The Scope *of* Autonomy

Kant and the Morality of Freedom

KATERINA DELIGIORGI

The Scope of Autonomy

Kant and the Morality of Freedom

Katerina Deligiorgi

Preface

'Whatever be the consequences of my experiment, I am resolved to judge with my own eyes of the various conditions of men, and then to make deliberately my *choice of life*' (Johnson 1985: 69). *The Scope of Autonomy* is about the important philosophical issues that arise from the aspiration, expressed here by Rasselas in Samuel Johnson's tale, to give deliberate shape and direction to our lives. Finding what is worth doing and choosing accordingly is, I want to argue, essential to possessing autonomy. But as Johnson makes clear, this project is fraught and he warns his readers that it may be chimerical: one of those 'absurd' projects which we 'familiarise . . . by degrees, and in time lose sight of their folly' (Johnson 1985: 135). The folly in the ambition of making one's choice of life, Johnson suggests, consists in becoming forgetful of our finitude; we assume that 'what now acts shall continue its agency and what now thinks shall think on for ever' (Johnson 1985: 149). The message is reinforced during Rasselas's visit of the Egyptian catacombs, 'those that lie here stretched before us . . . warn us to remember the shortness of our present state: they were, perhaps, snatched away while they were busy, like us, in the choice of life' (Johnson 1985: 149). The theory of autonomy I present here offers a way of asserting our freedom without forgetting our humanity. Essential to this task is showing how the aspiration to be an agent—to make our own choice of life—is properly considered in conjunction with the expectation that we shape our behaviour in response to the legitimate claims others make on us. The defence of this moral theory of autonomy, which I develop in response to recent arguments about agency, morality, and practical reasoning, has Kantian roots and is Kantian in orientation.

For many contemporary critics, this Kantian heritage is part of the problem. Most recently, Charles Larmore has argued that a morality of autonomy is fundamentally misguided and that the Kantian conception of autonomous agents as self-legislators promotes the mistaken belief that the authority of moral norms depends on individuals (Larmore 2008: 88).[1]

---

[1] It is debatable whether Larmore's charges against Kant and Kantian constructivist readings in Larmore 2003 and 2008 are always on target; see Ameriks 2003: 279–82 and Stern 2009.

Underpinning this individualistic conception of autonomy, which Larmore attributes to Kant and to contemporary Kantians, is a hollow view of agency and an ultimately egoistic view of morality.[2] It is certainly hard to see how autonomy can express a moral ideal, if it comes down merely to individual self-endorsement. So it is understandable that the search for answers to the question of what is a life worth living have often been sought among contextual or communal conceptions of the good that are ultimately justified by reference to social norms, natural facts, or *sui generis* normative entities, as in Larmore's own work (Larmore 2008: 63). I have some sympathy with the criticisms that motivate these projects, but I hope to show that the charge of individualism is not applicable to Kant's conception of autonomy. Rather these criticisms are more directly applicable to contemporary psychological and naturalist accounts of personal autonomy that focus on how an agent organizes her desires and intentions to act. These accounts, which I examine in the last chapter of this book, are vulnerable to the problems Larmore describes precisely because they are highly individualistic and explicitly non-moral.

Other problems traditionally raised in the context of Kant's conception of autonomy relate to the emphasis Kant places on the role of rational self-reflection. On the familiar picture, the autonomous agent is one who prioritizes reason over desire and so acts on what reason demands without regard to wants, interests, and the like. This reading has generated considerable literature on the question of motivation, in particular whether it is at all possible to act in the absence of desires or interests.[3] One of the aims of this book is to move on from this debate by showing that autonomy is centrally about *wanting* to do the right thing. A more recent generation of Kantians, including Christine Korsgaard, Barbara Herman, and Marcia Baron, have sought to correct the traditional view of Kantian rational agency by arguing that autonomy is about realizing a certain value; it is about achieving a standard of good agency by means of independent reflection on one's actions.[4] These philosophers share a more generous approach to notions of identity, which encompasses the

---

[2] See Larmore 2008: 44. In Samuel Johnson's story, the prince and his companions, as soon as they escape the confines of their homeland, 'seeing nothing to bound their prospect, consider themselves as in danger of being lost in a dreary vacuity' (Johnson 1985: 73).

[3] A good discussion of how the entirely disinterested and desire-less agent borders on the pathological is Langton 1992; see also Deligiorgi 2006.

[4] See Korsgaard 1998 and 2002; Herman 1993; Baron 2000.

agent's commitments and interests. As many critics have pointed out, however, the notions of agency and of self-determination that underpin these neo-Kantian interpretations do not suffice to support moral commitments. It is one thing to be a good agent in the required sense and quite another to be someone who takes moral considerations into account.

The theory of autonomy I develop in this book has at least two aspects that make it distinctive. First, I argue that 'autonomy' is not capturable in simple definitions such as 'acting on one's higher order desires', or 'acting on general principles'.[5] To do justice to the practical ideal of autonomy a complex theory is required that places autonomy at the intersection of concerns with morality, practical rationality, and freedom. Second, whereas autonomy has primarily been understood in terms of our relation to ourselves, I show that it also centrally involves our relation to others. I thereby open up the concept of autonomy to investigation of its intersubjective dimension. This is key to the moral claim autonomy makes on us. Showing how autonomy pertains to morality, practical rationality, and freedom without exhausting these topics is the task of tracing the scope of autonomy and forms an essential part of the project. To that extent, the account I offer here is a modest one. It requires, however, that we do not give up certain philosophical ambitions with respect to moral philosophy.

Bernard Williams warned that 'various versions of moral philosophy share a false image of how reflection is related to practice, an image of theories in terms of which they uselessly elaborate their differences from one another' (Williams 1985: 198). What causes this 'false image' is a tendency to overstate the coherence of the phenomenon studied, seeking to show that there are correct answers to all our moral questions and that these form a consistent theory; all we need is the right calculus and breeding. In marked contrast to this view is what we might call the tragic outlook, which takes as given that we can never know what is right, that morality is beset by insolvable dilemmas and so the task of philosophy is to remind us of this or to reconcile us to this. The former can be gratingly

---

[5] The former definition relates to the family of views I called 'psychological naturalistic' and comes from Harry Frankfurt's early work; see Frankfurt 1988 and for this characterization of the personal autonomy family of views see Taylor 2005: 2. The latter is Kantian, see Hill 1991: 44.

upbeat and inadvertently cruel, the latter insufferably complacent and frustratingly quietistic.

I treat the idea that morality can be characterized as a coherent set of practices initially as a working assumption to show that the search for an ordering principle is plausible and desirable; and since the principle in question is also critical, that is, it enables us to revise some of our commitments, it can be seen to contribute to the establishment of such order. This is not to deny that questions can arise that we may not know how best to answer. It is to deny that these difficulties are created, exacerbated or left untouched by philosophical reflection. So the arguments contained in this book are aimed against a certain type of scepticism, which considers that moral knowledge is impossible under conditions of philosophical reflection.[6] Philosophy, I argue, is in a position to shoulder the weight of the task of helping us find an objective answer to the question of what it is right to do.

One of the most important features of Kantian autonomy is that it turns the search for objectivity into a search for what is shareable, whilst at the same time seeking to stretch our conceptions of shareability. This goes to the heart of issues not only of justification but also of normative content. Central to the claim that autonomy is a moral ideal is the argument that autonomy describes an intersubjective rather than intra-subjective norm. The moral content of autonomy is given by the notion of the law (*nomos*) and the demand that one think of oneself in relation to others under a shared law. Autonomy requires that we engage in a 'perspectival ascent' from a subjective viewpoint to one that putatively embraces all rational others.[7] The intersubjective dimension of autonomy reveals common ground between Kant and his successors, Schiller and Hegel, who sought to revise Kantian ideas in order to strengthen intersubjective commitments.

The approach adopted in this book balances reconstruction with analysis. I prioritize contemporary over historical texts because I seek to address

---

[6] Although Michael Williams characterizes in this way a different philosophical trouble (Williams 1991: 130), the purported self-defeat of philosophical reflection describes well the basic diagnosis underpinning counsels of caution issued within moral philosophy; examples include Prichard's criticisms of moral rationalism in Prichard 1912, and indeed Bernard Williams's reservations about the usefulness of certain types of philosophical reflection in Williams 1985.

[7] I take the notion of perspectival ascent from Sacks 1989: 188; see too Ch. 4 Sect. 3.

contemporary concerns. Yet I look back to Kant in doing so because I think that Kant's writings provide us with conceptual tools that enables us to transcend some of our own historical limitations, certain ingrained habits of thought that do not allow us to make progress in the way we conceive of our autonomy.[8] It is not always easy to establish a dialogue between historical and contemporary writings on ethics, because of shifts in philosophical concerns as well as in terminology. In the shorthand employed in current meta-ethics, I defend the possibility of a moral standpoint that combines an unusual yet coherent set of attributes: it is objectivist and cognitivist without being intuitionistic; realist but non-naturalist; anti-Humean about practical reasons but not necessarily anti-Humean about motivation.

Chapter 1 is introductory; it aims to locate the central concepts, ideas, and questions that arise in the context of current debates about autonomy. This chapter provides the basic Kantian orientation for the theory of autonomy I present in the main part of the book.

Chapter 2 deals with questions of moral knowledge; how we know right and wrong, and how we can justify such knowledge on Kantian grounds. The chapter takes 'right' as its central moral concept, a thin concept that captures the core features of Kant's concept of 'duty'. This chapter covers the epistemic component of autonomy and supports a cognitivist, objectivist, and Kantian realist interpretation of right.

Chapter 3 deals with questions of moral action. It fulfils both a negative and a positive task. The negative task consists in contextualizing the questions of moral psychology, which have tended to dominate discussions of Kantian ethics. I argue that the kind of motivation an agent has and the things she counts as reasons have the importance they do because of the metaphysics of free agency Kant holds. The positive task consists in developing a psychologically plausible account of motivational autonomy.

---

[8] I think that we need not be bound by the choice Frederick Beiser offers when he asks that we choose 'between anachronism and antiquarianism' (Beiser 2008b: 7, see also Beiser 2007: 84–7). Avoidance of Beiser's dilemma can take the form of austere engagement with philosophy's history as described in Danto 1997: 1–5 or of stepping with circumspection into the 'hermeneutic circle' as described in Ameriks 2006: 33–50, here 49). My approach is somewhere in this continuum: I generalize the maxim that it makes sense sometimes to 'step beyond the perspective of Kant scholarship alone and to reflect on the basic features that contemporary philosophers would insist that any acceptable moral theory treat with sensitivity' (Ameriks 2006: 89).

Chapter 4 draws some of the broader consequences of the anti-naturalist assumptions and implications of my reading. First, I show that naturalism provides the framework for influential contemporary interpretations of Kantian autonomy. I then address the external reasons and categoricity debates. My aim is to show that it is possible to develop an account that is externalist, and so anti-Humean, about practical reasons but not necessarily anti-Humean about motivation. Having established this, I return to the ethical substance of autonomy, its intersubjective normative content. This discussion links up with the first chapter and addresses the universalizability formula of right and apriority in ethics.

Chapter 5 returns to some of the issues about motivation discussed in Chapter 3 by looking at Schiller's arguments about the role of emotions in ethics. This enables me to bring to the foreground two divergent conceptions of ethics, between which Schiller equivocates, one that is intersubjective and one that is self-perfecting. I argue that emotions can be accommodated within an ethic of autonomy, provided they complement the cognitive and motivational components that make up the theory of autonomy defended here.

Chapter 6 starts with locating the theory of autonomy presented here in the contemporary discussion about personal autonomy using the issues raised by Schiller as a conceptual bridge. Whereas the theory avoids some of the familiar pitfalls, the question of the metaphysics of freedom, postponed from Chapter 3, returns with considerable force. I examine first the substantive model defended by Hegel, who, in common with some contemporary theorists, seeks to defend a socially embedded conception of agency. I argue that what is gained in terms of substance is lost in terms of morality and also in terms of freedom. On the other hand, the Kantian conception of freedom relies on the obscure and controversial concept of a 'causality of reason'. So, drawing on earlier discussions from Chapters 3 and 5, I show how the concept can be understood as asserting agential control. I conclude with a discussion of the 'scope' of autonomy, arguing that appreciation of its scope is essential if we are properly to recognize the importance of autonomy to our moral lives.

Acknowledgements

List of Abbreviations

Arabic numerals below refer to volume numbers of the *Akademie* edition, *Kants gesammelte Schriften: Herausgegeben von der Deutschen Akademie der Wissenschaften* (formerly *Königlichen Preussischen Akademie der Wissenschaften*), 29 vols (Berlin and Leipig: Walter de Gruyter (formerly Georg Reimer), 1902–). References in the text give volume and page numbers of this edition. The list of English translations used or consulted is given in the Bibliography.

A 7	*Anthropology from a Pragmatic Point of View*
E 7	'An Answer to the Question: "What Is Enlightenment?"'
G 4	*Groundwork of the Metaphysics of Morals*
KpV 5	*Critique of Practical Reason*
KU 5	*Critique of the Power of Judgement*
L 9	*Logic (Jäsche)*
LR 28	*Lectures on the Philosophical Doctrine of Religion*
M 6	*Metaphysics of Morals*
O 8	'What Is Called Orientation in Thinking'
R 6	*Religion Within the Limits of Reason Alone*
U 8	'Idea for a Universal History from a Cosmopolitan Perspective'

For the *Critique of Pure Reason* the usual reference to the A/B editions and pagination is used.

Contents

1

Introduction. Autonomy: Specification of a Term, Recognition of a Problem

Autonomy gives expression to the idea that we can lead our lives rather than suffer them. This liberating sense of purpose, however, comes perilously close to self-exaltation. As its critics point out, autonomy fosters a deceptive and treacherous view of the self as 'man-god . . . free, independent, lonely, powerful, rational, responsible, brave, the hero of so many novels and books of moral philosophy' (Murdoch 1970: 80).[1] The concern Iris Murdoch voices here is that even if such a way of life were a realistic possibility—rather than a philosophical or novelistic fantasy—it does not seem in any obvious way to be a *moral* way of life. Certainly, if we are to be held accountable for the things we do, we need to be able to hold on to the belief that we are 'autonomous' in the general sense Thomas Nagel identifies, namely 'the sense that we are authors of our actions' (Nagel 1986: 168).[2] And yet, it is just this identification of autonomy with authorship that, some would argue, gives us the wrong idea about ourselves, leading us to forget our finitude and the multiple contingencies and

---

[1] See too 'Fact and Value' in Murdoch 2003: 25 ff. and O'Neill 1989: 75–7. More recently, Seiriol Morgan has shown how this Luciferian conception of free agency actually forms the basis for Kant's argument about evil: the evil will, on Morgan's reconstruction of Kant's argument, 'yearns for the kind of freedom and power possessed by a very different kind of will, the infinite unlimited will of God' (Morgan 2005: 85). In the non-Kantian literature the problem raised by Murdoch is addressed in, among others, Dworkin 1976 and Feinberg 1986; I return to these debates in the context of discussing substantive and relational theories of personal autonomy in Ch. 6.

[2] Nagel raises this type of autonomy as a problem. Others seek to show that autonomy for the purposes of ascribing moral responsibility is tenable without being overly demanding (Mele 1995: ch. 13) or that the link between autonomy and moral responsibility is questionable (McKenna in Taylor 2005: 205–34).

dependencies that shape our lives.[3] Even if one grants the importance of authorship in connection with accountability, its value as a distinct moral notion and aspiration is not immediately perspicuous. This is because, however gratifying or comforting it is to think of oneself as the author of one's actions, it is not clear what moral authority the autonomous agent claims for her choices or indeed why such choices deserve to be thought of as moral.

A standard way of addressing such doubts, and also of answering Murdoch's charge, is by giving reason a defining role in the autonomy-based ethics: it is our rational self we obey when we act autonomously. Self-direction is morally desirable because steering our conduct is reason itself. This solution is not without its own difficulties. Questions can be raised about how exactly reason is supposed to perform its steering role or indeed about the desirability of tying morality so closely with rationality. The familiar spectre of grim dutifulness lurks in reconstructions that show the autonomous agent moved to action by the thought of duty alone; such an agent can appear 'alienated', as lacking the human responses we value in each other (Railton 1984: 94).[4] The task of this book is to examine how a theory of autonomy can be formulated that captures both our capacity for self-determination, identified in Nagel's sense of authorship, and our responsiveness to reasons—in particular, moral reasons.

Current views of personal autonomy tend to emphasize the psychological features of authorship: agents possess autonomy with respect to their desires and motivations. On this model, inspired by early work by Harry Frankfurt and Gerald Dworkin, an agent is autonomous provided she endorses a desire or motivation.[5] The attractions of this conception of

---

[3] The critics of the autonomy as authorship view are many. Classical objections can be found in Sandel 1984 and Taylor 1989. See also Taylor 1991, Menke 2000, Friedman 2003. Responses include Benson 1991 and Christman 1995.

[4] See also Williams 1985 and Stocker 1976; I discuss Williams in Ch. 4.

[5] The non-Kantian models of autonomy current in the relevant literature tend to focus on motivation. Whatever their other philosophical commitments (which range from Platonic to Lockean), what matters is psychology and to that extent they fit a broadly naturalistic philosophical framework. Motivation, in the sense of promoting an ideal of coherence among one's basic practical commitments, usually expressing a hierarchical model is relevant in Dworkin 1976 and Frankfurt 1988; also Benson 1983: 5–17 and Bratman 2007; and also Watson 2009. At a further remove, because of his concern with practical reflection and the way practical reflection is tied to the free development of a conception of the good, is Raz 1994: esp. 78–9. I examine contemporary views on personal autonomy in Ch. 6.

autonomy are its consistency with a widely held view of what it is to be a person, its naturalism, and its neutrality with respect to the good.[6] What has proved remarkably resistant to analysis, however, is the 'self' whose autonomy is asserted in particular instances. As critics point out, a desire I have been manipulated into endorsing cannot count as my 'own'. But then the acceptable range of what is to count as my 'own', as against what is perniciously 'external' or 'alien', becomes difficult to define, given that very often we endorse and reject motivations in the context of getting advice, of being educated into different ways of thinking, of being persuaded, and so forth. Another dimension of the problem of identifying what is truly my 'own' regards pinpointing the authorizing higher order desire that ensures that endorsements of first-order desires count as autonomous. The move up the hierarchy in search for such an authorizing desire confronts the familiar problem of regress. Indeed, more needs to be said about the authority of the authorizing desire, about why it has a claim on the agent so that it should be treated as reason-giving. One obvious response is that such or such desire is autonomy-promoting, but, well apart from its incipient circularity, this claim begs the question about the value of autonomy. In recent years, these pressures have led the discussion of personal autonomy into new directions characterized by attempts to forge links between the basic psychological model and moral responsibility or responsiveness to reasons, or in more radical proposals, by the introduction of substantive normative concerns on the grounds that it is not incompatible with autonomy to place restrictions on the sorts of preferences the agent may endorse. The pressures that have led to the exploration of these theoretical options can also be taken as indicating the need for a more radical rethinking of the concept of autonomy. This is what I seek to do in this book, using the resources that are available in Kant's formulation of the practical ideal of autonomy.

The view I defend here takes as its starting point 'the idea of *the will of every rational being as a universally legislating will*' (G 4:431). This formulation is the closest we come to a definition of autonomy. To grasp fully the idea contained in this formulation, however, we need to identify—and justify—the commitments that are operative in four distinct but related fields. (1) In the domain of knowledge, to be autonomous requires that we are able

---

[6] See Taylor 2005: 1–2.

to work out *for ourselves* what it is right to do. (2) With respect to motivation, attribution of autonomy entails that we want to do the right thing *because it is right*. (3) The third field is best described as 'metaphysical' since it contains the claim that freedom should be understood in terms of a capacity to respond to rational practical principles. In other words, autonomy is not just about the kind of motivation people have or the things they count as reasons: it is also about the conception of freedom that needs to be assumed for these things to carry the weight they do. It is a central contention of this book that psychology alone cannot provide the basis for the robust conception of agency that underpins autonomy. Although metaphysics is rejected by many who write on this topic, an address to metaphysical issues is necessary if we are to make sense of the cognitive and motivational commitments that autonomy requires. (4) The fourth field is normative ethics. Central to the claim that autonomy is a moral ideal—and not, for example, an ideal of self-fulfilment—is the argument that autonomy describes an intersubjective rather than intra-subjective norm.[7] The moral content of autonomy is given by the notion of the law (*nomos*) and the demand that one think of oneself in relation to others under a shared law. I show that autonomy requires that we engage in a perspectival ascent from a subjective viewpoint to one that putatively embraces all rational others. This interpretation casts new light on the links between Kant and his successors, especially Schiller and Hegel, both of whom acknowledge their debt to Kant's moral thinking while arguing for the need to develop a stronger account of our intersubjective commitments.

In addition to these links with the post-Kantian tradition, other relevant references are to the early history of the concept. Besides the Pauline reference to the Gentiles as being a 'law unto themselves', perhaps most familiar is the political ideal of self-determination that directly relates to the original sense of *autonomia*.[8] *Autonomia* in the classical sense refers to a claim

---

[7] The distinction I use here between self-fulfilment and a moral ideal does not preclude that pursuing a moral ideal may not be self-fulfilling. So I do not assume self-fulfilment to be merely a matter of gratification. However, I also think that we can and may differentiate conceptually between self-fulfilment—however broadly understood—and morality; see also Ch. 5. For a morally rich conception of self-fulfilment see Gewirth 1998. However, Gewirth follows the usual Kantian path of seeking to deduce moral identity from some basic agential identity, a strategy which I think fails; see n. 12.

[8] The reference to Paul, from Romans 2:14, is frequently cited but its sense is far from clear. See Martin 2009.