kant's Theoretical Philosophy,' *American Philosophical Quarterly* 19 (1982) 1-23. For a discussion of this issue specifically in relation to the problem of freedom and determinism, see Watkins, 317-25.

50 See, for example, Robert Paul Wolff, 'Remarks on the Relation of the Critique of Pure Reason to Kant's Ethical Theory,' in Bernard den Ouden and Marcia Moen, eds., *New Essays on Kant*, 139-53; and Lewis White Beck, *A Commentary on Kant's Critique of Practical Reason* (Chicago: University of Chicago Press 1960), 191-4.

appraised on the basis of the strength or effectiveness of its virtuous motives. This makes it easier to understand how the cultivation of character can have moral aims besides merely acting for the sake of duty alone, and how progress in acquiring Kantian virtue can be expected to make enduring, empirical differences in the person, and consequently also in the phenomenal world.

Received: November 2005

Revised: August 2006

