

The Inaugural Address

AUTONOMY: THE EMPEROR'S NEW CLOTHES

by Onora O'Neill

ABSTRACT Conceptions of *individual autonomy* and of *rational autonomy* have played large parts in twentieth century moral philosophy, yet it is hard to see how either could be basic to morality. Kant's conception of autonomy is radically different. He predicated autonomy neither of individual selves nor of processes of choosing, but of principles of action. Principles of action are *Kantianly autonomous* only if they are law-like in form and could be universal in scope; they are heteronomous if, although law-like in form, they cannot have universal scope. Puzzles about claims linking morality, reason and autonomy are greatly reduced by recognising the distinctiveness of Kantian autonomy.

I

I*ntroduction.* In the last half-century appeals to autonomy have played a larger and larger part in ethical and political debate. Yet the advocates of autonomy still disagree about what it is, and why it is important. At times it seems that they agree only that autonomy has a noble, Kantian pedigree that links it closely to morality.

They are certainly right that Kant links autonomy to morality. For example, he claims both that '*Morality* is thus the relation of actions to the autonomy of the will'¹ and that '*Autonomy* of the will is the sole principle of all moral laws and of duties in keeping with them.'² However, I believe that there is little evidence for strong links between morality and twentieth century conceptions of autonomy. Recent conceptions of autonomy have no claim to be 'the sole principle of all moral laws and of duties in keeping with them', and their claims to Kantian ancestry are greatly exaggerated. We have been admiring a naked Emperor of questionable legitimacy.³

1. Kant, 1785, 4:439.

2. Kant, 1787, 5:33.

3. Differences are often pointed out; yet the persistence of claims to Kantian ancestry suggests that many contemporary protagonists of autonomy overlook them. See Hill, 1992, who notes that current conceptions of autonomy have been 'cut loose from their Kantian roots' (77), and that 'Autonomy is a central concept in contemporary

These negative claims tell us nothing about Kantian autonomy, or about its supposed links to morality and reason. We can test the links that Kant thought he had established only by going back to the Kantian texts. I hope that a short canter through those texts will reveal a more interesting landscape than the limited terrain so energetically charted in recent decades.

II

Autonomy as Independence. I first realised quite how bewildering the array of differing conceptions of autonomy in contemporary debates had become when I read Gerald Dworkin's book, *The Theory and Practice of Autonomy*. Dworkin offers a list of about a dozen different conceptions of autonomy, which, he claims, has variously been equated with

Liberty (positive or negative) ... dignity, integrity, individuality, independence, responsibility and self-knowledge ... self-assertion ... critical reflection ... freedom from obligation ... absence of external causation...and knowledge of one's own interests.⁴

This list is far from complete. For example, Ruth Faden and Thomas Beauchamp in their interesting work *The History and Theory of Informed Consent* note that autonomy has been equated with a quite different list of concepts, including

privacy, voluntariness, self-mastery, choosing freely, choosing one's own moral position and accepting responsibility for one's choices.⁵

Dworkin contends that despite all these variations, all conceptions of autonomy share two features:

The only features that are held constant from one author to another are that autonomy is a feature of persons and that is a desirable quality to have.⁶

moral debates as well as in the discussion of Kant; but the only thing that seems completely clear about autonomy in these contexts is that it means different things to different writers' (76).

4. Dworkin, 1988, 6.

5. Faden and Beauchamp, 1986, 7.

6. Dworkin, 1988, 6.

I doubt whether either claim is true. It is certainly not true that all conceptions of autonomy view it as a feature of persons. The original use of the term *autonomy*—literally *self-legislation*—in antiquity referred to a property not of persons, but of polities. Autonomous city-states made their own laws; colonies were given laws by their mother cities. And, as we have already seen, Kant ascribes autonomy not to persons, but to the will or, more accurately, to determinations of the will or principles. Moreover, some twentieth century writers—structuralists, behaviourists—dismiss the very idea that autonomy could be a feature either of persons or of the will.

There is also no general agreement that autonomy is 'a desirable quality to have'. On the contrary, many distinguished recent writers maintain that there is something morally unsavoury about autonomy. Often they condemn autonomy as no more than a form of individual independence whose manifestations may be morally admirable, corrupt or merely trivial.

III

Rational Autonomy. Many late twentieth century protagonists of autonomy have taken this point and do not identify autonomy with *mere* independence, of the sort advocated by pop-existentialists. They often insist that autonomous action must not only be chosen (so to some extent independent), but rationally chosen. They have advocated one or another form of *rational autonomy*. Rational autonomy (unlike autonomy conceived as mere, sheer independence) might, it seems, be linked to morality.

The principal source for most conceptions of *rational autonomy* is, I think, not Kant, but John Stuart Mill's *On Liberty*. Mill explicitly repudiates the thought that mere, sheer independence or choosing is morally important. He ascribes value to intelligent and reflective choosing, by persons of well-developed *individuality* and *character*. He claims that

A person whose desires and impulses are his own—are the expression of his own nature, as it has been developed and modified by his own culture— is said to have a character. One whose desires and impulses are not his own, has no character, no more than a steam engine has a character.⁷

7. Mill, 1859, 189.

He then argues that the choosing of persons with greater individuality and character benefits each and all, and so that Utilitarians have reason to promote and protect the liberties that promote reflective and intelligent choosing, and thereby individuality and character:

In proportion to the development of his individuality each person becomes more valuable to himself, and is therefore capable of being more valuable to others.⁸

However, Mill does not call choosing that reflects individuality or character autonomous. So far as I can discover, he never speaks of the *autonomy of persons* or of *autonomous choosing*, although I have found references to the *autonomy of states*.⁹ I suspect that for Mill the term *autonomy* was a term of art that belonged either in constitutional discussions, or in Kant's non-naturalistic account of action, and had no place in his own naturalistic account of action.¹⁰

Despite this divergence in terminology, I think that twentieth century advocates of rational autonomy are close to Mill when they say that what is ethically important is not *mere choosing*, or *mere independence*, but specifically *rational choosing*, hence (some version of) *rational autonomy*.¹¹ However, unlike Mill, many recent admirers of rational autonomy are not Utilitarians. They do not view intelligent and reflective choosing, or the liberty that respects and protects it, as valuable because it is an effective means to human happiness. Some of them promote versions of 'rational autonomy' not as an instrumental but as a fundamental human good or value.

The twentieth century writers who follow Mill in claiming that some version of rational autonomy (rather than mere, sheer independence) is ethically important, also disagree about what it is.

8. Mill, 1859, 192.

9. Mill, 1862, Ch. 16.

10. Mill rejected Kant's ethics because he thought that the Categorical Imperative was not action-guiding, so concluded that ethics must be based on claims that Kant would have classified as heteronomous rather than autonomous.

11. Mill no more speaks of *rational autonomy* than he speaks of *autonomy*. But many of his interpreters use both terms in reporting his position. For example, John Skorupski writes of 'Mill in whose philosophy naturalism and ... rational autonomy are the two deepest convictions'. Skorupski, 1989, 43.

For example, Harry Frankfurt in a now classic paper¹² first published in 1970 distinguished between routine choosing and choosing that reflects *second-order desires*, and argued that capacities for the latter sort of choosing set persons apart and are morally significant. His famous example contrasts the wanton addict, who in choosing her fix is driven by mere desire, with the person who also has second-order desires to be a person with certain immediate desires (the determined addict who both wants her fix and to be an addict, the admirable addict who wants not to be addicted and struggles with her first-order desire). Autonomous choosing has been variously characterised by other writers as choosing that is well informed, or fully informed, or reflective, or reflectively endorsed, in short as choosing that, while ultimately based on desires or preferences, follows certain rational processes.

I will not linger on the many ingenious accounts that advocates of various conceptions of rational autonomy provide of the proper processes for rational choosing. I suspect that it would be hard to show that rationally autonomous choosing always (or even generally) leads to ethically superior choices. Recent conceptions of rational autonomy allow desires and preferences, as well as rational procedures for choosing, to determine what is autonomously chosen: why then should we suppose that rational autonomy secures ethical acceptability? Indeed, as opponents of autonomy often point out, hunch, tradition and intuition sometimes do better; they may reach ethically sound results, and choosing with rational autonomy may fail to do so.

On reflection it should not surprise us that protagonists of rational autonomy, whose theories of action build not only on conceptions of rationality but also on specific accounts of motivation, find that motivation duly reflected in rationally autonomous choices, often with ethically disturbing implications. Choosing that is rationally autonomous is likely to endorse whatever individuals prefer, and to veer towards egotistical choosing. At best, rationally autonomous choosing is likely to steer those with egotistical motivation towards more 'enlightened' self-interest. Many of the ethical objections raised about autonomy conceived of as mere, sheer independence then recur as objections to conceptions of rational autonomy.

12. Frankfurt, 1970.

And there is worse to come. Proponents of rational autonomy may hope to show that certain rational processes *for choosing* generally produce more *valuable choices*. But they will not be able to show even this much without independent criteria for identifying valuable choices. In Mill's hands, where the value of choices is settled by Utilitarian arguments, there is a context for arguing that some ways of choosing are more likely than others to produce valuable choices. But without an independent account of valuable choices, it is unclear how we could show that one or another way of choosing, such as those favoured by various conceptions of rational autonomy, is more valuable. Attempts to promote rational autonomy as a *fundamental* rather than an *instrumental* value need quite different sorts of support, and I am not myself at all sure where it could be found.

IV

Kantian Autonomy in Context. Both Kant's admirers and his detractors agree that Kantian autonomy is distinctive. In my view it differs markedly *both* from mere, sheer individual independence *and* from conceptions of rational autonomy, with their inescapable links to preference-led and desire-driven accounts of motivation.

Kant's views on autonomy were also innovative. No earlier writer had made such strong claims about the moral importance of any conception of autonomy. Jerry Schneewind is, I believe, correct when he writes at the beginning of his recent book, *The Invention of Autonomy*, that 'Kant invented the conception of morality as autonomy.'¹³ But what he invented has, I believe, little to do with twentieth century conceptions of autonomy either as *individual independence* or as *rational autonomy*. The most convincing evidence that Kant was thinking of something quite different is that very few of his central claims about autonomy make much sense if we equate Kantian autonomy *either* with individual independence *or* with current conceptions of rational autonomy.

On the surface, Kant may seem to be promoting some version of rational autonomy. For example, in the *Critique of Practical Reason*, and in many other works, he predicates autonomy of

13. Schneewind, 1998, 3.

reason and links the autonomy of reason to morality. He writes, for example, that 'the moral law expresses nothing other than the *autonomy* of pure practical reason.'¹⁴ In some texts, such as his late essay, *The Conflict of the Faculties*, he goes even further and equates all reasoning, not only practical reasoning, with autonomy, remarking that 'the power to judge autonomously—that is, freely (according to principles of thought in general)—is called reason.'¹⁵

Taken out of context these claims could be—and often are—read as very strong and very confused versions of the idea that autonomy is some form of rational choosing. They seem even less plausible than contemporary accounts of rational autonomy because they require us to read Kant as making exorbitant assertions about the links between autonomy, reason and morality. Rather than dismissing Kant's account of autonomy as bizarre on this account, I shall consider a reading that takes his specific claims about the structure of autonomy seriously.

The context of Kant's account of autonomy is a distinctive view of action. Kant looks at action from the agent's standpoint; that is, from a practical point of view. He depicts agents as having a *power of choice* (*Willkür*) that they can use in innumerable ways. Agents exercise their power of choice by adopting one or another *determination of the will*. In doing so they freely adopt some practical principle (or rule, or law, or plan) making it their maxim.¹⁶ Maxims specify at a fairly general level some aspect of the way agents set about leading their lives: I may make it my maxim to build a mill,¹⁷ or to save for my old age,¹⁸ to embezzle funds,¹⁹ to avenge insults,²⁰ to overcharge gullible customers,²¹ to pursue my self-interest,²² or not to make false promises.²³ This heterogeneous sample includes maxims that Kant thinks of as

14. Kant, 1787, 5:33.

15. Kant, 1798, 7:27.

16. Timmerman, 2000, 39–52.

17. Kant, 1787, 5:26.

18. *Ibid.*, 5:20.

19. *Ibid.*, 5: 28.

20. *Ibid.*, 5:19.

21. Kant, 1785, 4:397.

22. Kant, 1787, 5:23–26, probably contains the most sustained discussion; but there are many other discussions of self-love and self interest.

23. Kant, 1787, 5:21; Kant, 1785, 4:403.

morally worthy and others that he thinks of as morally unacceptable, and some that he thinks merely optional or at most a matter of prudence.

In speaking of agents' *maxims* or *determinations of the will* Kant is not making a claim about the efficient causes of action. The principle (law, rule, plan) that an agent adopts does not cause him or her to do anything (how could abstract entities such as *principles* (*laws, rules* or *plans*) be efficient causes?). Rather the principle (law, rule, plan) that is adopted is the formal cause of action: it articulates *what* an agent chooses to do. This account of action does not require Kant to deny that (from a theoretical standpoint) action can be causally explained: he says in many passages that acts have efficient causes. But in choosing what to do we do not identify the causes of our future action.

Determinations of the will are a promising focus for an account of reasons for action. Since any principle, rule, law or plan that is or could be adopted as a determination of the will must have propositional structure and content, it will be apt for reasoned assessment. Moreover, reasoned assessment need not be confined to judging whether proposals for action are efficient or effective means to given ends: for Kant instrumental rationality need be only one aspect of practical reason. Here I shall set his account of instrumental reasoning on one side in order to concentrate on the basis of his distinction between heteronomous and autonomous reasons for action.

V

Kantian Autonomy: Heteronomy as a Clue. The contrast between *heteronomy* and *autonomy* is a useful way into understanding Kantian autonomy. Heteronomy is not a term we use in daily life, so may not seem a promising clue. However—for once!—Kant offers a helpfully simple account of what he means. He writes:

If the will seeks the law that is to determine it *anywhere* else than in the fitness of its maxims for its own giving of *universal* law ... heteronomy always results.²⁴

24. Kant, 1785, 4:441.

The difference between heteronomous and autonomous choosing is a difference between the sorts of principles adopted as determinations of the will. Heteronomous choosing and Kantianly autonomous choosing are both a matter of *seeking, choosing, adopting* or *willing*, some principle (law, rule, plan), thereby adopting some *determination of the will*. The difference between them is not that those who choose heteronomously are not free agents, or not capable of any independence in action, or that they have no rational capacities, or that they cannot seek, choose, adopt or will laws or principles. Heteronomous choosing is choosing. Agents with the *power to choose (Willkür)* are capable both of autonomous and of heteronomous choosing. The difference between them is that agents look in different directions in choosing autonomously and choosing heteronomously. The two types of principle are drawn or derived from *different types of assumption*.

Kant frequently contrasts heteronomous and autonomous principles by saying that the former take their justification *from elsewhere*, whereas autonomous principles take their justification *from nowhere else*. Yet here we may well lose patience. Surely justifications *must* begin somewhere *else*: isn't the whole point of justification to *derive* authority? And if a principle could be non-derivative, why would that make it morally special, as opposed to especially arbitrary? Why should a principle that is *merely* non-derivative (whatever that means) have *any* standing, let alone be fundamental to a conception of reason? Have we not ended up with something worse than the fantasy that individual autonomy is the basis of morality? Have we not descended from pop existentialism to postmodernism?

Kant's examples of principles or laws adopted by those who choose heteronomously are extremely varied. Heteronomous agents may defer to the dogmas of a Church, to the edicts of rulers, to immediate inclination or to the will of the majority.²⁵ The common core to all these examples is that the heteronomous chooser makes some *arbitrary*, hence unreasoned, move in choosing a determination of the will, whereas the autonomous chooser does not. A heteronomously chosen principle is 'justified' by imputing authority to something or other, for whose standing or

25. For numerous examples see Kant, 1784 and 1785.

authority either no reasons, or (at most) incomplete reasons, are given. Any reason to act on such principles *reflects that arbitrary assumption*, and heteronomous reasons for action are therefore always conditional upon it. Kant puts the point as follows:

Wherever an object of the will has to be laid down as the basis for prescribing the rule that determines the will, there the rule is none other than heteronomy; the imperative is conditional, namely: *if or because* one wills this object, one ought to act in such and such a way; hence it can never command morally, that is, categorically.²⁶

So the common core of all sorts of heteronomous willing is that it is *not fully reasoned*. It depends on arbitrarily ascribing authority to something or other: a desire or a dogma, the edicts of Church or State. Kant often depicts those who arbitrarily impute authority to such assumptions metaphorically as submitting to *alien* or *foreign* authorities.

Kant's numerous accounts of heteronomous willing reveal fundamental differences between lacking *individual autonomy*, lacking *rational autonomy* and lacking *Kantian autonomy*. Beings who wholly lack individual autonomy will not be able to choose *either* heteronomously *or* autonomously. In Kantian vocabulary, such beings lack the power of choice, so lack free will, and are incapable of moral choice or action. Having a power of choice is a precondition for heteronomous as well as for Kantianly autonomous choosing, so cannot be equated with Kantian autonomy. Kant speaks of beings without a power of choice, who cannot act either heteronomously or autonomously, as having no more than animal capacities to choose, an *arbitrium brutum* as opposed to *arbitrium liberum*.²⁷

A person with power of choice can choose either autonomously or heteronomously. Agents who choose heteronomously, so defer to arbitrarily selected standards and authorities, can give *at most* conditional reasons for their actions. They need not, however, be wholly capricious, and often embrace one or other version of what Kant calls *heteronomous ethics*. For example, they may choose not to follow immediate inclinations but to live with an eye to long-term personal advantage (egoism), or to the general happiness (Utilitarianism) or to the claims

26. Kant, 1785, 4:444.

27. Kant, 1797, 6:211–14.

of a supposed moral sense, or to some conception of perfection.²⁸

Kant would view twentieth century proponents of rational autonomy as endorsing various forms of heteronomy in ethics. They do not admire mere, sheer wilfulness; they claim that rationally autonomous agents can offer reasons for action, but they also accept that these reasons are always less than complete. Heteronomous choosers ultimately have to fall back on ascribing authority to desire or ideology, public opinion or certain powers that be.

Although heteronomous choosers make an arbitrary move in according certain desires, demands and dogmas the status of reasons for action, they may have quite a lot of moral luck. Their heteronomous choices are often expressed in morally acceptable action (the shopkeeper who is honest for the sake of his reputation, the self-interested chooser whose interests coincide with altruistic choices). But in other situations heteronomous choosers may act in capricious, self-centred or even malign ways. Common or garden heteronomy is reflected in lives that with luck are morally conformist, but without luck may be morally unacceptable.

VI

Kantian Autonomy and Self-Legislation. The limitations of heteronomous reasons are easily seen and constantly noted. An agent chooses heteronomously by adopting a principle in order to achieve something for which he offers only conditional reasons. But can we expect more? Kant thinks so. He claims that autonomous choosers adopt principles of action that are not conditional on any arbitrary assumption or posit. In the *Groundwork* he puts the matter as follows:

Autonomy of the will is the property of the will by which it is a law to itself (independently of any property of the objects of volition).²⁹

In this and many similar passages, we meet the most difficult aspect of Kant's account of autonomy. What is a 'property of

28. Kant, 1785, 4:442–3; see Kant, 1787, 5:39–41, for a more differentiated classification of heteronomous ethical positions.

29. Kant, 1785, 4:440.

the will by which it is a law to itself (independently of any property of the objects of volition)? How can the will adopt a determination simply on the basis of 'the fitness of its maxims for its own giving of universal law'? Kant's claims have a crabbed reflexivity that is often perplexing and hard, but not I think impossible, to sort out.

A common approach to Kantian autonomy harks back to the etymology of the word *autonomy*, and identifies autonomous willing with some conception of *self-legislation*. What sense can we assign to Kant's use of this venerable phrase? Far the most popular way of looking at the matter is to interpret *self-legislation* as *legislation* done by a *self* or *subject*. On this individualistic reading we picture each of many wills as legislating for all. Two questions immediately arise. First, why should the legislative action of my will and your will point in the same direction? Second, why should we think that the 'legislative action' of anybody's will must point in a morally acceptable direction, and how could such 'legislative action' convince us that the 'principle of autonomy is the sole principle of morals.'³⁰ If Kantian autonomy is pictured merely as *legislation by individual selves*, the coordination of different wills remains a mystery and the moral importance of autonomy is just as obscure as it is in conceptions of individual autonomy that make no mention of self-legislation. Unsurprisingly, many passages in Kant's writings are reduced to nonsense, or at the very least to implausibility, if we equate his conception of autonomy with 'legislation' by individuals, although this reading remains very popular.

Could this problem be resolved if Kantian autonomy were pictured as *legislation by co-ordinated individual selves*? This is the strategy of Rousseau's famous account of self-legislation, where the problem of possible divergence between numerous self-legislating wills is resolved by positing an extraneous source of convergence. On Rousseau's view 'The general will is always for the common good'³¹ and 'The general will is always upright and always tends to the public utility.'³²

30. Kant, 1785, 4:440.

31. Rousseau, 1755, 8.

32. Rousseau, 1762, 58.

Of course, Rousseau did not think that individuals *actually* will as the General Will demands. On the contrary, any convergence of particular wills on the General Will is a contingent convergence of 'corrected' wills, not necessarily or even generally mirrored in real time by the will of each or the will of all. As Kant would see it, Rousseauian self-legislation is a form of heteronomy: it assigns authority to a conception of the general good, and defines 'corrected' wills as pointing in this direction. Even if the problem of divergence were solved by this strategy—I personally think that Rousseau defines the problem out of existence rather than solving it—we have not been shown any reason to think that the supposed point of convergence defines morality, or is morally significant. Rousseau's account of *legislation by co-ordinated selves* resolves indeterminacy and disagreement by positing the authority of the general will or of the general good: for Kant this is heteronomy.

VII

Kantian Autonomy: Law-Like and Lawless Willing. Could an alternative interpretation of the idea of self-legislation make better sense of Kant's claims, and avoid conflating autonomy and heteronomy? It may seem that we are faced with a dilemma. If we view self-legislation simply as a matter of choosing or willing for oneself (independently, or even using some rational process), then the very distinction between heteronomy and Kantian autonomy is erased. If we view self-legislation as a matter of choosing or willing principles *with a certain sort of content or aim* we apparently fall into *heteronomy* by arbitrarily ascribing moral weight to that content or aim.

As is apparent from countless passages, Kant thinks that an essential feature of autonomous willing is that it have the *form of law*, so is expressed in *law-like* determinations of the will. He contrasts *law-like* choosing with choosing determinations of the will are literally *lawless*. But what, we may wonder, is the problem with lawless choosing? Why shouldn't we embrace the most extreme forms of lawlessness or lack of structure both in thought and action? Isn't any claim that thought or action *must* meet any requirements, including being law-like, intrinsically dogmatic and arbitrary? Postmodernists, after all, have asserted that we can

dispense with all and any requirements or structure or reason: the claims of law-like form surely need arguing.

It is fascinating to find Kant engaging with the postmodernists of his day—the advocates of spiritual enthusiasm or *Schwärmerei*—to show why the postmodernist fantasy of dispensing with all authorities, with all reasons, with all principles or laws for organising thinking or action, is deceptive. Like postmodernists, Kant sees clearly why people *imagine* that ‘lawless’ thinking is not merely feasible but attractive; but he also sees what it threatens. He depicts the pleasures of the advocates of lawless thinking with pointed irony:

First *genius* is very pleased with its bold flights, since it has cast off the thread by which reason used to steer it. Soon it enchants others with its triumphant pronouncements and great expectations and now seems to have set itself on a throne, which was so badly graced by slow and ponderous reason, whose language, however, it always employs. Then its maxim is that reason’s superior law-giving is invalid— we common human beings call this **enthusiasm**, while those favoured by beneficent nature call it *illumination*.³³

Kant believes that this heady liberation ends not merely in confusion, but in cognitive and practical disaster:

Since reason alone can command validly for everyone, a confusion of language must soon arise among them; each one now follows his own inspiration.³⁴

Communication breaks down and superstition rides high; morality and civil society fail. Attempts to achieve unlimited freedom in thinking and acting prove self-defeating. Lawlessness undermines thinking and acting because it undercuts the very possibility of offering others reasons for believing or for acting.

As Kant sees it, any reasoned use of human freedom in thinking and in acting must be *law-like* rather than *lawless*:

Freedom in thinking signifies the subjection of reason to no laws except *those which it gives itself*; and its opposite is the maxim of a **lawless use** of reason (in order, as genius supposes, to see further than one can under the limitation of laws). The natural consequence of declared lawlessness is that if reason will not subject

33. Kant, 1786, 8:145.

34. Kant, 1786, 8:145.

itself to the laws it gives itself, it has to bow under the yoke of laws given by another; for without any law, nothing—not even nonsense—can play its game for long. Thus the unavoidable consequence of *declared* lawlessness in thinking (of a liberation from all the limitations of reason) is that freedom to think will ultimately be forfeited and—because it is not misfortune but arrogance which is to blame for it—will be trifled *away* [*verschertzt*] in the proper sense of the word.³⁵

Only law-like thought and action offers others with whom we live, think and interact proposals that they can follow in thought or action, so could potentiality evaluate as reasons for action. Whatever else reasons are, they must be *followable by others*, and hence the sorts of things that we can offer and refuse, accept and challenge. This is why practical reasoning cannot cut loose from *law-like* determinations of the will. If we are cavalier about law-likeness, we no longer deal in reasons for acting or reasons for believing.

Needless to say, the demand that we act on law-like principles is an extremely weak constraint, that is met both by heteronomous and by autonomous action. Those whose principles of action are heteronomous through and through act on law-like principles. Even if Kant is right in diagnosing lawlessness in thought and action as catastrophic, heteronomy might be the *only* other option for conducting our thinking and acting. Perhaps, contra Kant, morality is at most a system of hypothetical imperatives. Perhaps all reasons for action are ultimately conditional.

VIII

Kantian Autonomy: A Law to Itself. If Kantian autonomy is a possibility, there must be *two* sorts of law-like principles, and reasons for thinking that heteronomous principles sometimes offer insufficient reasons for action. Kant's constant refrain in his practical philosophy, from the first Chapter of the *Groundwork* onwards, is that morally important principles must not only be law-like in *form*, but universal in *scope*. Since heteronomous principles arbitrarily take for granted specific desires, convictions or interests, or specific institutions or cultures, they are not capable

35. Kant, 1786, 8:145.

of being principles for all. By contrast, law-like principles that are capable of being principles for all, that have universal scope, are Kantianly autonomous principles.

Kant connects the demands of scope and law-likeness by insisting that

It is requisite to reason's lawgiving that it should need to presuppose only *itself*, because a rule is objectively and universally valid only when it holds without the contingent, subjective conditions that distinguish one rational being from another.³⁶

Kantian autonomy is not a matter of persons being independent (although, of course, Kant holds that persons are independent to a degree, since they have a power of choice). Rather Kantian autonomy is a matter of adopting *law-like principles that are independent of extraneous assumptions that can hold only for some and not for other agents*. Kant often encapsulates this requirement in phrases such as a 'lawgiving that needs to presuppose only itself' or in compressed references to the notion of 'a lawgiving of its own' or 'non-derivative lawgiving'. Principles that meet this standard are not merely law-like, but 'hold without the contingent, subjective conditions that distinguish one rational being from another'. They are potentially principles for all, and not merely for those who uncritically assume the authority of some desire or a particular dogma, some local institutions of powers that be, who can at most converge on heteronomous principles.

The idea of a 'lawgiving that needs to presuppose only itself' is I think the key to Kant's distinctive understanding of self-legislation. As he presents the matter, it is the *principle* of willing or *legislating*, and not the agent, that 'presupposes only itself'. For Kant the term *self-legislation* cannot mean that there are some terrific acts of the self (or terrific acts of the terrific self?) that are morally important, or definitive of morality, but only that there are some principles of action that are not derivative from supposed, but ultimately arbitrary 'authorities', and that these are morally important. The element *self* in the notion of *self-legislation* is *reflexive rather than individualistic*; it applies to certain justifications of principles rather than to certain agents or

36. Kant, 1787, 5:20; freedom in the positive sense, as we know from *Groundwork* III, is another term for Kantian autonomy.

'legislators'. Kant takes himself to be giving an account of the sort of law-giving that is reason's own, and not an account of lawgivings that are an agent's own. His understanding of self-legislation places the emphasis on the notion of legislation, rather than on any notion of the self. Kantian autonomy is *reason's lawgiving* rather than the *lawgiving of individual agents* (whatever that might mean).

Reason's lawgiving is not merely a matter of adopting one or another law-like determination of the will: heteronomous action does as much. Kantian autonomy is expressed in adopting principles, willings, reasonings that are *both* law-like in structure or form *and* do not derive that law-likeness from arbitrary assumptions that are open to some but not to others.

Heteronomous reasoning, by contrast, relies on arbitrary assumptions about the basis of morally significant principles, which may be available to some and not to others. Heteronomous principles may be widely shared: those who take for granted the authority of Church or state, public opinion or local ideology will generally have plenty of company. Unlike postmodernist gestures, heteronomous practical reasoning does not end in incomprehension or cognitive shipwreck. Its defect is only that its intelligibility coasts on arbitrary assumptions that some others cannot share, so cannot provide reasons for action for all others. Heteronomous reasons do not aspire to be

a law-giving of its own on the part of pure and, as such, practical reason [which] is freedom in the *positive* sense.³⁷

Hence, in Kant's view, heteronomous reasons are ultimately defective, incomplete or 'private' reasons. Reasons need to be in principle *followable* by their presumed audiences; fully reasoned claims and demands must be followable by all and any audiences. That is why they must be law-like, or have the form of law. Heteronomous reasons are law-like in form but presuppose some shared desire or belief, or other cultural or institutional source of agreement. They may provide in-group reasons to those others who have deferred (heteronomously) to the same 'authority': but they offer no basis for reasoning among those who do not presuppose allegiance to the same desire or dogma or doctrine,

37. Kant, 1787, 5:33: 'diese *eigene Gesetzgebung* ... der reinen und als solche praktischen Vernunft ist Freiheit im *positiven* Verstande.'

hence no basis for reasoning with the world at large, no basis for fully 'public' reason, hence in Kant's view no sufficient basis for morality or for justice. This too is perhaps quite a weak conception of practical reason: but it is not covert reversion to heteronomy.³⁸

Correspondingly, the phrase 'a lawgiving of its own' (*eine eigene Gesetzgebung*) is no mere awkwardness of locution. It expresses the requirement that anything that can count as the self-legislation of practical reason must be a not only *law-giving* (*gesetzgebend*) but also *non-derivative* (*eigen*). It is this demand for living by principles that could be described as *law-givings of reason*, so fit for all (regardless of their particular faith or ideology, desires or culture), that underpins Kant's distinctive use of the metaphor of self-legislation and that links his conception of autonomy to fitness for universal law. It is this picture that allows him to claim that autonomy and universalisability provide alternative formulations of the Categorical Imperative, inseparable from one another and equally, indeed equivalently, fundamental to morality. As Kant puts it,

The principle of autonomy is, therefore: to choose only in such a way that the maxims of your choice are included as universal laws in the same volition³⁹

As Kant sees it, combining a *formal* requirement (law-likeness) with a *scope* requirement (universality) allows us to derive certain substantive constraints, which he views as basic principles of morality. Morality is fundamentally a matter of setting aside principles of action that could not be adopted by all, so could not be universal laws. If we adopt only law-like determinations of the will that could be universal laws, we must adopt only principles that (we judge) all and any others too could adopt, and so must reject many tempting and interesting principles of action. Kantian autonomy bypasses the problem of possible divergence

38. Indeed, Kant sometimes presents it as the basis of all reasoning and not just of practical reasoning, including that part of practical reasoning that supports morality: 'To make use of one's reason means no more than to ask oneself, whenever one is supposed to assume something, whether one could find it feasible to make the ground or the rule on which one assumes it into a universal principle for the use of reason.' Kant, 1786, 8:146n.

39. Kant, 1785, 4:439; see Wood, 1999, 163–4 for a listing of versions of the Formula of Autonomy.

of individual choices, which would have to be resolved if self-legislation were a matter of each individual legislating for all. The key to Kant's thought is the explicit identification of self-legislation or autonomy with adopting only law-like principles that can 'hold without the contingent, subjective conditions that distinguish one rational being from another'.

Once we have shifted our conceptions of self-legislation in the way that a coherent reading of the Kantian texts requires, it is no longer so hard to see why he thinks that autonomy is a demand of practical reason. If we think of reasons as given and received, exchanged or refused, accepted or challenged, putative reasons that cannot be followed by some of those to whom they are offered will be defective or incomplete: they offer only limited, incomplete reasons for action.

IX

From Practical Reason to Morality. Practical reason at least demands that principles we offer to others as basic reasons for action are indeed fit to be reasons for others, so are Kantianly autonomous. Kant, I think, assumes that once we have an adequate account of practical reason, an account of morality will not be far away. He writes in the *Groundwork* 'That the above principle of autonomy is the sole principle of morals can well be shown.'⁴⁰

I think this too optimistic; and it is certainly a task for another day. Although I hope I have set out why the Kantian principle of autonomy is fundamental to reason giving, it would need more work to determine whether it is the sole principle of morals, so equivalent to the other formulations of the Categorical Imperative, and to set out the role of other considerations in moral reasoning. This task constitutes the programme of Kant's ethical and political writings. There he aims to show that

Autonomy of the will is the sole principle of all moral laws and of duties in keeping with them; *heteronomy* of choice, on the other hand, not only does not ground any obligation at all but is instead opposed to the principle of obligation and to the morality of the

40. Kant, 1785, 4:440 'Allein daß gedachtes Princip der Autonomie das alleinige Princip der Moral sei, läßt sich gar wohl darthun.'

will. That is to say, the sole principle of morality consists in independence from all matter of the law (namely, from a desired object)⁴¹ and at the same time in the determination of choice through the mere form of giving universal law that a maxim must be capable of⁴²

In the fairy tale the emperor processed stark naked, and only a child dared to point this out. As I see it, the newer versions of autonomy that have played so large a part in discussions of morality and politics since the mid-twentieth century, and that have penetrated the innermost and outermost reaches of public and professional life (especially in the English speaking world) are pretty scantily clad. Neither mere, sheer independence nor so-called rational autonomy has much to commend it, though each, no doubt, can be contrasted with even nastier possibilities. We face a choice. Either we accept some contemporary conception of autonomy, so must conclude that it is at best a minor (and sometimes suspect) aspect of the moral life. Or we could take the Kantian conception of autonomy seriously, and have at least reason to consider whether it just might be 'the sole principle of all moral laws and of duties in keeping with them'.

REFERENCES

- Dworkin, G., 1988, *The Theory and Practice of Autonomy* (Cambridge: Cambridge University Press).
- Faden, R. and Beauchamp, T., 1986, *The History and Theory of Informed Consent* (New York: Oxford University Press).
- Frankfurt, Harry G., 1971, 'Freedom of the Will and the Concept of a Person', *Journal of Philosophy* 68, 5–20.
- Hill, T. E., 1992, 'The Kantian Conception of Autonomy', in his *Dignity and Practical Reason* (Ithaca, New York: Cornell University Press), 76–96.
- Kant, I., 1784, *An Answer to the Question: What is Enlightenment?*, in *Practical Philosophy*, trans. Mary J. Gregor (Cambridge: Cambridge University Press, 1996), 8:35–42.
- Kant, I., 1785, *Groundwork of the Metaphysic of Morals* in *Practical Philosophy*, trans. Mary J. Gregor (Cambridge: Cambridge University Press, 1996), 4:385–463.
- Kant, I., 1786, *What Does it Mean to Orient Oneself in Thinking?* in *Religion and Rational Theology*, trans. Allen W. Wood and George di Giovanni (Cambridge: Cambridge University Press, 1996), 8: 133–46.

41. And, on a fuller account of heteronomy, from all ideologies, cultures and institutions that 'distinguish one rational being from another'.

42. Kant, 1787, 5:33.

- Kant, I., 1787, *Critique of Practical Reason* in *Practical Philosophy*, trans. Mary J. Gregor (Cambridge: Cambridge University Press, 1996), 5:1–163.
- Kant, I., 1797, *Metaphysics of Morals*, in *Practical Philosophy*, trans. Mary J. Gregor (Cambridge: Cambridge University Press, 1996), 6:203–493.
- Kant, I., 1798, *The Conflict of the Faculties*, in *Religion and Rational Theology*, trans. Allen W. Wood and George di Giovanni (Cambridge: Cambridge University Press, 1996), 7:5–116.
- Mill, J.S., 1859, *On Liberty*, in *Utilitarianism and other Writings*, ed. Mary Warnock (Glasgow: William Collins and Sons, 1962), 126–250.
- Mill, J.S., 1862, *Considerations on Representative Government*, in *Collected Works of John Stuart Mill* (London: Routledge, 1991), 29: 371–577.
- Rousseau J.-J., 1762, *Social Contract*, trans. Victor Gourevitch (Cambridge: Cambridge University Press, 1997), 39–152.
- Rousseau, J.-J., 1755, *Political Economy*, trans. Victor Gourevitch (Cambridge: Cambridge University Press, 1997), 3–38.
- Schneewind, J.B., 1998, *The Invention of Autonomy* (Cambridge: Cambridge University Press).
- Skorupski, J., 1989, *John Stuart Mill* (London: Routledge).
- Timmerman, Jens, 2000, 'Kant's Puzzling Ethic of Maxims', *Harvard Philosophy Review*, 39–52.
- Wood, Allen, 1999, *Kant's Ethical Thought* (Cambridge: Cambridge University Press).