KANT ON THE POSSIBILITY OF ACTION FROM DUTY BUT NOT IN ACCORDANCE WITH DUTY

by

GEORG URICH

Submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy
at Trinity College, Dublin University, in the Department of Philosophy

June 2010

DECLARATION

This thesis has not been submitted as an exercise for a degree at this or any other University or any other degree awarding body. This thesis is entirely my own work.

The librarian of Trinity College, Dublin may lend or copy the thesis upon request and without reference to me on the understanding that such authority applies to the provision of single copies made for study purposes, subject to normal conditions of acknowledgement.

Signed: ..... Date: 8/6/2010

Meinen Eltern

CONTENTS

PART II.

MISTAKEN MORAL JUDGEMENT AND THE POSSIBILITY OF ACTION FROM DUTY BUT NOT IN ACCORDANCE WITH DUTY

ACKNOWLEDGEMENTS

First and foremost I am indebted to my supervisor, Dr. Vasilis Politis. His seminar on Kantian ethics many years ago first awoke my interest in issues which have occupied me ever since and some of which are central to this study. I thank him also for taking me on as a research student, and especially for the continuing patience and loyalty – beyond the call of duty – he showed towards me.

I would like to thank Professor David Berman and Dr. Lilian Alweiss for reading and commenting on earlier drafts of several chapters and for their help and encouragement.

Thanks are also due to members of the TCD Postgraduate Philosophy Forum for their constructive criticism of two papers I presented there, and to attendees at the joint UCD-TCD Dublin Graduate Philosophy Work-in-Progress Seminar, especially Dr. Timo Jütten, who responded to a presentation of much of the penultimate draft of part II. with a number of helpful suggestions.

This thesis was submitted on 30 October 2009 and defended on 1 March 2010. I thank the examiners, Dr. Jens Timmermann of the University of St Andrews and Dr. James O'Shea of University College Dublin, for their very thorough, insightful, and helpful comments and suggestions. Some of these are reflected in this slightly revised version of the thesis, while others will no doubt prove valuable for my future work in the field of Kantian ethics.

Finally, I am deeply grateful to Julie Cruickshank for her love and support and for relieving me of my child-rearing duties during crucial periods of writing.

This thesis is dedicated to my parents.

Georg Urich

Dublin, June 2010

SUMMARY

The central theme of this study is a curious and controversial feature of Kant's account of acting from duty in the *Groundwork*: his apparent omission of the possibility of action from duty but not in accordance with duty.

The aims of part I. are largely preparatory. Chapter 1 explains the importance of Kant's twofold division of the concept of duty and situates it in the context of Kant's ethical theory. Chapter 2 examines the pervasiveness – both in a historical and a philosophical sense – of Kant's distinction. I argue that in its most basic form the distinction itself and the criterion of morality which it gives rise to can be found in both Aristotle's and Mill's moral theory, that is, in the two main alternatives to a Kantian approach to ethics. Chapter 3 is an examination of Kant's concept of a good will, which the twofold distinction is said to explicate. I identify in this Kantian notion two related principles, the Control Condition and the Effort Condition, which, like Kant's twofold distinction itself, are of crucial importance for the question concerning the possible moral worth of some actions from mistaken moral judgement which I consider in the final chapter.

The investigation of the possibility of action from duty but not in accordance with duty is the subject of part II. and constitutes the main business of this study. Chapter 4 outlines in some detail the background to the problem and its origin in Kant's twofold division. I identify two important assumptions common amongst Kant scholars, the Logical Compatibles View and the Identity Thesis, which taken together give rise to the impression of a missing type of action in Kant's account of duty. This apparent lacuna is generally taken to imply the denial of the possibility of action from mistaken moral judgement, or at least the denial of the moral worth of such actions. Chapter 5 considers three suggestions which might be thought to explain why Kant never mentions, let alone discusses, actions from duty yet contrary to it. I critically examine and eventually reject all three. In the following three chapters (6-8) I discuss in considerable detail the textual evidence as well as the various interpretations available in the literature, before offering an alternative reading. Chapters 8 and 9 argue for what I believe to be a novel and coherent interpretation of both the text and the underlying issues which separates entirely the question concerning the possibility of action from duty but contrary to it from the issue of moral error, and which also has certain implications for a number of

related issues in Kantian moral theory. I conclude that the notion of 'action from duty but not in accordance with duty' is, in the last analysis, incoherent.

Chapter 10, finally, links back with some of the concepts developed in part I. and outlines – in accordance with the main findings of this study – the parameters of the continuing debate. The investigation ends on an aporetic note by suggesting that Kant's control and effort conditions may be appealed to both to reject and to support the view that Kant should allow that some actions from mistaken moral judgement can have moral worth.

ABBREVIATIONS AND TRANSLATIONS OF KANT'S WORKS

All references to and quotations from Kant's works are by volume and page number to *Kant's gesammelte Schriften, herausgegeben von der Königlichen Preussischen* (later *Deutschen*, then *Berlin Brandenburg*) *Akademie der Wissenschaften*, 29 volumes (Berlin: Georg Reimer, later Walter de Gruyter & Co., 1900-), the *Akademie Ausgabe* (Ak.) of Kant's works, which appear in the margins of most translations. References to the *Critique of Pure Reason* are to the first (A) edition of 1781, or the second (B) edition of 1787 as is customary. Translations of Kant's works are listed below and are referred to by means of the relevant abbreviation, followed by the Akademie edition volume and page numbers.

In the case of the *Groundwork* I have consulted Paton's and Ellington's translations, but generally follow Wood's. On occasion I specify the exact line(s) in Kant's text to which I am referring; thus IV:397.11 points to line 11 on page 397 of volume IV of the Akademie edition. When the wording of a passage does not follow that of the relevant translation listed below, I indicate that the translation is my own.

KrV **Kritik der reinen Vernunft** (A: 1781, B:1787)

Critique of Pure Reason
transl. N. Kemp-Smith, London: Macmillan, 1929

GMS **Grundlegung zur Metaphysik der Sitten** (1785, Ak. IV:385-463)

Groundwork for the Metaphysics of Morals
ed. & transl. Allen W. Wood, New Haven and London: Yale University Press, 2002

Other translations consulted:
Groundwork of the Metaphysic of Morals
transl. H.J. Paton, New York: Harper & Row, 1956 (3rd ed.)
Grounding for the Metaphysics of Morals
transl. James W. Ellington, Indianapolis/ Cambridge: Hackett, 1981 (3rd ed)

KpV **Kritik der praktischen Vernunft** (1788, Ak. V:1-163)

Critique of Practical Reason
in: *The Cambridge Edition of the Works of Immanuel Kant: Practical Philosophy*,
ed. & transl. Mary J. Gregor, Cambridge: Cambridge University Press, 1996,
pp. 133-271

Rel. **Die Religion innerhalb der Grenzen der bloßen Vernunft** (1793, Ak. VI)

Religion within the Boundaries of Mere Reason
in: *The Cambridge Edition of the Works of Immanuel Kant: Religion and Rational Theology*, ed. & transl. Allen W. Wood & George di Giovanni, Cambridge: Cambridge University Press, 1996

T&P **Über den Gemeinspruch: Das mag in der Theorie richtig sein, taugt aber nicht für die Praxis** (1793, Ak. VIII)

On the common saying: That may be correct in theory, but it is of no use in practice
in: *The Cambridge Edition of the Works of Immanuel Kant: Practical Philosophy*, ed. & transl. Mary J. Gregor, Cambridge: Cambridge University Press, 1996, pp. 277-309

MS **Metaphysik der Sitten** (1797, Ak. VI:203-493)

The Metaphysics of Morals
in: *The Cambridge Edition of the Works of Immanuel Kant: Practical Philosophy*, ed. & transl. Mary J. Gregor, Cambridge: Cambridge University Press, 1996, pp. 353-603

LE ***Lectures on Ethics*** (Ak. XXVII/ XXIX)

The Cambridge Edition of the Works of Immanuel Kant, ed. Peter Heath & J.B. Schneewind, transl. Peter Heath, Cambridge: Cambridge University Press, 1997

KANT ON THE POSSIBILITY OF ACTION FROM DUTY
BUT NOT IN ACCORDANCE WITH DUTY

PART I.

KANT, ARISTOTLE, AND MILL ON MORAL MOTIVATION

Chapter 1

KANT'S TWOFOLD DIVISION OF THE CONCEPT OF DUTY

In the preface to the *Groundwork* Kant insists that the sole purpose of that work is "nothing more than the search for and establishment *of the supreme principle of morality*" (GMS IV:392), and, as we know, at the end of the first section he does indeed identify this principle in the universal law formula of the Categorical Imperative (GMS IV:402). Between the statement of intent in the preface and the discovery of the supreme principle lies the famous argument of the first section of the *Groundwork*. Beginning with the identification of the good will as the only thing which is good without limitation – a concept which "dwells already in the naturally healthy understanding" (GMS IV:397) – Kant proceeds in the first section by analysing in philosophical terms what he takes to be the basic assumptions of everyday morality, or "common rational moral cognition" (GMS IV:392). The concept of a good will is then explicated and developed in an analysis of the concept of duty, before, via a set of propositions regarding the role of maxims and the moral feeling of respect for the moral law, which according to Kant is the "immediate determination of the will through the law and the consciousness of it" (GMS IV:401, fn.), that is, a rational feeling which is "self-effected through a concept of reason" (ibid.), he identifies the universal law formula of the Categorical Imperative.

However, in the process of searching for and establishing the supreme principle of morality, Kant develops several important concepts which are central to his moral theory in general, and the significance of which is therefore not limited to their function within that search. Traditionally, the theory of moral motivation which constitutes much of Kant's argument for and derivation of the first formula of the Categorical Imperative has been considered equally as important in its own right as the establishment of the supreme principle of morality itself; indeed, even a brief survey of the literature might give the impression that it is in fact the former which is Kant's most significant and original contribution to moral theory.[1] The comparative importance in Kant's moral

[1] This is not uncontroversial. The first section of the *Groundwork* has led to an amount of criticism and commentary which is out of all proportion to its size, and according to some, is even inappropriate if it leads to the above-mentioned impression. Thus most recently Allen Wood has insisted repeatedly that "the

philosophy of the value of the rational nature of human beings and the lawfulness of their actions on the one hand, and his theory of moral worth and moral motivation on the other, is an interesting question in its own right. However, my primary concern in this thesis is with one of the central concepts Kant develops in the first section of the *Groundwork* – the concept of duty – and in particular with Kant's notion of acting *from duty* and the moral worth such actions are said to have. I will begin by outlining that distinction in its most familiar guise and in the most general terms possible, before further clarifying it in the following chapters.

As early as in the preface to the *Groundwork*, Kant insists that morality, even in its popular form, is concerned with actions insofar as they exhibit a particular quality or value which is uniquely ethical. In the context of explaining the need for moral philosophy, and in particular the need for a moral philosophy which is *pure* in the sense that it gives to the human being, *qua* rational being, laws *a priori* (cf. GMS IV:389) Kant maintains that

> a metaphysics of morals is indispensably necessary not merely from a motive of speculation, in order to investigate the source of the practical principles lying *a priori* in our reason, but also because morals themselves remain subject to all sorts of corruption as long as that guiding thread and supreme norm of their correct judgement is lacking. For as to what is to be morally good, it is not enough that it *conform* to the moral law [dem sittlichen Gesetze *gemäß* sei], but it must also happen *for the sake of this law* [*um desselben willen*]; otherwise, that conformity is only contingent and

---

first section of the *Groundwork* gives us only very limited information about Kantian ethics", and that "the two chief elements of Kant's argument in these pages – the unlimited goodness of the good will and the moral worth of acting from duty – although (properly understood) they belong to Kantian ethics, are not particularly central or fundamental to it". See his *Kantian Ethics* (Cambridge, Cambridge University Press, 2008), p. 41f. On the other hand there are those who not only consider the first section of the *Groundwork* to contain some of Kant's best writing (cf. Harald Köhl, *Kants Gesinnungsethik* (Berlin: de Gruyter, 1990); Ernst Tugendhat, *Vorlesungen über Ethik* (Frankfurt am Main: Suhrkamp, 1993)), but who also argue that it presents a stringent argument from the common notion of the good will via the analysis of duty, maxims, and respect for the law, right through to the first statement of the supreme principle of morality (cf. Samuel Kerstein, *Kant's Search for the Supreme Principle of Morality* (Cambridge: Cambridge University Press, 2002)).

> precarious, because the unmoral ground will now and then produce lawful actions, but more often actions contrary to the law. (GMS IV:389f.)

This distinction between an action's *conformity with* the moral law on the one hand, and its being done *for the sake of* the moral law on the other, is the first and most basic formulation in the *Groundwork* of the central theme of this thesis: Kant's twofold division of the concept of duty, that is, his famous distinction between *action in accordance with duty* and *action from duty*. In the cited passage Kant gives several clues as to why, and in what sense, this distinction might be thought to be significant and why a metaphysics of morals – and a groundwork for it – are necessary: There is, of course, the philosopher's "motive of speculation" (ibid.), and from this point of view alone an investigation into the grounds of morality ("the source of the practical principles lying *a priori* in our reason", ibid.) is needed. But furthermore – and perhaps more importantly – Kant thinks that such an investigation is required from a *practical* point of view, "because *morals themselves* [die Sitten selber]" (ibid., my emph.), that is, human conduct, "remain subject to all sorts of corruption as long as that guiding thread and supreme norm of their correct judgement is lacking" (ibid.). This practical need is grounded in the very distinction between acting merely in accordance with duty and acting from duty, for without such a distinction – and more specifically, without the action's being performed from an appropriate moral motive ("for the sake of this law", ibid.) – Kant thinks that moral *conduct itself* will be corrupt. The reason he adduces for this last claim is that actions done merely in conformity with duty but not from duty are only *contingently* good because our performing them depends on our having a motive other than duty to perform the action.[2] Thus the practical relevance of the distinction lies in the fact that, according to Kant, in the absence of the motive of duty the accordance with the moral law of an agent's actions becomes *arbitrary* and *unreliable*, occasionally leading to actions which are in fact in accordance with the moral law, but, not merely on occasion, but "*more often*" (ibid.), to actions which are contrary to it. Kant's insistence on actions which are done from the right sort of motive thus has important *practical* implications in that only actions

[2] Cf. Jens Timmermann, *Kant's Groundwork of the Metaphysics of Morals – A Commentary* (Cambridge: Cambridge University Press, 2007), pp. 26-32.

done from the right motive are likely be *right* actions, that is to say, actions which constitute good *conduct*.

On the other hand, there is also present in Kant's argument in the *Groundwork* a more general concern with the sort of things we *value*, and in that context Kant's point is a related but slightly different one: a value which is uniquely ethical is to be found only in actions which are done from the right sort of motive. For an action to have this particular value, or *moral worth*, it must not only be the right sort of action, but it must also be performed from a particular moral motive, that is, for the right sort of reason. We will see in chapter 3 that at the most fundamental level the basis for this claim is that a right action done because it is right is the most common way in which a human being may exhibit a good will, which Kant identifies as the only unqualified good.

Three points can be taken from the above which are worth reiterating: Firstly and most obviously, Kant's distinction itself is of central importance, an its precise nature and pervasiveness is the main subject of this study. Secondly, it is worth bearing in mind that according to Kant, quite apart from and in addition to the philosopher's speculative interest in an analysis and grounding of the concept of duty, we take a *practical* interest in the distinction between, generally speaking, an action's mere conformity with a principle or rule, and its being done from the recognition of the value of that rule, that is, its being a case of *principled* action. Also, Kant believes – and this is the working assumption of the first, if not the first two, sections of the *Groundwork* – that the importance of acting *from duty* is already evident "from the common idea of duty and of moral laws" GMS IV:389). Thus his aim there is primarily to explicate "the moral cognition of common human reason" (GMS IV:403), that is, to spell out in philosophical terms the implications of everyday morality.

My aim in this first chapter is to outline Kant's twofold division of the concept of duty both as a substantial claim in its own right and in the most general terms, and – as much as this is possible with a philosopher as systematic as Kant – in isolation from the various other elements of his moral theory. The rest of part I. will be concerned with clarifying Kant's distinction, firstly, in relation to a number of other Kantian concepts, in particular the good will, and secondly by drawing certain parallels with Aristotle's and Mill's ethics. Only in part II. will I consider the types of actions the distinction gives rise

to, and in particular one troubling feature of it: the apparent impossibility of action from duty but not in accordance with duty.

1. Kant's normative criterion: Rightness

According to Kant's analysis of the common notion of duty in the first section of the *Groundwork*, all actions are either *right* (obligatory or permissible)[3], namely when they are in accordance with duty, or *wrong* (impermissible) when they are contrary to duty. Whether an action is in accordance with duty is evaluated by use of the Categorical Imperative[4], which is intended to enable the agent to recognise what his duty is in a

---

[3] Kant divides right actions into those that are permissible and those that are obligatory. But he also holds that to act from duty is to act from the notion that the action is morally required. Therefore, merely permissible actions cannot be done from the motive of duty and hence cannot have moral worth. Since, as will become apparent shortly, my primary concern is with the possibility of acting *from duty*, I will henceforth ignore merely permissible actions. (Kant does so himself in the analysis of the notion of acting from duty in *Groundwork I.*: All the examples he adduces are of actions which are morally required and the omission of which is impermissible. See GMS IV:390, 397ff. Cf. Henry Allison, *Kant's Theory of Freedom* (Cambridge: Cambridge University Press, 1990), pp. 113, 266 n.15; Samuel J. Kerstein, *Kant's Search for the Supreme Principle of Morality* (Cambridge: Cambridge University Press, 2002), p. 99; Steven Sverdlik, 'Kant, Nonaccidentalness and the Availability of Moral Worth', *The Journal of Ethics* 5 (2001) p. 297.
For the same reason, actions which are performed above and beyond what duty requires (supererogatory acts) cannot be done from duty, simply because they are not obligatory. In the *Critique of Practical Reason* Kant is unashamedly dismissive of such actions: "But I do wish that educators would spare their pupils examples of so-called *noble* (supermeritorious) actions, with which our sentimental writings so abound, and would refer everything only to duty and to the worth that a human being can and must give himself in his own eyes by consciousness of not having transgressed it; for, whatever runs up into empty wishes and longings for inaccessible perfection produces mere heroes of romance (Romanhelden) who, while they pride themselves on their feeling for extravagant greatness, release themselves in return from the observance of common and everyday obligation, which then seems to them insignificant and petty." (KpV V:155)

[4] I will set aside for the purposes of this study all questions concerning Kant's derivation of the Categorical Imperative, its various formulas, its application in particular maxims and actions, which actions it prescribes or condemns, etc. For my purposes it is sufficient to note that with the Categorical Imperative Kant provides at the very least a criterion for the *rightness* of actions.

particular context by universalizing the maxim of his action. The Categorical Imperative is Kant's objective[5], *normative* criterion for what is morally required and permissible. If a maxim can be universalised without contradiction, then an action conforming to it is objectively good (or: in accordance with the law), and, if performed, the action is (at least) done in accordance with duty: "The concept of duty requires of an action, objectively, accordance with the law." (KpV V:81; cf. GMS IV:397ff.) "[W]ith this compass in its hand", which "common human reason [...]always actually has before its eyes and uses as the standard of its judgement", Kant believes, it "knows very well in all the cases that come up to distinguish what is good and what is evil, what is in conformity with duty or contrary to duty". (GMS IV:403f., my transl.) Thus the Categorical Imperative, as a criterion of the *rightness*, and only the rightness[6] of an action's maxim, is concerned with the outward conformity of an action with the moral law. In the *Metaphysics of Morals* Kant puts the point in the following way:

> A principle that makes certain actions duties is a practical law. A rule that the agent himself makes his principle on subjective grounds is called his *maxim*; hence different agents can have very different maxims with regard to the same law.
>
> The categorical imperative, *which as such only affirms what obligation is*, is: act upon a maxim that can also hold as a universal law. – You must therefore first consider your actions in terms of their subjective principles; but you can know whether this principle also holds objectively only in this way: that when your reason subjects it to the test of conceiving yourself as also giving universal law through it, it qualifies for such a giving of universal law. (MS VI:225)

---

[5] See part II. for a discussion of the difficult question concerning the possibility of distinguishing between action that is objectively right and action that the agent mistakenly believes to be objectively right.

[6] Nico Scarano and Otfried Höffe take the Categorical Imperative to be the criterion of both the action's rightness and its morality, although they fail to argue for this claim. Cf. Nico Scarano, 'Necessity and Apriority in Kant's Moral Philosophy: An Interpretation of the Groundwork's Preface', in: Christoph Horn & D. Schönecker (eds.): *Groundwork for the Metaphysics of Morals* (Berlin: de Gruyter, 2006), p. 8 (fn. 5); Otfried Höffe, *Immanuel Kant* (München: C.H. Beck, 1992), pp. 76-86. It seems clear that while actions based on maxims which conform with the Categorical Imperative, when performed from duty, have moral worth, the Categorical Imperative itself cannot be used to establish whether an action has been done from duty and thus has moral worth.

He goes on to recast the familiar *Groundwork* distinction in the terms of an action's *legality* and its *morality* which he had first introduced in the *Critique of Practical Reason*.

> The conformity of an action with the law of duty is its *legality* (*legalitas*); the conformity of the maxim of an action with the law is the *morality* (*moralitas*) of the action. A *maxim* is a *subjective* principle of action, a principle which the subject himself makes his rule (how he wills to act). A principle of duty, on the other hand, is a principle that reason prescribes to him absolutely and so objectively (how he *ought* to act). (MS VI:225)[7]

But rightness alone is not nearly sufficient for actions to have any genuinely moral quality. Action merely in conformity with duty, "however it may conform to duty, and however amiable it is" (GMS IV:398) lacks a uniquely ethical value. Despite its rightness, therefore, such action "nevertheless has no true moral worth" (ibid.). Most importantly in the context of the *Groundwork*, where Kant's aim is to explicate the concept of a good will, rightness alone does not allow us to do this, because by itself it is not concerned with the will which can be good or bad, i.e. with the inner motivation of the agent. So in order to determine what distinguishes a good will, a further distinction must be drawn (and according to Kant is in fact always drawn by common human reason).

2. Kant's motivational criterion: Moral worth

Since Kant believes that neither the good will's actions and nor its ends can account for the good will's goodness, he argues that only volition itself can: moral goodness lies in the will's reasons for willing (choosing) certain ends and actions, insofar as these reasons motivate the agent to act. Common human reason already draws this distinction by separating merely *what* someone does from *why* they do it and conferring different values on each (and, generally speaking, as Kant repeatedly insists, a *higher* value on the latter

[7] Cf. MS VI:224, where he defines rightness as accordance with duty vis-à-vis actions (deeds), rather than maxims. "A deed is *right* or *wrong* (*rectum aut minus rectum*) in general insofar as it conforms with duty or is contrary to it (*factum licitum aut illicitum*); the duty itself, in terms of its content or origin, may be of any kind."

when an agent acts from the right sort of reason)[8]. So the agent's motive for acting is where 'true moral value' is to be found. While right actions may be done for various reasons, that is, from different motives, the good-willed agent not only does what is right, but does it *because* it is right: the agent's recognition that an action is required motivates his acting. Hence such actions are done *from* (the motive of) duty, and actions have moral worth if, and only if, they are done from duty.

In cases of action which merely coincides with duty, Kant, as we saw earlier, believes that the agent's motivation to act is incidental or even alien to the action itself and to its rightness. It is one of Kant's deeply held convictions that in such cases the agent is not morally praiseworthy. For the action to have *moral worth*, the rightness of the action and the agent's acting would have to be connected in some non-accidental way. In general terms this means that the action's rightness itself, or some part of it, or something necessarily connected with it, functions as the agent's motivation to perform the action.[9] So in addition to the criterion of rightness, Kant insists on a further, *subjective* and *motivational*, condition: While it might be the case that the objective rightness of an action can be read off actions or their effects, this criterion is *subjective* insofar as it applies to an act only *qua* act of an *agent*.

The following passages from the *Metaphysic of Morals* are instructive. Kant here again makes the point already insisted on in the *Groundwork*, namely that ethics is concerned with actions insofar as they are done from the right sort of motive, but perhaps more explicitly: while all lawgiving is concerned with duties and with motives (incentives) to comply with those duties, by its very nature ethical lawgiving, when it makes an action a duty, *also makes this duty the incentive*.

---

[8] See, for example, GMS IV:398 f., where the naturally unsympathetic and therefore unfortunate "friend of man" is said to be able to "find a source within himself to give himself a *far higher* worth than that [of] a good-natured temperament" (emph. added) whose beneficent acts are merely in conformity with duty. "[W]orth of character, which is *moral* and *the highest without any comparison*" (emph. added), begins when he acts *from duty*.

[9] Kant's own notion of the moral feeling of *reverence for the law* fulfils this role: its presence and motivating force guarantees that the agent's action has moral worth. It thus serves as a non-accidental nexus between the agent's recognition of what is right and his acting upon that recognition.

> In all lawgiving (whether it prescribes internal or external actions, and whether it prescribes them a priori by reason alone or by the choice of another) there are two elements: **first** [Kant's emphasis], a law, which represents an action that is to be done as *objectively* necessary, that is, which makes the action a duty; and **second** [Kant's emphasis], an incentive, which connects a ground for determining choice to this action *subjectively* with the representation of the law. Hence the second element is this: that the law makes duty the incentive. By the first the action is represented as a duty, and this is a merely theoretical cognition of a possible determination of choice, that is, of practical rules. By the second the obligation so to act is connected in the subject with a ground for determining choice generally.
>
> All lawgiving can therefore be distinguished with respect to the incentive (even if it agrees with another kind with respect to the action that it makes a duty, e.g., these actions might in all cases be external). That lawgiving which makes an action a duty *and also makes this duty the incentive* is *ethical.* But that lawgiving which does not include the incentive of duty in the law and so admits an incentive other than the idea of duty itself is *juridical.* It is clear that in the latter case this incentive which is something other than the idea of duty must be drawn from *pathological* determining grounds of choice, inclinations and aversions, and among these, from aversions; for it is a lawgiving, which constrains, not an allurement, which invites.
>
> The mere conformity or nonconformity of an action with law, irrespective of the incentive to it, is called its legality (lawfulness); but that conformity in which *the idea of duty arising from the law is also the incentive to the action* is called its morality. (MS VI:218f., emphasis partly mine)

> Ethical lawgiving (even if the duties might be external) is that which *cannot* be external; juridical lawgiving is that which can also be external. So it is an external duty to keep a promise made in a contract; but the command to do this merely because it is a duty, without regard for any other incentive, belongs to *internal* lawgiving alone. So the obligation is assigned to ethics not because the duty is of a particular kind (a particular kind of action to which one is bound) – for there are external duties in ethics as well as in right – but rather because the lawgiving in this case is an internal one and can have no external lawgiver. For the same reason duties of benevolence, even though they are external duties (obligations to external actions), are still assigned to ethics because their lawgiving can be only internal. – Ethics has its special duties as well (e.g., duties to oneself), but it also has duties in common with right; *what it does not have in common with right is only the kind of obligation.* For what is distinctive of ethical lawgiving is that one is to perform actions *just because they are duties and to make the principle of duty itself, wherever the duty comes from, the sufficient incentive for choice.* So while there are many *directly ethical* duties, internal lawgiving makes the rest of them, one and all, indirectly ethical. (MS VI:220f., emphasis partly mine)

By way of summary and because it is of crucial importance for the various issues which I will consider in both parts of this study, let me reiterate that Kant's account of duty rests

on this *twofold division*: Firstly an action is either right (obligatory), namely when it conforms to the requirements of morality (such actions are said to be '*in accordance with duty*'), or it is wrong (impermissible) when it does not (Kant refers to these actions as '*contrary to duty*', or 'not in accordance with duty'). Secondly, an action either has moral worth, namely when it is done *because* it is recognised as required ('*from duty*'), or it lacks such worth when it is done from some other motive, typically a desire or inclination ('*not from duty*').

3. The aims of the remainder of part I. and the confines of this study

As readers of Kant's writings often find to their frustration, his ethical theory contains many intricately interwoven claims regarding the nature of morality, the unconditionality, universality, and apriority of the moral law, the nature of freedom, the phenomenal and noumenal standpoints, etc. To make matters worse, it is often impossible to say where one claim ends and another begins and whether, or how, they could possibly be evaluated on their own terms. Despite this, I believe that Kant's twofold distinction of the concept of duty, and the resulting claim concerning the importance of the agent's motives (the criterion of morality), can be isolated from the other elements of his moral theory and evaluated separately. The aims of the remainder of part I. of this study are the following:

I will first attempt to show that Kant's twofold distinction is noteworthy – perhaps this even makes it unique amongst central concepts of Kantian moral theory, although I will not argue for this claim – in at least two important respects: firstly, it is, historically, far from an original Kantian notion, and secondly, it can be shown to be quite compatible with rival ethical theories. It will be the business of chapter 2 to illustrate that both Aristotle's virtue ethics and Mill's utilitarianism, albeit in very different ways, are compatible with or even explicitly embrace the idea that something like the Kantian distinction is central to the very nature of morality.

In chapter 3 I will attempt to illuminate the origins, significance, and implications of the Kantian distinction beyond the basic concept outlined above by drawing on a number of central Kantian themes: the good will, moral worth, and Kant's so-called

control condition. Of course, these notions easily merit a detailed study in their own right. I will consider them here only insofar as doing so serves the purposes of clarifying Kant's distinction with respect to the confines of this study.

Finally, in part II. of this study I will turn to some of the particulars of Kant's distinction. My main concern is with the question whether Kant's distinction is one between two logically compatible features which can be evaluated independently of each other, or whether, on the contrary, a certain asymmetry and dependence between the criteria is to be found which must be accounted for. I will consider the textual evidence, discuss a number of specific difficulties which the distinction gives rise to, and the implications of their solution for Kant's moral theory.

Chapter 2
ARISTOTLE AND MILL

The claim that actions have moral worth insofar as they are performed from a moral motive is commonly regarded as *the* distinguishing feature of Kant's moral philosophy, as well as his most original contribution to the subject. Of course Kant, like every moral philosopher, is concerned with establishing and justifying an ultimate principle for determining what is right and wrong. His theory of moral motivation, on the other hand, is commonly regarded as a Kantian answer to a peculiarly Kantian problem, since most rival moral philosophers simply reject the Kantian concern with agent's inner motives and acting from duty. My aim in this chapter is to outline some of the reasons why this common view might be thought to be mistaken. A comparison of Kant's views on the issue with those of his opponents is illuminating for at least three reasons: Initially, there is a primarily *historical* point: the insistence on a motivational criterion in morality is not Kantian *in origin*, but is present (sometimes prominent) in the work of many of his predecessors. Secondly, we need to consider the fact that Aristotle's account of virtue and happiness and Mill's utilitarianism probably constitute the *main alternatives* which present the most serious objections to a Kantian account of morality. This in itself suggests that their stance on the motivational criterion should be of some importance in the present context. Thirdly, I want to suggest that the Kantian motivational claim is compatible with other, non-Kantian or even distinctly anti-Kantian normative criteria. So, finally, I need to argue that even a theory which explicitly rejects a Kantian approach, such as Mill's utilitarianism, and does not itself contain a version of the Kantian motivational criterion (as I argue Aristotle's does), can be shown to be compatible with such a criterion.[10]

---

[10] Another interesting point of contact between these rival theories which, however, I will not consider in the present context, is the following: Kant himself, as well as his apologists and critics alike, observe (in varying degrees) that, far from being entirely opposed to them, Kantian moral theory already contains or accommodates elements of eudaimonism, virtue ethics, teleology, and consequentialism. On the influence of eudaimonism, virtue ethics, and teleology in Kant's moral philosophy see, for example, Barbara

Herman, 'Making Room for Character', Stephen Engstrom, 'Happiness and the Highest Good in Aristotle and Kant', Jennifer Whiting, 'Self-Love and Authoritative Virtue: Prolegomenon to a Kantian Reading of *Eudemian Ethics* viii 3', Christine Korsgaard, 'From Duty and for the Sake of the Noble: Kant and Aristotle on Morally Good Action', and Julia Annas, 'Aristotle and Kant on Morality and Practical Reasoning', all in: Stephen Engstrom & Jennifer Whiting (eds.) *Aristotle, Kant, and the Stoics – Rethinking Happiness and Duty*, Cambridge: Cambridge University Press, 1996. On the vexed question of categorising ethical theories, see Korsgaard, pp 232f (fn20) in the same volume. For a consequentialist interpretation of Kant see e.g. David Cummiskey, *Kantian Consequentialism* (Oxford: Oxford University Press, 1995). Of course there are also numerous attempts to read Kant as a rule-utilitarian. For one such example, see R. F. Harrod, "Utilitarianism Revised", *Mind* 45 (1936), 137-156.

Some commentators go as far as claiming that, especially if certain elements are stressed enough at the expense of others, his ethical thought collapses into a consequence-, virtue-, or happiness-centred ethics. For an example of the former, see John Stuart Mill, *Utilitarianism* (London: Longmans, Green & Co., 1897, 13th ed.) p 5f., where Mill criticises Kant for being a de facto utilitarian who, while claiming to reject consequentialism, cannot but revert to consequentialist considerations in the employment of his supreme principle of morality, the Categorical Imperative. Cf. Mill, op cit., p5f.: "[There is no] school of thought which refuses to admit that the influence of actions on happiness is a most material and even predominant consideration in many of the details of morals, however unwilling to acknowledge it as the fundamental principle of morality, and the source of moral obligation. I might go much further, and say that to all those *a priori* moralists who deem it necessary to argue at all, utilitarian arguments are indispensable. [...] One of the most illustrious of them [...], Kant, this remarkable man, [...] when he begins to deduce from this precept [i.e. the categorical imperative] any of the actual duties of morality, he fails, almost grotesquely, to show that there would be any contradiction, any logical (not to say physical) impossibility, in the adoption by all rational beings of the most outrageously immoral rules of conduct. All he shows is that the *consequences* of their universal adoption would be such as no one would choose to incur."

As regards the latter, some commentators argue that Kant's duty ethics turns into a *virtue ethics* in his mature philosophy: The earlier rigorism (exclusion of all non-rational motives) predominant in the *Groundwork* makes way for an Aristotelian concern with virtue and the cultivation of character in the *Metaphysics of Morals* and the *Religion*. Others identify a type of *eudaimonism* in the culmination of Kant's ethical thought: his conception of the highest good.

2. The moral motive in Aristotle's *Nicomachean Ethics*

Although it is usually, and to some extent justifiably, considered a rival theory, Aristotle's *Nicomachean Ethics*[11] contains an account of morality that anticipates basic elements of Kant's views on moral motivation. Aristotelian virtues are states of character or of mind and have an affective and an intellectual aspect, i.e. they involve both our feelings, especially of pleasure and pain, and our intellect (in reasoning and practical intelligence). There can be three different conditions of soul concerned with action: feelings (*pathē*) such as desire, fear etc., capacities (*dunameis*) in respect of which we are capable of having these feelings, and states or dispositions (*hexeis*)[12] in respect of which we are inclined towards the feelings. Aristotle argues by a process of elimination that the genus of virtues is state, rather than feeling or capacity. They cannot be the latter since we have feelings and capacities 'by nature', which as such are not subject to moral praise or blame (which virtues, or a lack of them, are), and as they do not involve deliberate choice or decision (*prohairesis*)[13], whereas virtues do.[14] But it remains to be seen what exactly it is

---

[11] *Nicomachean Ethics*, transl., with introd., notes, and glossary, by Terence Irwin (Indianapolis/ Cambridge: Hackett, 2nd ed, 1999). Hereafter referred to as *NE*. Commentary parts referred to as Irwin, op. cit. Additions in square brackets [ ] are Irwin's unless indicated otherwise.

[12] On translating *hexis* (a stable state, disposition, or habit) and *diathesis* (a less stable disposition) see Julia Annas, *The Morality of Happiness* (Oxford: Oxford University Press, 1993), p. 50, and W.F.R. Hardie, *Aristotle's Ethical Theory* (Oxford: Clarendon Press, 1980), pp. 103f. I will use 'state' and 'disposition' interchangeably.

[13] Decision is a kind of 'rational desire' involved in deliberate choice. Susan Meyer observes that in "contexts where he is not discussing his specific conception of virtue and vice of character, Aristotle generally follows Plato in recognizing a single kind of rational desire. His name for this desire is 'wish' (*boulēsis*). When articulating his own account of virtue and vice of character, however, he distinguishes wish from decision (*prohairesis*) and insists that decision is the desire most important to virtue and vice of character." [See NE 1105a31-33, which I discuss below.] It is important primarily because it contains a motivational element: "A wish may simply be a favourable evaluative attitude toward an opportunity that has no actual effect on one's inclinations to act. A decision, by contrast, is a determinate intention to act that actually moves the agent to act. It is not always successful, for it can be impeded by an opposing desire, as in the case of continence or incontinence. However, it always yields some impulse (*hormē*) toward action. This is the impulse that non-rational desire opposes in the case of incontinence, and that

about a disposition that differentiates it enough from feelings and capacities to make it, unlike these, a possible bearer of virtue.

How does a person become virtuous? Aristotle believes that learning how to become virtuous is in many respects not unlike learning to master a craft, or some similar skill (technē).[15] The craft analogy is prominent in many of his works, but particularly so in his ethics. We will need to consider here in some detail the extent to which Aristotle utilises this analogy as well as the extent to which he rejects it. It will become clear that at a certain point the analogy breaks down entirely and the *differences* between a skill with its concern with *making* things (*poiēsis*) on the one hand, and virtue with its focus on *doing* or action (*praxis*) on the other, are more illuminating than the similarities.

Aristotle divides the virtues into intellectual (*dianoētikai*) and moral (*ēthikai*), both of which are acquired rather than natural. The former arise and grow from teaching, whereas the latter result from habit (*ethos*).[16] This feature of virtues of character invites the aforementioned skill analogy:

> Virtues, by contrast [with the senses, which arise in us by nature (my addition)], we acquire, just as we acquire crafts, by having first activated them. For we learn a craft by producing the same product that we must produce when we have learned it; we become builders, for instance, by building, and we become harpists by playing the harp. Similarly, then, we become just by doing just actions, temperate by doing temperate actions, brave by doing brave actions. (*NE* II 1 §4, 1103a27-1103b2)
>
> To sum it up in a single account: a state [of character] results from [the repetition of] similar activities. (*NE* II 1 §7, 1103b21f)

---

overcomes non-rational desire in cases of continence [...] It is an 'origin of action' (*archē praxeōs*) (NE 1139a31)." See her *Aristotle on Moral Responsibility – Character and Cause* (Oxford: Blackwell, 1993), pp. 24ff.

[14] See *NE* II 5. Although I cannot discuss this in this context, it is worth mentioning that this is another important conviction that Aristotle shares with Kant: the belief that we are only responsible and can be held accountable for what we have freely chosen from within the sphere of our influence, thereby excluding feelings, inclination, and consequences beyond our control from moral evaluation and rejecting the idea that there can be such a thing as 'moral luck'.

[15] "He is not, of course, the first to do so; the analogy with *technē* is extensively developed by the Platonic Socrates, and Aristotle is clearly at times reacting to this." Julia Annas, op. cit. p. 67 (fn58).

[16] NE II 1, 1103a15.

The analogy stresses an important feature of virtues of character: they are not only acquired, but are learnt and grow through a process of repetition and habituation which involves the same virtue both as an activity (*praxis*) and as the intended result (*telos*):

> But these actions are not only the sources and causes both of the emergence and growth of virtues and of their ruin; the activities of the virtues [once we have acquired them] also consist in these same actions. [...A]bstaining from pleasures makes us become temperate, and once we have become temperate we are most capable of abstaining from pleasures. It is similar with bravery; habituation in disdain for frightening situations and in standing firm against them makes us become brave, and once we have become brave we shall be most capable of standing firm. (*NE* II 2 §8-9, 1104a28-39).

Aristotle argues that a person's repeated performance of actions which he has recognised as virtuous will allow him to habituate himself in such a way that he will develop a disposition to act virtuously. This disposition will in turn become part of the person's stable character and manifest itself in the continued performance of the virtuous actions in which it originated. Thus what Aristotle presents here is more than an analysis of virtue, but an account of moral education which stresses that virtue is an activity that must be practised to acquire the right state of character. Habituation is required because virtue involves non-cognitive aspects of the soul (feelings, desires) as well as the recognition of the good, and these need to be trained and practised in order to become virtuous:[17]

> It is right, then, to say that a person comes to be just from doing just actions and temperate from doing temperate actions; for no one has the least prospect of becoming good from failing to do them. (*NE* II 4 §5, 1105b10-12)[18]

---

[17] Cf. Susan Meyer, op. cit., ch.5.4, and J.A. Stewart, *Notes on the Nicomachean Ethics of Aristotle* (Oxford: Clarendon Press, 1892), vol. I, p. 185.

[18] Furthermore, this point is an implicit criticism of Socrates' intellectualist account of virtue which claims that knowing the good is sufficient for acting virtuously - hence the Socratic denial of the possibility of weakness of will (*akrasia*). Aristotle attacks the implications of this account at the end of his own explanation of the process of character formation by habituation: "The many, however, do not do these actions. They take refuge in arguments, thinking that they are doing philosophy, and that this is the way to become excellent people. They are like a sick person who listens attentively to the doctor, but acts on none

But Aristotle anticipates a possible objection to his account: The process of learning and habituation just described appears to be circular: virtue depends on habitually acting virtuously, and being able to act virtuously depends on having virtuous dispositions. This invites the accusation that he is begging the question. It is at this point that Aristotle makes some important qualifications to the skill analogy (going as far as claiming that it breaks down entirely), and some strikingly 'Kantian' observations about the nature of virtue. He presents us with a puzzle (*aporia*) that appears to attack his account of habituation by way of a *reductio ad absurdum*:

> Someone might be puzzled, however, about what we mean by saying that we become just by doing just actions and become temperate by doing temperate actions. For [one might suppose that] if we do grammatical or musical actions, we are grammarians or musicians, and, similarly, if we do just or temperate actions, we are thereby just or temperate. (*NE* II 4 §1, 1105a17-21)

This is the practical equivalent of the learning paradox that Plato raises in the Meno.[19] "The puzzle arises because Aristotle has emphasized the similarity between the actions that we learn to do in habituation and the actions that we do when we are virtuous."[20] This similarity suggests that learning a skill is a good model for learning a virtue. Both skill and virtue are, certainly to some extent, the product of repetitive practice which might involve, at least initially, a fair amount of picking up what to do in particular cases

---

of his instructions. Such a course of treatment will not improve the state of the sick person's body; nor will the many improve the state of their souls by this attitude to philosophy." (NE II 4 §6, 1105b13-19)

[19] "MENO: But how will you look for something when you don't in the least know what it is? How on earth are you going to set up something you don't know as the object of your search? To put it another way, even if you come right up against it, how will you know that what you have found is the thing you didn't know? SOCRATES: I know what you mean. Do you realize that what you are bringing up is the trick argument that a man cannot try to discover either what he knows or what he does not know? He would not seek what he knows, for since he knows it there is no need of the inquiry, nor what he does not know, for in that case he does not even know what he is to look for. MENO: Well, do you think it a good argument? SOCRATES: No." Plato, Meno, 80d-81a, transl. W.K.C. Guthrie, in Plato, *The Collected Dialogues*, ed. Edith Hamilton & Huntington Cairns (Princeton, Princeton University Press, 1961).

[20] Irwin, op.cit. p. 195.

and copying from one's teachers. But now the proponent of the puzzle suggests that if it is possible to attempt to become skilful or virtuous by simply engaging in skilled or virtuous activity, then surely one has already mastered it and there is nothing to learn. In other words, attempting to *become* virtuous seems to presuppose already *being* virtuous, which would make it impossible to even begin to learn. In reply, Aristotle will have to show that in his account of the learning process he does not have to presuppose from the start that which is supposed to be the end result. He will have to explain "why his account of habituation does not self-defeatingly assume the existence of the state that is supposed to result from the habituation".[21]

Aristotle now argues that "the objector's argument (1) rests on an alleged feature of the crafts, and hence (2) assumes that virtues are analogous to crafts in the relevant ways".[22] In his first reply he still upholds the analogy but challenges the effectiveness of the *aporia* by arguing that the crafts do not actually support the objection: The analogy, so far as it holds, evidently does not support the suggestion that his account of habituation is self-defeating, but works against it:[23]

> "But surely actions are not enough, even in the case of crafts; for it is possible to produce a grammatical result by chance, or by following someone else's instructions. To be grammarians, then, we must both produce a grammatical result and produce it grammatically – that is to say, produce it in accord with the grammatical knowledge in us." (*NE* II 4 §2, 1105a22-26)

Even in the crafts, but certainly in virtue, Aristotle argues, it is not enough to look only at the outwardly manifested doing of the action. Even in the crafts we make a distinction between getting something right by chance, or being told how to do it, and doing something deliberately and knowledgeably. Writing words correctly does not amount to being literate. So to the extent that the analogy is a good one it supports the view that in

---

[21] Ibid.

[22] Ibid.

[23] Cf. Stewart, who observes that "the analogy between *areté* and *techné*, appealed to by the promoters of the *aporia* (which, it may be observed, is an *ignava ratio* in morals – or excuse for inactivity – [...] similar to that in science refuted in the Meno) [...] does not help them." Op. cit. p. 183. Cf. Irwin, op.cit. p. 195.

both the crafts and virtue there is a difference between doing something correctly, 'to the letter', and knowing how to do it. Therefore the *aporia* rests on a false equivocation.[24]

This is as far as the analogy can be stretched. Having taken the wind out of its sails by showing that the *aporia* is not an objection to his account of habituation, Aristotle now explicitly rejects the objector's assumption that virtues are analogous to crafts in the relevant ways. In turning against the analogy, he insists that *virtue itself* is not a skill:

> Moreover, in any case, what is true of crafts is not true of virtues. For the products of a craft determine by their own qualities whether they have been produced well; and so it suffices that they have the right qualities when they have been produced. (*NE* II 4 §3, 1105a26-29)

This is the claim that the excellence of products lies in themselves rather than in the way they were made.[25] "The goodness of the work of art is a quality which we see at once in the work itself – Hermes is dug up at Olympia, and we find him beautiful as soon as we see him."[26] We will see how virtue differs from this in a moment, but one immediate problem that arises is Aristotle's apparent retraction of a claim he made earlier: part of the point of the rejection of the *aporia* seemed to be that the crafts, like the virtues, imply an underlying intellectual grasp of the processes involved, which distinguishes the 'truly skilled' from the 'accidentally successful' craftsman. But "Aristotle is not taking back the point he has made in the first reply, that someone might produce a good product accidentally. He means that the goodness and badness of production is determined by its usefulness for producing the product; a better method of production is better because it

---

[24] Cf. Stewart: "This *aporia* we may resolve by pointing out (which Aristotle does not do explicitly) that [just acts] before the formation of the Habit are, so far as the agent is concerned, [...] only apparently *his* just acts; really, *qua* just, they are expressions of the wise intention of the law-giver, or ruler, who enjoins and enforces them. The moral agent himself has to begin by acting under the compulsion of the law, until by repetition a habit is formed, and he performs the acts in question *proprio motu*, the habit bringing with it an insight, more or less clear, into the significance of the acts, and a belief that they are good. Then they are really the agent's own just acts – i.e. they are 'just' in the strict, and not in an equivocal sense." Op.cit., p. 183.

[25] Cf. Annas, op. cit., p. 68.

[26] Stewart, op. cit., p. 183.

is better at producing the right sort of product."[27] Virtue should be put in sharp contrast with the crafts because

> for actions in accord with the virtues to be done temperately or justly it does not suffice that they themselves have the right qualities. Rather, the agent must also be in the right state when he does them. First, he must know [that he is doing virtuous actions]; second, he must decide on them, and decide on them for themselves; and, third, he must also do them from a firm and unchanging state. (*NE* II 4 §3, 1105a29-35)

This passage, I believe, clearly demonstrates that Aristotle anticipates Kant's insistence on a motivational criterion in morality. He argues that "it is unsafe to estimate an action [...] from the outside, judging of its moral value from its splendour or immediate utility; before we come to a decision, we ought to know the state of the agent – whether he is aware that he performs the act, whether he chooses it because it is right, and that, of fixed habit, having often performed it before, and therefore being likely to perform it often again."[28] Irwin also supports such a reading of this crucial passage:

> The value of virtuous action, as opposed to the craftsman's production (the process), is not simply determined by its efficiency in producing a product; it also has its characteristic motive. The value of virtue is intrinsic; virtuous action is not valuable simply as a means to some further result (e.g., acting kindly is not simply a means to making someone feel better). The intrinsic value of virtue reflects the virtuous person's *motive*, shown by the second condition in a32. The *demand for a specific motive* differentiates virtue from craft, and hence differentiates the training required for each of them; this is Aristotle's answer to the puzzle raised in the chapter.[29]

---

[27] Irwin, op. cit., p. 195. Cf. Annas' explanation of the somewhat counterintuitive way in which Aristotle makes his point here. She observes that, contrary to appearances, Aristotle in fact acknowledges many similarities between skill and virtue which the above passage, on the face of it, seems to deny. He stresses the intellectual aspects of skill, but "chooses to express this position by contrasting skill and virtue. But this is not the only possible response; one could say that virtue, since it shares these features of intellectual grasp and need for habit and practice, is a special kind of skill. Presumably Aristotle insists that virtue is not a skill because skills other than virtue can be practiced in comparative independence of the agent's moral character." Op. cit., p. 68.

[28] Stewart, op. cit., p. 184.

[29] Irwin, op. cit., p196, my emph. Cf. also W.F.R. Hardie, op. cit., ch.VI., and J.A. Stewart's summary of the passage which highlights the parallels with Kant's motivational criterion nicely:

With the first of the three conditions Aristotle rules out involuntary and non-deliberate actions which as mere accidents would be without any moral significance. The agent must perform his duty *deliberately* and *knowingly* for it to qualify as an act at all. As I mentioned before, Aristotle shares with Kant the conviction that we are only morally responsible for what we have freely chosen. The third condition, i.e. acting from a state that is a permanent fixture in a person's character as a result of education and habituation, is arguably the feature of this definition of virtue that most marks it as Aristotelian. But from a Kantian perspective, the second condition is immediately recognisable as the feature of moral goodness that is not only crucial for Kant's conception, but often cited as its distinguishing quality: that is the claim that the agent must be motivated to choose and perform virtuous actions for their intrinsic rightness.[30] The most important point here is that Aristotle differentiates between actions that are only incidentally virtuous (cf. Kant's *action merely in accordance with duty*) and those that are done knowingly, of choice, and from a virtuous disposition (cf. Kant's *action from duty and in accordance with duty*).[31]

"[I]n the realm of conduct we do not say 'Well done!' as soon as we see an action, with certain good qualities in it, performed. Before we pronounce, we look at the agent also, and ask – 'How did he perform it?' [...] Did he choose deliberately to perform it, and that, because he thought it right, not because he thought it would lead to some ulterior pleasure or profit? [...] Unless these conditions in the agent be fulfilled, we do not speak of the moral value of actions [...]."

"[T]he goodness of a moral action is not a quality in the action itself, which we can appreciate apart from the goodness of the agent."

"[B]efore we come to a decision [when estimating the moral value of an action], we ought to know the state of the agent – whether he is aware that he performs the act, whether he chooses it because it is right [...]". Op. cit., pp. 182, 183, 184.

[30] Cf. Korsgaard, op. cit., on this point.

[31] Discussing Kant's *Groundwork*, chapter I, Ernst Tugendhat mentions in passing that Kant makes "the distinction, *common in all ethics since Aristotle*, between an action merely 'in accordance with duty' and an action 'from duty'" (my transl.) ["die *in aller Ethik seit Aristoteles übliche Unterscheidung* zwischen einer bloß 'pflichtmäßigen' Handlung und einer Handlung 'aus Pflicht'"], but he does not elaborate on the Aristotelian origins of the distinction, nor indeed on its presence in pre-Kantian ethics. See his *Vorlesungen über Ethik*, (Frankfurt am Main: Suhrkamp, 1993), p. 110.

Aristotle continues chapter 4 by stating that,

> As conditions for having a craft, these three do not count, except for the bare knowing. As a condition for having a virtue, however, the knowing counts for nothing, or [rather] for only a little, whereas the other two conditions are very important, indeed all-important. And we achieve these other two conditions by the frequent doing of just and temperate actions. (*NE* II 4 §3, 1105b1-6)

Apart from including another gibe at the Platonic intellectualist account of morality, this point further supports his total rejection of the skill analogy. But as Irwin observes, it is also noteworthy that Aristotle here "passes from (a) the contrast between products produced well and actions done well, to (b) the contrast between good craftsmen and virtuous agents. The two contrasts are connected, because an action is done well only if a good agent does it from the right state (whereas a product can be produced well even if a good craftsman does not produce it)."[32]

It would of course be worthwhile to consider whether the Aristotelian and the Kantian categories match exactly, whether they have the same prominence in their overall outlook, and so on, but my point here is simply that the notion of taking some fact about the agent to be indicative of moral value, and most crucially, making that agent's *motive* (which, amongst other things, constitutes 'the state of the agent') the criterion of morality, is a notion that is common to and indeed decisive for both Aristotle's and Kant's account of morality. This basic point is supported by the passage at *NE* II 4 §3, 1105a29-35.

[32] Irwin, op. cit., p. 196.

3. The moral motive in Mill's *Utilitarianism*

We saw earlier there are two core features of Kant's moral theory. The Categorical Imperative is intended to enable the agent to recognise what his duty is in a particular context by universalizing the maxim of his action. It is Kant's objective, *normative* criterion for what is morally required and permissible. If a maxim can be universalized without contradiction then the action is objectively good (or: in accordance with the law), and, if performed, the action is (at least) done in accordance with duty: "The concept of duty requires of an action, objectively, accordance with the law."[33] It meets Kant's normative criterion of moral goodness (the legality requirement of action). Of course, as mentioned earlier, the peculiarity of Kant's particular moral theory is that he insists on a second, much stricter, subjective, or *motivational*, criterion of morality that attaches not to the action but to the agent: An action has moral worth only if it is in accordance with duty and also done from duty.[34] This criterion is subjective insofar as the attribute 'done from duty' cannot be read off actions or their effects; rather, something must be known about the state of the agent.

One interesting consequence of Kant's clear distinction between his normative and his motivational criterion of morality is that they can be critically examined and evaluated independently, and as a consequence, as I will now argue, it can be shown that the motivational claim can be quite compatible with otherwise rival ethical theories. I would like to illustrate this point using as an example a comparison with Mill's utilitarianism.[35] (I will be considerably briefer here as my sole aim is to show that the acceptance of a version of the principle of utility does not, by itself, preclude a concern with moral motivation.)

---

[33] *KpV* 5:81. Cf. *GMS* 4:397ff.

[34] Cf. *GMS* 398f., 406.

[35] Of course, there are important differences which are just as illuminating. My aim here is merely to show that despite these differences both distinguish between, on the one hand, actions which are right by their respective normative standards, and a richer notion of moral value which rests on the motivational state of the agent.

Mill's utilitarianism has at its root the Greatest Happiness Principle, which, notwithstanding significant differences, fulfils in Mill's theory the same kind of role that the Categorical Imperative plays in Kant's account, at least in the following sense: it is an ultimate normative principle which sets up a moral standard for the moral deliberator and allows him to decide in any given context what, objectively, is the right action:

> The creed which accepts as the foundation of morals, Utility, or the Greatest Happiness Principle, holds that actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness. By happiness is intended pleasure, by unhappiness, pain, and the privation of pleasure.[36]

According to Mill the rightness of an action is to be judged solely by considerations regarding its consequences: one ought to do whatever is estimated to produce the best consequences. The utilitarian criterion is, like Kant's supreme principle, a criterion of rational choice. As such it stands in direct contrast to Kant's Categorical Imperative as a rival principle for distinguishing right from wrong. In fact, it is explicitly rejected by Kant in the first section of the *Groundwork* by his claim that the consequences of actions cannot be the object of moral goodness. It is clear, then, that Kant's and Mill's normative criteria of goodness are directly opposed. But where does this leave the motivational criterion concerning the agent's motive?

Mill is generally dismissive of the very question. Insisting that there is a very important distinction to be made between "the rule of action" and "the motive of it" (which, of course is precisely Kant's), he goes on to claim that

---

[36] John Stuart Mill, op.cit., p9f. For a 20th century version of such an act-utilitarian principle see J.J.C. Smart, *An outline of a system of utilitarian ethics*, in: J.J.C. Smart & B. Williams: *Utilitarianism for and against* (Cambridge: Cambridge University Press, 1973), p30: "The only reason for performing an action *A* rather than an alternative action *B* is that doing *A* will make mankind [...] happier than will doing *B*, [or more precisely stated] that *A* will produce more probable benefit than will doing *B*." Or again: "According to the act-utilitarian, then, the rational way to decide what to do is to decide to perform that one of those alternative actions open to us [...] which is likely to maximize the probable happiness or well-being of humanity as a whole [...]. The utilitarian position is here put forward as a criterion of rational choice." (p. 42)

> It is the business of ethics to tell us what are our duties, or by what test we may know them; but no system of ethics requires that the sole motive of all we do shall be a feeling of duty; on the contrary, ninety-nine hundredths are done from other motives, and rightly so done, if the rule of duty does not condemn them."[37]

Rejecting the notion that a "system of ethics requires that the sole motive *of all we do* shall be a feeling of duty" Mill here insists that from a moral point of view it is sufficient that "ninety-nine hundredths" of actions are done in accordance with duty but not from duty. Of course, even those who insist on the moral motive allow that there are actions which simply do not fall into the sphere of the moral: most actions are not subject to moral evaluation, and even of those that are, not all are required to be done from a moral motive (and Kant certainly acknowledges this). But Mill seems to conclude from this that it could not even be the case that the sole motive of *some* things we do is required to be duty, which simply doesn't follow. But that seems to be precisely what Kant, at the very least, does claim: *Some* actions ought to be done from the motive of duty, and their moral worth, if any, lies in their being performed from the motive of duty.

Despite his dismissiveness concerning such motivational considerations in ethics,[38] and despite the fact that he certainly never develops a theory of moral motivation, Mill does make certain concessions to those who insist on them, stating that

> utilitarian moralists have gone beyond almost all others in affirming that the motive has nothing to do with the morality of the action, *though much with the worth of the agent.* [...]

---

[37] Ibid., p. 26.

[38] Smart is equally dismissive of motivational considerations. On the one occasion where he does mention moral motives, he defines them entirely in terms of the likelihood and frequency of their moving the agent to produce actions that maximize happiness; i.e. entirely in terms of resulting consequences. "What we do need is a pair of terms of appraisal for *agents* and *motives* [as opposed to *actions*]. I suggest that we use the terms 'good' and 'bad' for these purposes. A good agent is one who acts more nearly in a generally optimific way than does the average one. A bad agent is one who acts in a less optimific way than the average. A good motive is one which generally results in beneficent actions, and a bad motive is one which generally ends in maleficent actions." Op. cit., p. 48. "'My favourite food!' the Gruffalo said. 'You'll taste good on a slice of bread!' – 'Good?' said the mouse. 'Don't call me good! I'm the scariest creature in this wood.'" Cf. Julia Donaldson & Axel Scheffler, *The Gruffalo* (London: Macmillan, 1999), no page no.

> [T]he motive, that is the feeling which makes him [the agent] will so to do, when it makes no difference in the act, makes none in the morality: *though it makes a great difference in our moral estimation of the agent* [...].[39]

Although rejecting motives as irrelevant for the evaluation of *actions*, Mill very much allows that they are crucial when it comes to estimating the moral worth of the *agent*. Despite the fact that Mill makes the point in terms of a contrast between actions and agents, the distinction he makes here between the morality of the action and the moral worth of the agent perfectly mirrors Kant's own distinction between actions which are merely in accordance with duty because they conform to the normative criterion and actions which are done from duty and thereby meet the motivational criterion. The point, of course, is that in order to determine moral worth, something must be known about the motivational state of the agent.

Despite the fact that Mill focuses almost exclusively on actions and their consequences and shows no concern with developing a theory of the moral worth of acts (or agents), as opposed to their conformity with the utilitarian 'rule of action', there does not appear to be any inherent systematic incompatibility of such an undertaking with a utilitarian account of action. While Kant insists that true moral value is only found the agent's disposition, Mill simply thinks that this is not as important or interesting a question as the one concerning 'the rule of action'. But he does not show that it is not a *valid* question, nor does he show that a concern with it would somehow conflict with an otherwise utilitarian outlook.[40]

The closest Mill comes to explicitly rejecting the concern with motives is in the following passage:

> He who saves a fellow creature from drowning does what is morally right, whether his motive be duty, or the hope of being paid for his trouble: he who betrays the friend that trusts him, is guilty of a crime, even if his object be to serve another friend to whom he is under

---

[39] Mill, op. cit., p. 26, p. 27fn, my emph.

[40] Perhaps the only possible explanation of Mill's lack of interest in the inner motives of the moral agent is a socio-historical observation along the following lines: "These people an empire to run – of course they had no time for *Innerlichkeit*!" (Vasilis Politis)

> greater obligations. But to speak only of actions done from the motive of duty, and in direct obedience to principle: it is a misapprehension of the utilitarian mode of thought [...].[41]

The context is important here: Mill is replying to the accusation that utilitarianism itself *entails* the requirement of some form of 'purity of motive'. This of course it does not and he makes it very clear that he does not think such a requirement is necessary. Yet, in the context of my present point, it is important that Mill only makes the weaker claim that utilitarianism *does not entail* a motivational requirement. That is, those who find such a criterion objectionable (and he clearly includes himself amongst them) do not have to accept it by accepting utilitarianism. He does *not* show, however, that such a requirement, if supported by other (non-utilitarian) considerations, could not possibly be combined with an otherwise utilitarian outlook, such that it becomes a requirement of at least certain actions that they be done from the motive of duty.

So while Kant positions the motive of duty at the core of his ethical theory, Mill claims that – at least as far as theory is concerned – it is not the business of ethics to contemplate agents' possible motives as long as their actions conform to utilitarian considerations. My point here is simply that this is all that Mill's response is: a declaration of a lack of interest. It is *not* a consequence of his normative utilitarian views.

[41] Ibid., p. 26f.

Chapter 3
KANT'S ACCOUNT OF THE GOOD WILL

1. The good will

Kant famously opens the *Groundwork* with the statement that the only thing which is good *without qualification* is a good will.

> There is nothing it is possible to think of anywhere in the world, nor indeed even outside it, which could be held to be good without limitation, except only a *good will*. Understanding, wit, power of judgement, and whatever else the *talents* of the mind might be called, or courage, resoluteness, persistence in intention (Beharrlichkeit im Vorsatze), as qualities of *temperament*, are without doubt for many an intention (in mancher Absicht) good and to be wished for; but they can also become extremely evil and harmful, when the will that is to make use of these gifts of nature, and the peculiar constitution of which is therefore called *character*, is not good. (GMS IV:393, my transl.)

Having introduced the notion of the good will as the only unqualified and unlimited good, Kant immediately turns his attention to a discussion of the classic virtues of the ancients. His ultimate aim is to show that they are all qualified goods, and thus that unlike any of these, the good will is alone in possessing unqualified goodness. Under the heading of *gifts of nature* he mentions two categories, *talents of the mind* and *qualities of temperament*. In the following paragraph he lists an additional set of "qualities" (ibid.) which would equally fall under these headings (presumably under *qualities of temperament* in the case of the first two, and under *talents of the mind* in the third case), but which are singled out because, even more so than the aforementioned gifts of nature, they might be thought to be linked to the inner worth of a person:

> Some qualities are even conducive to this good will itself and can make its work much easier, but still have despite this no inner unconditioned worth, yet always presuppose a good will, which limits the esteem that one otherwise rightly has for them, and does not permit them to be held absolutely good. Moderation in affects and passions, self control, and sober reflection not only

> are good for many aims, but seem even to constitute a part of the *inner* worth of a person; yet they lack much in order to be declared good without limitation (however unconditionally they were praised by the ancients). For without the principles of a good will they can become extremely evil, and the cold-bloodedness of a villain makes him not only far more dangerous but also immediately more abominable in our eyes than he would have been held without it. (GMS IV:393f.)

Under the heading of *gifts of fortune*, Kant goes on to consider another group of qualities or states of the person, which might be, and have been, identified as possessing a higher worth than any other, pre-eminently amongst them happiness.

> It is just the same with *gifts of fortune*. Power, wealth, honour, even health, and that entire well-being and contentment with one's condition, under the name of *happiness*, make for courage and thereby often also for arrogance (Mut und hiedurch öfters auch Übermut), where there is not a good will which corrects and makes universally purposive their influence on the soul (Gemüt) and thereby on the entire principle of action; not to mention that a rational impartial spectator can never take satisfaction even in the sight of the uninterrupted welfare of a being which is not adorned by a trace of a pure and good will, and thus the good will appears to constitute the indispensable condition even of the worthiness to be happy. (GMS IV:393, my transl.)

There are some important implications of Kant's comparison between these goods on the one hand, and the good will on the other, to which I will return shortly. Presently it is worth noting that the upshot of the centrality of the concept of a good will in Kant's ethical theory is that the agent's will becomes the bearer of moral goodness. An agent is morally good or bad primarily in terms of his volition. Moral goodness lies in the will's reasons for willing (choosing) certain ends and actions, insofar as these reasons motivate the agent to act. There are also a number of issues concerning the nature and the validity of Kant's arguments which I will not entertain in any detail in the present context. But we may note briefly that his claim is not that the good will is the *sole* good: many things are said to be "without doubt" good "for many an intention [...]and to be wished for" (ibid.), and some qualities of temperament "are good for many aims" (ibid.). But if the good will is not the *only* good, then it follows that by itself it can neither be the highest good or *summum bonum*, for the latter is "a whole in which the greatest happiness is represented as connected in the most exact proportion with the greatest degree of moral

perfection (possible in creatures)" (KpV V:129f.), for which reason moral philosophy "is not properly the doctrine of how we are to *make* ourselves happy but of how we are to become *worthy* of happiness" (KpV V:130). Nor, again, does Kant think that the good will is the only good which is good for its own sake or good-in-itself, although one can easily read such a claim into the passage that follows:

> The good will is good not through what it effects or accomplishes, not for its efficacy for attaining any intended end, but only through its willing, i.e., good in itself, and considered for itself, without comparison, it is to be esteemed far higher than anything that could be brought about by it in favour of any inclination, or indeed, if you prefer, of the sum of all inclinations. (GMS IV:394)

The good will is, of course, as Kant says, good in itself, because its goodness lies in itself, that is, in its willing, rather than in "any intended end" (ibid.). But Kant does not claim that a good will is the only thing that we value on account of its *intrinsic* qualities. Happiness, as both Aristotle and Mill argue so convincingly, is also something we value for its own sake, because we never value it for the sake of something else, but we do value many other things for its sake.

At this point a comparison with both Aristotle's and Mill's theories of intrinsic goodness is instructive, for in broad terms Kant shares with both of them a certain argumentative strategy. Each of these three philosophers attempts, with the use of various criteria, to single out from amongst the virtues and goods that we all value one good which stands above the rest in some decisive respect. But contrary to Aristotle and Mill, who by virtue of its intrinsic goodness elevate happiness – of course, understood in very different terms – to the highest good and ultimate end, Kant does not think that the quality of intrinsic goodness singles out any one good. Rather, the only consideration which can do this is his argument for the unqualified goodness of the good will, which is of special worth and the *supreme* good (*bonum supremum*).

2. The control condition and the effort condition

In order to see why he feels justified in making this claim we must examine more closely the particular features that according to Kant all the qualified goods, but not the good will, share. If successful, this argument will single out the good will, at least in some respect. I will set aside the question whether that in itself is sufficient to prove its unqualified goodness. My interest here is rather in how this particular aspect of the good will highlights an important feature of Kant's twofold distinction of the concept of duty, which, after all, he considers to be an explication of the concept of a good will:

> But in order to develop the concept of a good will, which is to be esteemed in itself and without further aim, such as it already dwells in the natural common understanding and requires not so much to be taught as rather only to be clarified (aufgeklärt), [to develop] this concept, which in the estimation of the whole worth of our actions always holds the highest place and constitutes the condition of everything else: we will put before ourselves the concept of *duty*, which contains that of a good will, although under certain subjective limitations and hindrances, which, however, far from hiding it or making it unrecognisable, rather elevate it by contrast and let it shine forth all the more brightly. (GMS IV:397, my transl.)

It is worth noting in passing that Kant again confirms that he does not think that the notion of the unqualified goodness of a good will, which in an obvious sense is a novel philosophical concept, is a particularly original idea. Rather, he takes himself to be engaged in an analysis of an everyday concept as it "already dwells in the natural common understanding" (ibid).

But more importantly, when Kant insists in the third paragraph of the first section that even the goodness of a good will which is entirely unsuccessful in carrying out its aim would not be diminished, he insists on the difference between a mere wish and an act of willing.

> Even if through the peculiar disfavour of fate, or through the meagre endowment of a stepmotherly nature, this will were entirely lacking in the resources to carry out its aim, if with its greatest effort nothing of it were accomplished, and only the good will were left over (to be sure, not a mere wish, but *as the summoning up of all the means insofar as they are in our control*): then it would shine like a jewel for itself, as something that has its full worth in itself. Utility or fruitlessness can

> neither add to nor substract anything from this worth. It would be only the setting, as it were, to make it easier to handle in common traffic, or to draw the attention of those who are still not sufficiently connoisseurs, but not to recommend it to connoisseurs and determine its worth. (GMS IV:394)

The thought expressed in the emphasised clause has become known as Kant's *control condition* and has played a prominent role in the debate of the possibility of moral luck.[42] His claim, then, is that "the summoning up of all the means insofar as they are in our control" (ibid.) is all that can be morally required of us, and when an agent displays a good will in acting from duty, the moral worth of the action is not diminished by the lack of success.[43] My current concern with the control condition is its relevance for Kant's twofold distinction and how it might highlight his claims about the good will. But, as we shall see in the final chapter of part II., it will also play an important role when it comes to the question whether actions from mistaken moral judgement can have moral worth. Let us state the control condition in terms of the moral worth of actions in the following way:

The Control Condition of Moral Worth: *An action's moral worth cannot be diminished by what is beyond the agent's control.* Or, conversely: *An action's moral worth is determined only in terms of what is in principle within the agent's control.*

---

[42] Thomas Nagel labels his own, slightly narrower, version of this principle the 'condition of control'. See his 'Moral Luck', in: *Mortal Questions*, (Cambridge, 1979), p. 26. Bernard Williams suggests that such a condition is attractive because it presents the successful moral life "as a career open not merely to the talents, but to a talent which all rational beings necessarily possess in the same degree". Cf. his *Moral Luck*, (Cambridge, 1981), p. 21. Köhl refers to it as the 'Prinzip der Einflußnahme': *We are morally responsible only for whatever we have sufficient influence on.* See his *Kants Gesinnungsethik* (Berlin: de Gruyter, 1990), pp. 32-33, 41-44. Thus Kant wishes to exclude from the sphere of morality those elements of the agent, the action, and the situation which are beyond his control, especially those that are in any way 'accidental' or a result of 'luck'. In their discussion of *constitutive luck*, *resultant luck*, *circumstantial luck*, and *situational luck*, Nagel and Williams question the very possibility of this by pointing out that ordinary moral judgements do as a matter of fact often consider factors which lie beyond an agent's control. Of course, the Kantian will respond that such judgements must simply be mistaken, precisely because the 'condition of control' has been violated.

[43] Cf. also T&P VIII:279.

Furthermore, in the *Critique of Practical Reason*, in a passage where he is concerned with stressing that acting from duty involves constraint and often self-denial, Kant lays down a counterpart to the control condition, a principle which we might label the *effort condition* of moral worth:

> [C]onsidered in relation to human beings and to the individual human being; [the moral law] appears in a form that, though indeed highly deserving of respect, is not so pleasing as if it belonged to the element to which he is naturally accustomed but instead as it *constrains* him to leave this element, often not without self-denial, and to go to a higher element in which he can maintain himself *only with effort* and with unceasing apprehension of relapsing. In a word, the moral law demands obedience *from duty* and not from a predilection that cannot and ought not to be presupposed at all. (KprV V:158)

Allen Wood makes the same point in slightly different terms when he considers what Kant refers to as "certain subjective limitations and hindrances" (GMS IV:397) in the passage linking the concept of a good will and the twofold distinction of the concept of duty cited above. He describes how we can

> see here that acting from duty is a *special case* of the good will. It contains the concept of a good will, but it also contains some other features – in particular, 'certain subjective limitations and hindrances.' These consist generally in the fact that an agent subject to duties has needs and inclinations that might tempt it not to fulfill them, but more specifically in the fact that the good will that acts from duty must constrain itself to fulfill its duties, because in that case its inclinations do not suffice to secure what the good will wills it to perform. This *condition of moral adversity* is what Kant means by the 'subjective limitations and hindrances' that, far from concealing the good will, make it 'shine more brightly' – that is, display most conspicuously to us the inner, true, or authentic moral worth that elicits our moral esteem.[44]

While the control condition is a negative condition in that it denies that the moral worth of an action could be diminished by a lack of success insofar as this 'failure' can be said to be beyond the control of the agent, the effort condition (or the condition of moral adversity, in Wood's terms) makes the positive claim that an action can only be done *from*

---

[44] Cf. his *Kantian Ethics* (Cambridge: Cambridge University Press, 2008), p. 31 (second emph. added).

duty and thereby have moral worth if doing one's duty involves a *constraint* and an *effort* on the part of the agent.

The Effort Condition of Moral Worth: *Moral worth requires constraint and effort on the agent's part.* Or, conversely: *Moral worth is diminished by lack of constraint and effort on the agent's part.*

Thus the importance of the control condition and the effort condition lies in the fact that they spell out – each picking out a slightly different nuance of what is at bottom the same principle – how the moral worth of an action *depends* on, and depends *only* on, an effort of will on the part of the agent. That is to say, the goodness of the agent's willing is the necessary and sufficient condition of moral worth. Secondly, this clarifies the connection between Kant's basic concept of the good will and his twofold distinction of the concept of duty. Finally, in order to put these conclusions into the context of Kant's main argument for the unlimited goodness of the good will and the crucial difference he sees between the good will on the one hand, and the limited goods and qualities on the other, the following passage from the *Metaphysics of Morals* might be instructive. In his discussion of the duty of one's own perfection as one of the ends that are also duties, he makes the point that:

> When it is said that it is in itself a duty for a human being to make his end the perfection belonging to a human being as such (properly speaking, to humanity), this perfection must be put in *what can result from his deeds*, not in mere *gifts* for which he must be indebted to *nature*; for otherwise it would not be a duty. (MS VI:386f., emph. added)

Again, the principles underlying Kant's claim here are the control and effort conditions. The end of one's own perfection can be a duty only insofar as it requires of the agent moral *deeds*. Actions, which are chosen, or 'willed' in Kant's terms, and the performance or omission of which lies within the control of the agent, are here again contrasted with mere *gifts of nature*. This confirms the underlying theme of Kant's discussion of the good will. The single difference he aimed to establish and by virtue of which he separates the good will on the one hand, and the virtues of the ancients on the other, is that the latter, qua *gifts* of nature and *gifts* of fortune, are all in some fundamental respect, if not in their entirety, subject to luck and therefore beyond the control of the agent.

KANT ON THE POSSIBILITY OF ACTION FROM DUTY
BUT NOT IN ACCORDANCE WITH DUTY

PART II.

MISTAKEN MORAL JUDGEMENT AND ACTION FROM DUTY BUT NOT IN ACCORDANCE WITH DUTY

Chapter 4
OUTLINE OF THE PROBLEM

1. The origin of the problem in Kant's twofold division of the concept of duty

Kant's twofold distinction in the *Groundwork* between action in accordance with duty and action from duty has attracted an inordinate amount of commentary and criticism. However, one aspect of his account has been largely ignored by commentators, and, as I will attempt to show, has been misinterpreted by most of those who do address it: Kant's apparent oversight of the possibility of action from duty but not in accordance with duty. Since commentators commonly take the kind of action referred to by this description to be action from mistaken moral judgement, it is alleged that Kant fails to account for the possibility of such action. The question then is how this apparent failure is to be explained. My aim in the present chapter is to outline the common connection between the two issues and to argue for the need to separate them.

We saw in chapter 1 that Kant uses two criteria to distinguish actions, and to divide them into several types of action. As will become apparent in the course of this chapter, one of the more problematic aspects of this distinction is that it is not immediately obvious how many classes of action result from the twofold division.

Kant distinguishes actions, first, according to the criterion of rightness. All actions are either *right* (obligatory or permissible) when they conform to the requirements of morality *('in accordance with duty')*, or *wrong* (impermissible) when they do not *('contrary to duty')*. Since to act *from* duty is to act from the notion that the action is morally *required*, merely permissible actions can be ignored in this context.[45] Secondly, an action either has moral worth, namely when it is done *because* it is recognised as required (*'from duty'*), or it lacks such worth when it is done from some other motive, typically a desire or inclination (*'not from duty'*).

2. The claim that Kant's account of the motive of duty is incomplete: The Logical Compatibles View

There are two ways to approach the subject of action from duty but not in accordance with duty. The first arises from the notion, common amongst critics, that Kant's

[45] Cf. fn.3, ch. 1.1. above.

twofold distinction is a distinction between logical compatibles, the logical form of which is a double disjunctive such that either of the two attributes (here: accordance with duty and being done from duty) could be affirmed or denied of its subject (an action), without this predetermining the affirmation or denial of the other. Thus Kant's twofold distinction, "[l]ike any distinction between logical compatibles, [...] yields four logically possible classes of action"[46]. But since Kant fails to mention, let alone discusses one of the four, namely 'actions from duty but not in accordance with duty', commentators allege that Kant's account of duty, at least as it is presented in the *Groundwork*, is incomplete.[47] One of my main aims in the following chapters is to argue that this claim rests on an unwarranted assumption which arises from a misinterpretation of the textual evidence. For convenience's sake I will label the view that underlies this first assumption the *Logical Compatibles View* of Kant's twofold distinction of duty, according to which, when combined as logical compatibles, the criterion of rightness (accordance with duty) and the criterion of moral worth (being done from duty) give rise to four types of action, including the class of actions we are concerned with in the present context, namely 'actions from duty but not in accordance with duty'.

3. The claim that action from duty but not in accord with duty *is* action from mistaken moral judgement: The Identity Thesis

The second approach to our problem takes its starting-point from the phenomenon of actions which are motivated by an agent's sincerely held but *mistaken* belief that they are morally required.[48] Beginning with the idea that such actions are common and interesting and that therefore an appropriate moral theory must be able to account for them, the question then is how Kant (or Kantians) might do so. Here, the assumption

---

[46] John Hardwig, 'Action from Duty But Not in Accord with Duty', *Ethics* 93 (1983), reprinted in: *Immanuel Kant, Critical Assessments* (Vol. III), ed. Ruth Chadwick, London 1992, p. 68. The four classes of action are "(1) actions which are both in accord with duty and from duty; (2) actions which are neither from duty nor in accord with duty; (3) actions which are in accord with duty but not from duty; and (4) actions which are from duty but not in accord with duty." (Ibid.)

[47] Cf. Hardwig, op. cit., pp. 68-71; Kerstein, op. cit., pp. 119-132. Some go further and claim that Kant's omission of action from duty but not in accordance with duty is indicative of a serious inconsistency in his account of morality. See Hardwig, op. cit., pp. 73ff.

[48] I shall from now on refer to such action simply as 'action from mistaken moral judgement' or 'action from erroneous moral judgement'.

which I will question – again, it is one which is shared by most of the critics who address the issue – is that such actions are most appropriately captured in Kantian terms as 'actions from duty but not in accordance with duty':

> "Obviously, these would be actions which are motivated by the desire to do one's duty, but which do not correspond to the requirements of morality; they would be cases in which the agent is *mistaken* about what his or her duty is. The agent in question intends to do his duty; in fact, he does what he *takes to be* his duty and he does it because it *is seen to be* his duty. Thus his action must be described as action *from duty*. But he is mistaken about what duty (in general) is, about the morally relevant facts of the context of his action, or about how the general requirements of moral action apply to the context of his action. Because he is mistaken, he is in fact *not* acting *in accord with duty*, despite what may have been a conscientious attempt to formulate the correct moral judgement and despite his belief that he is doing his duty because it is his duty. He is, then, acting *from duty but not in accord with duty*."[49]

I will henceforth refer to the claim that action from mistaken moral judgement *is* action from duty but not in accord with duty as the *Identity Thesis*. Taken together, the assumptions expressed in these two theses lead to the following picture:

action / motive	RIGHT (required)	WRONG (impermissible)
NON-MORAL (inclination)	**Morally right action** 'Action in accordance with duty but not from duty' **(no moral worth)**	**Immoral action** 'Action neither in accordance with duty nor from duty' **(no moral worth)**
MORAL (duty)	**Morally right and praiseworthy action** 'Action both in accordance with duty and from duty' **(moral worth)**	**Action from mistaken moral judgement** 'Action from duty but not in accordance with duty' **(moral worth?)**

(Diagram 1)

[49] Ibid. p. 69 (my emphasis). More specifically, and in relation to Kant's universal law formula of the Categorical Imperative, "immoral actions are actions based on maxims which cannot be universalized without contradiction. But contradictions need not be apparent to one who holds contradictory beliefs, and they do not always reveal themselves to agents who conscientiously look for them. In Kant's terms then, an action from duty, but not in accord with duty would be an action which the agent believes he can will to be a universal law but which cannot in fact be universalized without contradiction." (Ibid.)

If accurate, this picture points to a serious shortcoming of Kant's account of duty. If, in accordance with the Logical Compatibles View of Kant's twofold distinction of duty, this distinction does indeed yield four logically possible classes of action, including 'action contrary to duty but from duty', and if in line with the Identity Thesis the latter is the appropriate description, in Kantian terms, of action from mistaken moral judgement, then it seems indeed "striking that [in the *Groundwork*, i.e.] in a work where Kant is at pains to explore the concept of duty, he mentions not one example of an action done from duty but which conflicts with duty"[50]. Moreover, as is well known, Kant's famous analysis of the concept of duty in the *Groundwork* is primarily concerned with the distinction between morally correct action which merely *coincides* with duty and morally correct action which is done *from* duty, thus giving the action *moral worth*. Since Kant holds that all actions done from duty have moral worth, but also appears to insist that only action in accordance with duty can be done from duty, there is a puzzle as to how action contrary to duty but from duty is even possible. If it is possible, there is the further question whether such action, *qua* action from duty, has moral worth, or whether, *qua* action contrary to duty, it lacks such worth.[51]

---

[50] Kerstein, op. cit., p. 208 (fn. 7). Cf. Hardwig, op. cit., p. 68: "Kant never considers or even mentions [...] action from duty but not in accord with duty. This is perhaps surprising in a philosopher with Kant's interest in logic and passion for thoroughness. One would have thought that he would mention this logical possibility, even if only in order to discount it as not really possible."

[51] Since one of my aims is to deny that action from erroneous moral judgement should be translated into Kantian terms as 'action from duty but not in accordance with duty', and, further, that, strictly speaking, the latter term is meaningless, I am somewhat reluctant to cite examples of such action here, for fear of further confusing the issue. My arguments in this part of the investigation imply that the proper place to discuss examples of such action would be in the context of action from mistaken moral judgement, *not* in terms of action from duty but not in accordance with duty, the analysis of which is my present concern. Nevertheless, if these points are kept in mind, the following examples of action from erroneous moral judgement are some of the most prominent in the literature, and hence they might be instructive: Jonathan Bennett thinks that Himmler committed atrocities despite his sympathy for his victims, i.e. acting contrary to his inclinations, because he erroneously thought it was his duty to do so in a desperate bid to save western civilization. Bennett's other example is Jonathan Edwards, who gave his approval to the notion of God's eternal tormenting of the damned, because he, again, erroneously took the endorsement and propagation of such views to be his duty. See his 'The Conscience of Huckleberry Finn', *Philosophy*, vol. 49 (1974), pp. 127-131. Hardwig mentions several more examples: "The inquisitor torturing a heretic in a last ditch attempt to save his soul; [...] a social scientist marshalling arguments to convince his students that morality is nothing more than social mores; a woman destroying her potential and her life in an attempt to honour, love and obey her husband." Cf. Hardwig, op. cit., p. 69. He also suggests that it could be argued "that anyone acting as Kant recommends in 'On the Supposed Right to Lie from Altruistic Motives' would merit precisely this description." (ibid.) In an example reminiscent of

4. The importance of this problem and the aims of part II.

So of the many interesting and controversial issues that Kant's account of duty has given rise to, I will here focus on a question which has been largely overlooked in the literature: the possibility of action from duty but not in accordance with duty and Kant's apparent oversight of this type of action, as well as the related issue concerning the analysis of action from erroneous moral judgement and the question whether such action does, or should, on Kant's account have moral worth.

This issue is, I believe, of considerable importance in the context of Kant's

---

Kant's description of Frederick the Great, who carried "a fast-acting poison with him, presumably so that if he were captured when he led his troops into battle he could not be forced to agree to conditions of ransom harmful to his state" (MS VI:423), Kerstein describes the case of Colonel Mikavitch: well-educated and morally reflective, she has embraced the Categorical Imperative as the supreme principle of morality. Finding herself in a situation where she will be captured and, under torture, reveal the location of innocent people who her captors are intent on killing, she considers committing suicide. After careful deliberation, against her own inclination, and with great courage she decides that taking poison in order to take her own life is the only way she can save innocent lives. Kerstein urges, for reasons which I will consider in chapter 10, that her action should have moral worth whether or not one considers her action to conform with duty.

To add to these a recent example: In 'The Heretic', a special investigation into an Al-Qaeda leader who turned his back on terror, Lawrence Wright describes the predicament which imprisoned members of the Islamic Group found themselves in when they were confronted with, and eventually accepted, the new doctrine of their spiritual leader, whose writings had previously justified global terror but now "denounced the use of violence – in a stroke undermining the entire intellectual basis of jihad". He goes on to describe the prisoners' concerns in the following terms: "What about the brother who was killed while carrying out an attack that we now realise was against Islam? Is he a martyr? If not, how do we console his family? One of the leaders proposed that if the brother who died was sincere, although genuinely deceived, he would still gain his heavenly reward; but because 'everyone knows there is no advantage to violence, and that it is religiously incorrect', from now on such actions were doomed." (in: *The Observer Magazine*, 13 July 2008, p. 31)

Of course, there are important differences between the examples and, more importantly, it is obviously questionable, if not unlikely, that the historical characters mentioned did in fact act autonomously and from duty, rather than from other motives. The point here is simply whether, on the assumption that they did indeed act from (what they perceived to be their) duty, their actions should be described as *action from duty but not in accordance with duty*, as commentators commonly assume it should be, and this alone is my present concern. There is a further question whether such action can, or should, on Kant's account be thought to have moral worth, a question which I will not consider until the last chapter.

ethics[52], for several reasons: Of course it would be curious if Kant, having made a distinction that gives rise to four types or classes, had committed the obvious blunder of overlooking one of the four. So the completeness or otherwise of his classifications is in itself of some interest. Secondly, if Kant's intention, at least in the *Groundwork*, is to identify the motive of duty in actions that are done *from duty* (the ultimate aim of which is to illuminate the concept of a good will), and since actions from duty but not in accordance with duty are, after all, actions *from duty*, one might expect that he would consider them relevant for the purposes of that identification, or, at the very least, that he would explain why he does *not* consider them relevant. Thirdly, if, as most commentators agree, the common phenomenon of action from mistaken moral judgement *is*, in Kantian terms, action not in accordance with duty but from duty, then Kant's oversight of this kind of action would suggest that he is ignorant of a significant feature of the life of the moral agent.

In the following chapters I will reject the Logical Compatibles View of Kant's twofold division by arguing that, contrary to appearances, Kant's classification of types of actions *is* complete. I will show why he *must* hold that actions not in accordance with duty but from duty are *impossible* (i.e. that he is consistent on his own terms); but also, and more importantly, that he is right to do so. I will then reject the Identity Thesis by showing that the very notion of action from duty but not in accordance with duty is incoherent. I will offer an alternative analysis of mistaken moral judgements, the upshot of which is a complete separation of this issue from the question concerning the possibility of action contrary to duty but from duty. Of course, this in itself does not address the question regarding action from mistaken moral judgements, but it will free it from the constraints that have resulted from the

---

[52] Or, for that matter, for any moral theory that recognises that agents' motives have a value which is distinct from the rightness of their actions. I have argued elsewhere (see chapter 2) that the insistence on a motivational criterion of moral worth, which is often thought to be a distinguishing feature of Kantian ethics, is in fact neither originally nor typically Kantian, but, on the contrary, widely held by moral philosophers of various convictions. Aristotle and Mill, whose thinking is opposed to Kant's on so many other issues, subscribe to such a requirement and recognise, however varied their criteria of rightness might be, the importance of motives in the moral estimation of actions. Like Kant they insist that the primary object of praise and blame is not the action's rightness or wrongness, but the motive that produced it. And like Kant they take this point to be no novel philosophical claim but an idea that is already implied by our common understanding of morality. Cf. Aristotle, *Nicomachean Ethics*, transl. Terence Irwin, Indianapolis/Cambridge: Hackett, 2nd ed. 1999, 1105a18-1105b6; John Stuart Mill: *Utilitarianism*, (London: Longmans, Green & Co.), 1897 (13th ed.), p. 26, p. 27 (fn.).

confusion of the two issues which dominates the literature and open the way for a new approach to the question. In chapter 10 I will consider briefly some of the ways in which this question might be addressed, especially to what extent Kant's theory can accommodate the notion that some actions from erroneous moral judgements might have moral worth (or possibly some similar sort of value). Finally, I believe that my analyses will shed light on some important features of Kant's notion of duty which are commonly ignored or misunderstood.

In chapters 5-8 I examine Kant's apparent oversight of the possibility of action from duty but not in accordance with duty. While the main business of part II. is the critique of the two assumptions (the Identity Thesis and the Logical Compatibles View), I will begin by addressing a number of considerations which could be thought to go at least part of the way towards accounting for Kant's apparent omission. In doing so, I will grant the two assumptions and endeavour to show that even on their own terms these attempts must fall short of the target or fail otherwise.

Chapter 5

THE POSSIBILITY OF ACTION FROM DUTY BUT NOT IN ACCORDANCE WITH DUTY: SOME GENERAL CONSIDERATIONS

Before examining the textual evidence, I will begin by considering briefly three common but inadequate explanations of Kant's apparent failure to deal with actions from duty but not in accordance with duty. They are (1.) the suggestion that the *Groundwork* is not the appropriate place to discuss *any* actions which are not in accordance with duty; (2.) the conceptual claim that Kant simply *defines* the concept of duty in a way that rules out the possibility of such actions; and (3.) the view that such actions are impossible, or at least practically irrelevant, because their possibility rests on a failure on the agent's part to ascertain correctly what his duty is, and that according to Kant there is no problem of knowledge in morality, making such failure impossible, or at least very unlikely and hence practically irrelevant.

1. The suggestion that the scope and purpose of the *Groundwork* precludes a consideration of wrong actions

The first two of these explanations are closely related and commentators do not always distinguish between them[53], but it seems clear that (1.) is a more general point about the scope and purpose of the *Groundwork*, i.e. about what Kant does and what he does not attempt to achieve in this work. The suggestion is that the discussion of the concepts of a good will and duty in the first section, or even the *Groundwork* as a whole, is not an appropriate place to consider *any* actions which are contrary to duty. Kant deliberately sets aside *all* such actions precisely because his purpose is to gain insight into the nature of morality and because he is driven by the overall aim of the *Groundwork*, which is "nothing more than the search for and establishment *of the supreme principle of morality*" (GMS IV:392). Consequently, he looks at actions that are *at least* in conformity with duty, i.e. with the supreme principle of morality, and then inquires whether and under what circumstances they might be done *from duty* (and therefore have moral worth).

[53] Consequently I will discuss them in conjunction in section 2. below.

> [Kant] sets out to find clear cases of action motivated by duty. [...] Of course, no [...] insight [into the nature of morality] is to be gained from an examination of actions that violate duty [...]. [...] A moral action must be done *because* the law commands it. Only a clear example of such action *from duty* will advance the present project of identifying its principle. [...] Kant proceeds to exclude actions contrary to duty, no matter how useful, for the reason already mentioned: their investigation does nothing to advance the search for the supreme principle of morality.[54]

2. The claim that moral worth presupposes rightness

According to (2.) a more specific conceptual point can be made about the logical relationship between the criterion of rightness and the criterion of moral worth. "It may be that Kant has chosen from the outset of the *Groundwork* to use the expression 'from duty' to refer only to actions that one does because one believes they are right *and that according to Kant's standard are indeed right.*"[55] In the preface of the *Groundwork* Kant makes it clear that "as to what is to be morally good, it is *not enough* that it conform to the moral law, but it must *also* happen for the sake of this law." (GMS IV:390, my emph.).

> In other words, for an action to have moral worth (be morally good), it must *both* be done from duty (for the sake of the law) and be in accordance with duty (conform with the moral law). Here Kant implies that if an action has moral worth, it is in accordance with duty. Since for Kant all actions from duty have moral worth, it follows that all actions from duty are in accordance with duty.[56]

According to this line of thought Kant makes conformity with duty a necessary but

---

[54] Jens Timmermann, *Kant's Groundwork of the Metaphysics of Morals. A Commentary* (Cambridge: Cambridge University Press, 2007), p. 27f. Cf. Paul Guyer, *Kant's Groundwork for the Metaphysics of Morals. A Reader's Guide* (London: Continuum, 2007), p.43, and Samuel J. Kerstein, *Kant's Search for the Supreme Principle of Morality* (Cambridge: Cambridge University Press, 2002), pp. 96ff., who notes that in "his famous exploration of cases in Groundwork I, Kant is attempting to elucidate the concept of a good will. With the help of the concept of duty, he is trying to clarify [...] when our actions are morally good – that is, when they have intrinsic, moral worth, which is the kind of worth characteristic of a good will. Ultimately, Kant aims to pinpoint the principle of a good will: the supreme principle of morality [...]. [...] [I]n Kant's view, we cannot hold a principle to be the supreme principle of morality unless we can maintain that no actions that fail to conform to it can have moral worth."

[55] Kerstein, op. cit., p. 97, emphasis added.

[56] Ibid., p. 98, emphasis added.

not sufficient criterion for moral worth, such that "it simply belongs to the concept of an action done from duty that it is in accordance with it"[57]. If that were the case it would indeed be futile to search for actions done from duty amongst the class of those that fail to conform with duty, since they are excluded from that class: Kant defines the term 'duty' such that no action from duty is contrary to it. Hence actions from duty but not in accordance with duty would simply be impossible *per definitionem*.

In replying to these first two points I will concede straight away that the *Groundwork* is not the most obvious place to consider the possibility of action from duty but not in accordance with duty, nor is a discussion of the possibility of such action Kant's aim there. Neither do I wish to question the validity of the general point (1.) *per se*, or even to imply that those who make it do so in order to address the possibility of action from duty but not in accordance with duty. I merely want to insist that in the present context it would not be sufficient to adduce this general consideration in an attempt to *explain* Kant's apparent omission of actions contrary to duty yet from duty. Rather, doing so would simply *attribute* to Kant the view that such actions are irrelevant in the particular context of the search for the supreme principle of morality, and that consequently they can be safely ignored. Significantly, this in turn leads to the natural assumption that such actions are very much possible and that Kant *should* consider them, if not in the *Groundwork* then perhaps elsewhere, or even that he *does* indeed consider them elsewhere.[58]

I also agree that *if* Kant did in fact hold that accordance with duty is a necessary condition for acting from duty then this would explain why he ignores actions contrary to duty but from duty. However, simply to stipulate this by claiming that Kant *defines* the term 'from duty' in a certain restricted sense does nothing to explain the fact: it

---

[57] Ibid., p. 97. This claim is also made or implied by others. Henson, for example, remarks that "the question of dutifulness is different from another interesting question which can be asked about an act, though *only about an act which is dutiful*, that is, required by duty: of such an act one can ask the *further* question whether it has moral worth." See Richard G. Henson, 'What Kant Might Have Said: Moral Worth and the Overdetermination of Dutiful Action', in: *Philosophical Review* 88 (1979), p.40 (emph. added).

[58] Thus some writers suggest that he does so in the *Metaphysics of Morals*. Cf Kerstein, op.cit., pp. 124-127, who argues that Kant in the later work "moves toward, though he dos not explicitly embrace, the possibility of morally impermissible actions having moral worth" (p. 124); and Allison op. cit., p.113, p. 265 fn. 4, who argues that in the *Groundwork* "Kant dismisses [...] the possibility [...] that an action done *from* duty might not, objectively speaking, accord *with* duty [...], although he does recognize it elsewhere" (p. 113).

would only explain *why* Kant sets aside all actions contrary to duty if we already assume what is at stake, i.e. that no action contrary to duty can be done from duty. To put this point another way; if Kant does indeed make rightness a necessary condition for moral worth then the question really is *why* and *how* he does so. Without such an explanation we do not understand why there could not be some actions which are contrary to duty yet done from a moral motive, nor can we see any ground for ruling them out as possibly illuminating in the attempt to pinpoint the moral motive. For instance, Kant might have described an agent who, sincerely but mistakenly taking some act to be his duty, performs this act, and then said about *this* agent what he does in fact say about the agent who sincerely *attempts* to act from duty and in accordance with duty, but "with [his] greatest efforts [...] achieve[s] nothing", failing entirely "to carry out [his] purpose" (GMS IV:494): since moral worth is determined by the agent's willing, his action has moral worth.[59] But Kant does *not* say this, and to claim that Kant simply stipulates that no action contrary to duty can have moral worth does nothing to explain this fact.

3. The view that there is no problem of knowledge in morality

Finally, there is a third line of argument which has been utilised in the attempt to explain Kant's apparent ignorance of the phenomenon of action from duty but not in accordance with duty. As we saw in the introduction, the Identity Thesis claims that to act on a mistaken conception of one's duty is, in Kantian terms, to do what one *takes to be* one's duty and do it *because* one takes it to be one's duty, thus acting *from* duty, but to be *mistaken* in one's judgement, thus acting *contrary* to duty. The possibility of such action obviously rests on the possibility that one might fail to ascertain correctly what one's duty is. But, according to (3.), perhaps Kant believed that knowing one's duty does not pose a problem, or at least not a significant one. Thus Hardwig observes that for

> Kant, as opposed to Plato, the paradigm problem of moral action is not a problem of knowledge, but a problem of volition. The paradigm moral situation is not one in which the agent may act wrongly because he mistakenly believes that he is doing what he ought, or one in which the agent struggles with the problem of discovering what he ought to do; rather it is the situation in which the agent struggles to do what he knows he should do. It is also clear that Kant thought that the primary role for a philosophical theory of ethics is to resolve the

[59] I consider this suggestion in chapter 10.

> theoretical problems of the philosopher, not to guide action by illuminating the practical difficulties of the moral agent.[60]

According to this picture, knowing what is right is not as much of an issue for Kant as it is for Plato (and others), nor is the problem of knowledge the most pressing problem facing the moral deliberator and agent. And when we consider various statements scattered throughout his ethical writings it is indeed clear that Kant attaches much greater significance to the problem of volition and motivation – i.e. actually carrying out what one knows to be one's duty – than to the problem of knowledge. *Knowing* one's duty then is for Kant not the paradigmatic moral problem. But we can go further and note that he was also convinced that determining one's duty is *easy*, and hence that the problem of knowledge in morality, insofar as it exists at all, can be overcome without much difficulty by 'common human reason'. In the following passage from the second *Critique*, Kant is at pains to stress the relative ease with which we know our duty, especially compared to the difficulties involved in figuring out how to satisfy our desire for happiness.

> What is to be done in accordance with the principle of the autonomy of choice is seen *quite easily and without hesitation by the most common understanding*; what is to be done on the presupposition of heteronomy of choice is difficult to see and requires knowledge of the world; in other words, what *duty* is, is plain of itself to everyone, but what brings true lasting advantage, if this is to extend to the whole of one's existence, is always veiled in impenetrable obscurity, and much prudence is required to adapt the practical rule in accordance with it to the ends of life even tolerably, by making appropriate exceptions. But the moral law commands compliance from everyone, and indeed the most exact compliance. Appraising what is to be done in accordance with it must, therefore, *not be so difficult that the most common and unpracticed understanding should not know how to go about it*, even without worldly prudence. (KpV V:36f., emph. partly mine)

And again in the *Groundwork*, immediately following the identification of the first formula of the supreme principle of morality (the universal law formula of the Categorical Imperative), Kant says that "common human reason, indeed, agrees perfectly with this [principle] in its practical judgement, and has the principle always before its eyes" (GMS IV:403), and goes on to claim that it "would be easy here to show how, with this compass in its hand, *it knows its way around very well in all the cases that come before it*, how to distinguish what is good, what is evil, what conforms to duty

[60] Hardwig, op. cit., p. 70.

or is contrary to duty" (GMS IV:404, emph. added).

Of course, it is one thing to say that knowing one's duty is *relatively* easy, or even that common human reason, generally speaking, knows what duty is. The question, however, is whether we can go further again and attribute to Kant the view that there is *no* problem of knowledge in morality, as perhaps the last citation might suggest ("it knows its way around very well in *all* the cases that come before it", GMS IV:404, emph. added). For obvious reasons it is this stronger claim that will need to be examined in the present context: for one might think that if what duty is is easily seen by the most common understanding, then there is no problem of knowledge in morality, and if there is no problem of knowledge in morality then there are no mistaken moral judgements. And if one assumes further (with the Identity Thesis) that actions from mistaken moral judgement *are* (in Kant's terms) actions contrary to duty but from duty, then it follows that there can be no such actions. Some commentators claim that the passages cited and similar remarks by Kant[61] are indeed significant enough to attribute to him the view that there is *no* problem of knowledge in morality, and further, that this in itself is the best explanation available for his apparent omission of actions from duty but not in accord with duty. Thus Kerstein, commenting on a crucial *Groundwork* passage which I will examine shortly, argues that attributing this view to Kant constitutes the most compelling explanation of Kant's apparent claim that no action which is contrary to duty can be done from duty. It is

> compelling because it reveals how remarks Kant makes elsewhere in the *Groundwork* might explain the claim. Consider Kant's emphasis in this work and elsewhere on how easy it is to determine what our duties are. [...] Perhaps, then, he reasons thus. The ultimate ground of an action done from duty is the agent's notion that the action is morally required. But it is very simple to figure out whether doing something is morally required. Therefore, if someone does something contrary to duty, he has obviously not been motivated by the notion that doing it was morally required. In short, Kant might hold an agent's duties to be so transparent to her that she just could not both be motivated by the notion that she is required to fulfill them yet violate them. [...] [W]hat duty requires is so transparent that any agent who genuinely acts from the notion that doing something is morally required will succeed in abiding by his duty.[62]

---

[61] Cf. also GMS IV:402-405, KpV V:27, Theory & Practice VIII:286 f., KrV A831/B859.

[62] Kerstein, op.cit., pp. 97f. Cf. Hardwig, op. cit., pp. 69f.: "To say that there are no actions from duty but not in accord with duty is to say that there are no mistaken moral judgements, perhaps even that the whole idea of a mistaken moral judgement is itself a mistake. If there are no mistaken moral judgements, then there is no moral problem of knowledge. [...] I wish to suggest that Kant may very well have held that there is no moral problem of knowledge and hence no actions from duty but not in accord with duty."

Again, as with the account that has Kant simply define the concept of duty in a restricted sense, the upshot is that according to Kant no action from duty is contrary to it and action from duty but not in accordance with duty is therefore impossible.

But does Kant hold that there is no problem of knowledge in morality, and that human beings fail to act in conformity with the moral law only when overcome by their desires, never because they fail to determine their duty correctly? Unfortunately the textual evidence is inconclusive. The passages that might support this view are balanced by others, especially in the later works, which highlight the frailty of human beings and stress the need for philosophical reflection in the service of moral enlightenment and the improvement of character. Although *typically* Kant attributes these shortcomings to the 'dear self' and its tendency to prioritise the satisfaction of its desires over reason's commands, he does not do so *exclusively*. Even very early on in the *Groundwork* there are suggestions that Kant does not hold that when an agent acts contrary to duty he *always* does so because he is overcome by his inclinations and despite his better knowledge. Knowledge can not be taken for granted in morality, for the recognition of a priori moral laws requires of the rational human being

> a *power of judgement sharpened through experience*, partly *to distinguish in which cases they have their application*, and partly to obtain access for them to the will of the human being and emphasis for their fulfilment, since he, as affected with so many inclinations, is susceptible to the idea of a pure practical reason, but is not so easily capable of making it effective *in concreto* in his course of life. (GMS IV:389, emph. partly mine)

Here Kant seems to allow that it is possible that an agent might try, yet fail, to judge correctly what his duties are. One might lack either experience or judgement (or both) to apply moral laws correctly to particular cases, or to distinguish where and how they are applicable. He goes on to insist that as a result of these human frailties there is a practical need for the philosophical clarification of the foundation of morality ("Thus a metaphysics of morals is indispensably necessary…", ibid.). In the discussion of the role of conscience in the *Metaphysics of Morals* Kant goes even further when he explicitly acknowledges that "I can indeed be mistaken at times in my objective judgement as to whether something is a duty or not". (MS VI:401)

Can these two seemingly contradictory strands of Kant's thought be reconciled or at least explained?[63] As we have seen, Kant, on the one hand, takes the "truth of

[63] To the cynical observer these tensions within Kant's thought might well have an obvious and

reflective common morality [to be] the default position"[64]. Further, his reasons for doing so are *themselves* largely ethical: "Human beings must have access to moral truth to be responsible agents at all."[65] On the other hand, Kant has of course equally good reason to insist that it is possible for human beings to err in their moral judgements. He might be overly optimistic in his assessment that pre-philosophical moral thought cannot be mistaken, or at least not radically mistaken[66], he might even be naïve in thinking that knowing our duty is easy; but it would be absurd to attribute to him the view that we *cannot* be mistaken in the determination of our duties – that would make human beings infallible.[67] It is important to recall that the issue here is *not* whether or to what extent failure to know one's duty is *imputable*. This is of course an important question when it comes to the evaluation of mistaken moral judgements, and I will return to this topic in chapter 10. But the suggestion I am considering here is that if there is no problem of knowledge in morality, i.e. if knowing one's duties is so easy as to make it impossible to go wrong, then this would explain Kant's apparent omission. Consequently the question here is merely whether Kant held that such failure – imputable or not – is *possible*. If such failure is possible then the explanation under consideration – whatever other merits it might have – fails.

The upshot of all of this in the present context is that there is a tension between Kant's remarks that suggest that moral error is possible on the one hand, and the attempt to attribute to him the strong thesis that there is *no* problem of knowledge in morality on the other. Only this strong claim will suffice if the question concerning the problem of knowledge in morality is to have any sort of bearing on Kant's

---

somewhat prosaic source: his attempt to be popular and original at the same time. Kant's apparently paradoxical view that, while the moral law is plain and obvious to even the commonest intelligence, there is still a great need for a metaphysics of morals for there to be any true virtue, would thus be explained by the constant tension between his attempt to base criteria for the supreme principle of morality on an analysis of commonly held views, and his insistence that his own philosophically derived (i.e. *not* common-sense) version of this supreme principle is a substantial and original philosophical achievement.

[64] Timmermann, op.cit. p. xii.

[65] Ibid.

[66] Ibid.

[67] One feature of his ethics that might be thought to explain, at least partially, Kant's confidence that common reason does not err in judging the requirements of duty is the fact that for Kant the strictest duties are often duties of omission (such as duty not to lie) which tend to present *fewer* cognitive problems than duties to perform positive actions (such as a duty of assistance). However, since the suggestion I am considering here is that there is *no* problem of knowledge in morality, any such partial explanation must fall short of the mark.

apparent omission of action from duty but not in accordance with duty. Such an explanation would only succeed if Kant did indeed hold that there is *no* problem of knowledge in morality; it fails if he merely thinks that knowledge is not the *main* issue or that it is *comparatively* easy to know one's duties. So the remarks which strongly suggest that Kant allows for the possibility of moral error constitute a significant hurdle to attributing the strong thesis to him.

But if Kant did not subscribe to the strong thesis, then we are left with only the weaker claim that failure to know one's duty, while not inconceivable, is insignificant enough or unlikely enough to make it irrelevant *in practice*. In relation to our current concern with explaining Kant's apparent omission this would allow us to venture that he

> did not [...] overlook the possibility of acting from duty yet contrary to it; rather, based on his conviction that it is very simple to determine what one's duty is, he rejected this possibility as *practically irrelevant*.[68]

As I will return to the issue of mistaken moral judgements in chapter 10, I will refrain from debating here whether or to what extent the problem of knowledge might be practically irrelevant, or to what extent Kant thought that it was. I also grant that the issue is not a priority in the *Groundwork*. For present purposes I consider it sufficient to have argued that for this issue to have any bearing on the question of Kant's apparent oversight of actions from duty but contrary to it, we must take him to make the strong claim that there is no problem of knowledge in morality, and that for the reasons outlined it is questionable that Kant would have held such a view.[69]

Lastly, to add to my critique of these three common but inadequate explanations of Kant's apparent lacuna (1.-3. above) one final consideration: If concrete textual evidence is available which suggests that Kant *does* address the possibility of action contrary to duty but from duty, then this should be preferred over more general

---

[68] Kerstein, op.cit., p. 98 (emph. added).

[69] Note that my overall argument concerning the possibility of action from duty but not in accordance with duty does not rest on this point. I have rejected the explanation under consideration because I believe it cannot show convincingly what it needs to show in order to account for Kant's apparent omission; namely that Kant holds that there is *no* problem of knowledge in morality. But this explanation, even if successful, rests on the Identity Thesis and the Logical Compatibles View, both of which I have granted here for the sake of argument. But as I will argue shortly, they should both be rejected.

considerations or conceptual stipulation aimed at explaining why he does not address it. It is to this textual evidence that I now turn. In chapters 6 and 7, I will critically examine the common interpretations of the textual evidence, before offering my own interpretation in chapter 8.

Chapter 6

KANT'S EXCLUSION OF WRONG ACTIONS FROM THE INVESTIGATION INTO THE MOTIVE OF DUTY

Having introduced the concept of duty in order to "explicate the concept of a [good] will" (GMS IV:397.1), Kant begins his search for actions which are done from duty, and in which therefore a good will "shine[s] forth" (GMS IV:397.9f.), with the following statement:

> I here pass over all actions that are already recognised as contrary to duty, even though they might be useful in this or that intent; for with these there is not even the question whether they might have occurred from duty, since they even conflict with it. (GMS IV:397.11-14)[70]

It has been a matter of some, although very limited, debate in the literature, whether the expression "actions that are already recognised as contrary to duty" here refers to actions which we, the readers and observers, recognise as contrary to duty (while an agent might mistakenly take the same action to be in accordance with duty), or whether, on the contrary, the recognition of the action's impermissibility is the agent's. This is an important question to which I will return shortly.

In the first place, however, it is noteworthy that the above statement is crucial because it contains the *only* mention of any actions contrary to duty in the context of

---

[70] My translation. As much of what I have to say in the following chapters centres on this passage, I deliberately translate rather literally from Kant's German, probably at the expense of smoothness. However, I do so merely to show that nothing in my arguments concerning the correct interpretation of this passage hinges on the subtle differences between translations. The original reads: "Ich übergehe hier alle Handlungen, die schon als pflichtwidrig erkannt werden, ob sie gleich in dieser oder jener Absicht nützlich sein mögen; denn bei denen ist gar nicht einmal die Frage, ob sie aus Pflicht geschehen sein mögen, da sie dieser sogar widerstreiten." Cf. Wood: "I pass over all actions that are already recognized as contrary to duty, even though they might be useful for this or that aim; for with them the question cannot arise at all whether they might be done *from duty*, since they even conflict with it." Paton: "I will here pass over all actions already recognized as contrary to duty, however useful they may be with a view to this or that end; for about these the question does not even arise whether they could have been done *for the sake of duty* inasmuch as they are directly opposed to it." Ellington: "I here omit all actions already recognized as contrary to duty, even though they may be useful for this or that end; for in the case of these the question does not arise at all as to whether they might be done from duty, since they even conflict with duty."

the entire discussion of the concept of duty in the first section of the *Groundwork*. Having made this point, Kant immediately turns his attention to actions that are at least in conformity with duty which he then investigates as to their possible moral worth. So within the analysis of the concept of duty in the *Groundwork* the passage quoted offers the only direct textual evidence of Kant's views on actions which are contrary to duty.[71]

Before I return to the main controversy I will have to address an exegetical issue concerning the quoted passage, for the answer to the following question is not as obvious as it might appear: Precisely which actions does Kant "pass over" (GMS IV:397.11) here? I will begin by considering a reading of the passage which has been suggested in a recent commentary on the *Groundwork*[72], but which seems to me to be *obviously* wrong.

It might seem clear that Kant, in the passage cited, "pass[es] over" (GMS

---

[71] In the well-known passages that follow the statement cited above, Kant goes on to "also set aside" (GMS IV:397.14) *some* actions which are in accordance with duty, namely those to which human beings have no *immediate* inclination, but which they typically perform because they are a means to something else which is in their self-interest. (Somewhat confusingly for his readers, and in contrast to actions contrary to duty – which he says he will 'pass over' and which he indeed never mentions again – Kant professes to 'also set aside', i.e. *not* engage with, this second type of action, but then proceeds immediately to entertain at some length an example of these: the prudent shopkeeper who does not overcharge even an inexperienced customer because it would harm his long-term business interests.) He disregards these latter actions because he believes they are *easily recognised* as being "done neither from duty nor from immediate inclination, but merely from a self-serving aim" (GMS IV:397.31). The distinction between the motive of duty and the motive of inclination "is *much harder to notice* [...] where the action is in conformity with duty and the subject also besides this has an *immediate* inclination to it" (GMS IV:397.19ff., my translation)

So having set aside all actions *contrary* to duty as well as self-interested actions that *conform* with duty, Kant is now left with only one type of action to throw light on the distinction between acting merely in accordance with duty and acting from duty (and ultimately to elucidate the concept of a good will): actions which accord with duty and towards which the agent has an *immediate inclination*. Kant discusses three examples: the duty to preserve one's life, the duty of beneficence, and the duty to assure one's own happiness. He argues that we commonly act in conformity with these duties because we have an immediate inclination to look after our lives, be happy, and be sympathetic to others. The unfortunate man who refrains from suicide without loving his life, the cold-hearted philanthropist who lacks all sympathy for others, and the gout sufferer who refrains from enjoying what he likes for the sake of his long-term happiness all act *in conformity* with these same duties, but since *ex hypothesi* they lack all natural inclination towards doing these duties, they act *from* duty.

[72] Sally Sedgwick, *Kant's Groundwork of the Metaphysics of Morals. An Introduction*, (Cambridge: Cambridge University Press, 2008), pp. 59ff.

IV:397.11), that is, excludes from further consideration, *all* actions which are recognised as contrary to duty. But according to an interpretation suggested by Sally Sedgwick, the passage is concerned more specifically with one particular type of action which is contrary to duty, namely with cases of "[a]cting in a way that is *obviously* contrary to duty"[73]. On this reading, actions which are "already recognised as contrary to duty" (GMS IV:397.11) are "the *most obvious cases* of what a good will is *not*"[74] and this explains why Kant does not consider them. According to Sedgwick we should read 'already recognised' as '*easily* recognised', and when we observe actions which are obviously wrong since we easily identify them as violations of duty, we can infer from the agent's observable behaviour that duty is not her motive. For instance, if we observed an act of cheating,

> Kant says, the question 'never arises' whether the action is performed from duty [...], because no one would for a moment consider it appropriate to characterize the cheater's will as good. Her cheating behavior is clear evidence that she acts contrary to duty. [...] [W]e already recognize in the classmate's act of peering furtively at her neighbor's work that she does not act in conformity with duty. In cases like this, we can confidently infer from the agent's behavior a less than honourable intention. This is most likely why Kant remarks that [...] the question 'never arises' whether the agent acts from duty.[75]

In response, I will first argue that this interpretation is internally inconsistent: Even on the assumption that Kant's statement is concerned only with actions which are *obviously* wrong, Sedgwick does not succeed in justifying her assertion that such actions can be known with certainty to lack the motive of duty. There is, to be sure, a way to make such a claim, but it is incompatible with a number of assumptions Sedgwick makes. In any event, my second and main aim is to show that her reading of 'already recognised' as 'easily recognised' fails as an interpretation of the text.

I will begin by reiterating a general point: By itself, the inference from the wrongness, obvious or not, of a particular action to the absence of the motive of duty in that action is only valid if one holds as a general principle that no action which is contrary to duty can be done from duty. But it is not obvious that this last claim – which is precisely the assumption in question in the present context – is true, or that Kant thinks it is. It is clear, however, that Sedgwick's argument rests on this premise: If (P1) the cheater's behaviour is sufficient to identify her action as (obviously) contrary

---

[73] Ibid., p. 60 (emph. added)

[74] Ibid. (my emph.)

[75] Ibid., p. 60f.

to duty, and if (P2) no action which is (obviously) contrary to duty can be done from duty, then (C) we are in a position to "confidently infer" (ibid.) from her behaviour that she has not acted from duty (and since Kant holds that only actions done from duty have moral worth we also know that her action does not have moral worth). This argument is not valid without (P2), but Sedgwick fails to explain why (P2) might be true or why Kant might subscribe to it, simply assuming that he does[76], and consequently on her account we remain in the dark as to why actions which are *easily* recognised as wrong could not be performed *from* duty.

1. Knowledge of other agents' motives

But more importantly in the present context, the assertion that actions which *obviously* violate duty can be known with certainty not to have been done from the motive of duty is incoherent in the way in which Sedgwick makes it: As she acknowledges herself[77], Kant is notoriously sceptical about our ability to know with any certainty the motivational grounds of even our own actions, never mind another person's. Her claim that we can confidently make inferences from an agent's behaviour to the underlying motive appears to be at odds with Kant's repeated insistence that agents' motives are unknowable, or at the very least that they can never be known with any certainty. Kant makes this point most frequently with reference to the possibility of knowing in any particular case that an action has been done from the motive of duty alone, i.e. that no inclination towards the action moved the agent.

> If we attend to *the experience of the deeds and omissions of human beings*, we encounter frequent and, as we ourselves concede, just complaints that one could cite *no safe examples* of the disposition to act from pure duty; that, even if some of what is done may accord with what duty

---

[76] Admittedly, she is not alone in attributing (P2) to Kant, and I discuss the reasons why Kant might have subscribed to it below. As will become apparent then, I do not think that ascribing this view to Kant does anything to explain the current difficulty – his apparent omission of the possibility of action from duty but not in accordance with duty. But in addition, there are further reasons – reasons which go beyond the issues raised in the current context – which make it controversial whether Kant did in fact hold, and even more so whether he should have held, that no action which an agent performs from the mistaken notion that it is required can have moral worth. (For an outline of that debate, see chapter 10.) In any case, my overall argument against Sedgwick's reading of the passage does not rest on this particular point.

[77] Ibid., pp. 84f.

> commands, nevertheless it always remains doubtful whether it is *really done from duty* and thus has a moral worth. (GMS IV:406, emph. partly mine)

He goes on by claiming that it is in fact

> absolutely impossible to settle with complete certainty through experience whether there is even a single case in which the maxim of an otherwise dutiful action has rested solely on moral grounds and on the representation of one's duty [...], because *when we are talking about moral worth, it does not depend on the actions, which one sees, but on the inner principles, which one does not see.* (GMS IV:407, emph. added)

So according to Kant the presence of a purely moral motive can never be ascertained in any particular act. The morally worthy disposition to act from duty is hidden from our view and in principle unobservable.[78] But perhaps there is no conflict between this point and Sedgwick's; perhaps she merely maintains that while it is impossible to know with certainty that an agent *does* act from duty, it might well be possible to know with certainty, namely in the case of actions which are *obviously* contrary to duty, that an agent does *not* act from duty. In other words, the underlying question here is whether Kant makes the limited claim that the purity of the motive of duty can never be established by observing agents' behaviour, or the more sweeping one that we are never in a position to draw *any* confident conclusions concerning agents' ultimate motives, neither that they did nor that they did not act from the motive of duty.

This much at least is clear: On Kant's account, we can *never* be certain that we have acted from duty, nor that another person has acted from duty. It also appears to be an obvious consequence of his view of the struggle between duty and inclination that we *can* at times know that we ourselves have *not* acted from duty or even from prudential considerations – at the very least when we are aware that our inclinations lead us to act contrary to what we know to be our duty, or even contrary to what is in our own best interest, as in cases of weakness of will. But now the problem which Sedgwick's account raises is the following: Is it ever possible to know, and to know with certainty,

---

[78] Cf. GMS IV:407, where Kant continues: "One does not need to be an enemy of virtue, but only a cold-blooded observer, who does not take the liveliest wish for the good straight away as its reality, in order (especially with advancing years, and a power of judgement grown shrewder through experience and more acute for observation) to become doubtful at certain moments whether any true virtue is ever really to be encountered in the world." See also T&P VIII:284 f.; Rel. VI:62f.

merely on the basis of our observation of someone else's behaviour and our belief that this behaviour constitutes a violation of duty, that *another person* is *not* acting from duty?

The problem with the possibility of such knowledge appears to be that the attempt is made to draw a conclusion regarding an agent's motives solely from his observable behaviour and the fact that this behaviour is (obviously) morally wrong, but *without* any reference to the agent's deliberation, motivations, or decision-making process. As we have seen, such an attempt must fail unless one already assumes that no action which is contrary to duty can be done from duty. If, on the other hand and unlike Sedgwick, one were to base the argument not on the action's wrongness but on the *agent's judgement* that the action is wrong, then it could in fact be made in such a way that it not only avoids the controversial assumption that no action which is contrary to duty can be done from duty, but that it also concludes that an observer can indeed make confident inferences concerning another agent's motives *from the observation of mere behaviour*. Thus, to return to Sedgwick's example, if (P1) the cheater's behaviour indicates that *she herself* considers her cheating to be (obviously) contrary to duty, and if (P2) for an action to be done from duty it is a necessary condition that the agent *believes* (rightly or wrongly) the action to be in accordance with duty, then (C) we can know with certainty – because we can confidently infer this from her behaviour – that she has not acted from duty. In other words, the relevant point for the argument is not that the cheater is in fact acting *contrary to duty* in cheating, nor that her cheating is so *obviously* a violation of her duty; the point is rather that she *knows* that she is cheating. Sedgwick herself is barred from arguing along these lines by her assumption that it is the *wrongness* of the action that is decisive, rather than the *agent's own awareness* of this fact.[79] Her argument that we can sometimes infer the absence of the motive of duty from our observation of another agent's behaviour must fail as long as this inference is made from the *act itself*, i.e. from the fact that an action is *contrary* to duty, and *obviously* so. The argument succeeds only if this conclusion is drawn from the *agent's own recognition* of the wrongness of her actions and the manifestation of this awareness in the agent's behaviour. So much for the internal coherence of Sedgwick's suggested reading.

---

[79] Sedgwick, along with most other commentators, focuses her attention primarily on the fact that the action is in fact – we might say 'objectively' – wrong, and that we, the objective observers in this case, judge it to be so. It is not at all clear that Kant shares this concern with the 'objective' impermissibility of the action or why he should do so. In chapters 8 and 9 below I will argue in some detail that such a reading should be rejected.

2. Kant's exclusion of *all* action contrary to duty

Leaving these issues aside, however, the main interpretative question is whether in the cited passage Kant passes over *all* actions which are recognised as contrary to duty, or only *some*, namely those which are *easily* recognised as such, as Sedgwick would have it. On this question I consider the following third objection to Sedgwick's reading to be conclusive: It is a mistake to read 'already recognised' as '*easily* recognised', thus restricting the class of actions which Kant sets aside at GMS IV:397.11-14 to those which are easily seen to be contrary to duty.

On a purely exegetical note, it appears that if Kant had said that only in *clear* cases of action contrary to duty the question concerning the motive never arises, and that this is why he will pass over them, then this would imply that in cases which are *not so clear* the question *does* arise and consequently he would have to consider these. That is to say, Sedgwick's scenario by implication has Kant indicating to his readers that he will *not* pass over *some* actions which are recognised as contrary to duty, presumably those which are only *with difficulty* recognised as such. We should therefore expect an examination of the possible moral worth of these actions. But in the discussion that follows Kant does not mention *any* cases of actions contrary to duty, including such where it might be *more difficult* to determine their deontic status. Clearly Kant does not entertain such a possibility anywhere in the *Groundwork*. Thus the text is absolutely clear that Kant not only considers himself to be setting aside *all* actions which are recognised as contrary to duty, but also that he does indeed do so, as evidenced by the simple fact that he does not mention any such actions again.

But a more important philosophical point is to be made, especially in the context of the investigation into the motive of duty: Kant is *right* to dismiss *all*, not just some, actions contrary to duty from all further consideration, albeit not for the reasons it is commonly thought. A full argument for this claim will have to wait until chapters 5 and 6, but for the moment I consider my three objections against Sedgwick's interpretation sufficient to draw the following preliminary conclusions regarding the question *which* actions Kant excludes from all further consideration in his short statement at GMS IV:397.11-14: Sedgwick's interpretation mistakes what perhaps might be a subclass of impermissible actions (those which are *obviously* wrong) for *all* impermissible actions and consequently her explanation does not account for Kant's dismissal of "*all* actions that are already recognised as contrary to duty". This leaves me in a position to return to the main issue in the controversy, which is Kant's

apparent omission of actions from duty but not in accordance with duty.[80]

[80] On the interpretation that has Kant draw a distinction between actions which are *easily recognised* as contrary to duty and those which are recognised as such only *with difficulty*, his readers would be perfectly entitled to expect him to follow the same argumentative strategy as the one he adopts in the passages that immediately follow. There he does indeed divide a class of action into two subclasses, of which he then proceeds to set aside one but not the other. Thus, out of the class of all actions which conform with duty, he sets aside the subclass of those "for which [...] human beings have *immediately* no inclination" (GMS IV:397.15 f., emph. mine), but which are done for the sake of mediate inclination – the reason being that in such cases "it is *easy to distinguish* whether the action in conformity with duty is done from duty or from a self-seeking aim" (GMS IV:397.17ff., my emph.; the example he gives for this type of action is the prudent shop-keeper). Such actions then are *obviously* not done from duty. He then goes on to examine in great detail the possible moral worth of action belonging to the other subclass, namely action which "is in conformity with duty and [to which] the subject also besides this has an immediate inclination" (GMS IV:397.20 f., my transl.; the examples he later discusses are the man who refrains from suicide without loving his life, the cold-hearted philanthropist who lacks all sympathy for others, and the gout sufferer who refrains from enjoying what he likes for the sake of his long-term happiness). The reason why he does discuss these latter cases is simply that "[i]t is *much harder to notice* this difference", i.e. it is much more difficult to tell whether the action "is done from duty or from a self-seeking aim" (GMS IV:397.18 f.) when there is an immediate inclination towards the action. That is to say, whether these latter actions are done from duty or not is *not obvious*. My point here is simply that if Kant did indeed make the distinction which Sedgwick attributes to him, then one could reasonably expect him to follow the same argumentative strategy and, having set aside the subclass of actions which *obviously* lack a certain quality, to engage with the other subclass of actions which might or might not have this quality, something which is recognised only *with difficulty*. So the fact that he does not do so in itself makes this reading questionable.

Chapter 7

TWO INTERPRETATIONS OF THE RECOGNITION OF AN ACTION'S WRONGNESS

I now turn to a number of interpretive issues concerning the very *recognition* of the contrariness to the moral law of impermissible actions. The first of these issues gives rise to two further ways in which the seemingly obvious reading of GMS IV:397.11-14 as a dismissal of *all* action contrary to duty might be questioned. Let us remind ourselves of what Kant says:

> I here pass over all actions that are already recognised as contrary to duty, even though they might be useful in this or that intent; for with these there is not even the question whether they might have occurred from duty, since they even conflict with it. (GMS IV:397.11-14)

As I mentioned earlier, it has been a matter of some, although very limited, debate in the literature how exactly this brief statement should be interpreted. Perhaps the most pressing question his readers will direct at Kant's announcement concerns the very notion of recognising one's duty: According to *whose* judgements are impermissible actions impermissible? The agent who deliberates about the permissibility of a certain course of action? We, the readers, who observe the cases Kant describes and who might or might not agree with the agent's reasoning? The ideal agent, that is, the fully rational agent? Furthermore, in what sense are they *already* recognised as such? And why exactly is there *no question* that such actions could be done from duty? On the face of it Kant's statement is tantalisingly ambiguous. Yet these are important questions, both in their own right and because much will depend on them for the possibility of action contrary to duty yet from duty.

In the literature we find two interpretations of the passage. They are distinguished primarily by the way in which they answer these questions. In the last chapter we saw that the interpretation under consideration divided Kant's class of all action contrary to duty into two subclasses according to the ease or difficulty with which actions are recognised as contrary to duty. Kant's statement regarding "all actions that are already recognised as contrary to duty" at GMS IV:397.11-14 was then interpreted as referring to only one of the subclasses, namely 'actions which are *easily* recognised as contrary to duty'. A similar line of argument is taken by some commentators with regard to the

recognition of the contrariness to duty of impermissible actions. Kant's statement is not explicit as to *whose* recognition he is referring to. Does he, when he speaks of "actions that are already recognised as contrary to duty" (ibid.), refer to actions which we, the readers and observers, recognise as contrary to duty (while an agent might mistakenly take the same action to be in accordance with duty), or does he, on the contrary, take the recognition of the action's impermissibility to be the agent's?

In an argument that mirrors the one I discussed in the last chapter, the ambiguity of Kant's expression is taken by some critics to open up the possibility of two distinct ways of interpreting 'recognised as contrary to duty' depending on whose judgement one takes to determine the action's deontic status as 'contrary to duty'. Thus Kant's expression could stand for 'all actions that are already recognised *by the observer* as contrary to duty', or for 'all actions that are already recognised *by the agent* as contrary to duty'. Kant's statement is then interpreted as referring to only one or the other of these two types of action, depending on whether the expression 'recognised as contrary to duty' is taken to stand for the agent's or the observer's moral judgement.[81]

---

[81] As will become apparent in the following two chapters (7 and 8), these two interpretations share one important common feature: they read Kant as working with an (implicit) distinction between acts recognised as contrary to duty by the agent and those recognised as wrong by an observer. They furthermore agree that Kant should be read as referring to, and thereby setting aside, one but not the other type of act, their disagreement concerning merely which of the two types of act is picked out by the statement.

Perhaps the most obvious strategy to criticise both of these readings – which are the most prevalent interpretations amongst Kant scholars – is to attack the idea that Kant makes the relevant distinction in the first place. One way to do this is to suggest an 'impersonal' reading of the passage, according to which Kant excludes all actions which are *generally* recognised as contrary to duty, whichever they in fact are. On this reading, the term 'recognised' (*erkannt werden*) is taken to be factive: thus in the passage Kant passes over all actions which one (i.e. everyone: agent, observer, we all) generally *knows* to be contrary to duty.

However, the strategy which I propose in the following chapters is a different one. For various reasons which will become apparent in due course I favour a qualified subjective (agent-centred) reading of the passage which, like the impersonal reading, offers the advantage of avoiding the distinction between objectively and subjectively wrong acts, but which, unlike the latter, offers several interpretive advantages.

1. Reading 'recognised as contrary to duty' as 'recognised by the observer as wrong' (*The objective reading*)

The standard interpretation is perhaps *prima facie* the most obvious reading of the passage. It is expressed by Allison when he notes that Kant starts the discussion of the notion of acting from duty "with the assumption, itself questionable, that actions performed from duty cannot, *objectively* speaking, be contrary to duty"[82]. Kerstein adopts the same interpretive stance when he remarks that he, "in agreement with other interpreters, take[s] 'already recognized' to mean already recognized by the reader – that is, by 'objective' observers – to be contrary to duty"[83]. Thus the standard reading takes actions which are 'recognised as contrary to duty' to be actions which – because the judgement that they are wrong is made by the observer – are '*de facto* contrary to duty' or '*objectively* wrong'.[84] According to this reading, then, Kant thinks that no action which in fact violates duty can be done from duty, and based on this premise he sets them aside, since they could not possibly be illuminating in the search for actions that are done from duty and hence have moral worth. Kant's somewhat puzzling claim that the question concerning the motive does not even arise when an action is already recognised as contrary to duty is explained with reference to the action's *objective contrariness* to the moral law: actions which are in fact contrary to the requirements of morality just cannot be done from the motive of duty.

Despite the *prima facie* appeal of the standard reading, commentators – including those who think we cannot avoid the conclusion that this is indeed Kant's view – are

---

[82] Allison, op. cit., p. 109.

[83] Kerstein, op. cit., p. 97. Other authors whose treatments of the possibility of action from duty but not in accordance with duty suggest or imply that they favour – or at least that they consider reasonable – such a reading include Guyer, op. cit. pp. 42f.; Marcia Baron, 'Acting from Duty (GMS, 397-401)', in: Christoph Horn & D. Schönecker (eds.): *Groundwork for the Metaphysics of Morals* (Berlin: de Gruyter, 2006), p. 73; George C. Kerner: *Three Philosophical Moralists: Mill, Kant, Sartre* (Oxford: Clarendon Press, 1990), pp. 84ff.; Dieter Schönecker & A. Wood: *Kant's Grundlegung zur Metaphysik der Sitten: Ein Einführender Kommentar* (Paderborn: Schöningh, 2007, 3rd ed.), p. 64; Barbara Herman, 'On the Value of Acting from the Motive of Duty', *The Philosophical Review* 90 (1981), p. 376; Richard Henson, 'What Kant Might Have Said: Moral Worth and the Overdetermination of Dutiful Action', *Philosophical Review* 88, (1979), p. 40; H.J. Paton, *The Categorical Imperative: A Study in Kant's Moral Philosophy*, (London: Hutchinson, 1947), p. 40.

[84] Henceforth I will refer to this interpretation of Kant's expression 'recognised as contrary to duty' as 'the objective reading' or 'the objective interpretation'.

all too aware that this interpretation has Kant relying either on the problematic *assumption* that no action which in fact violates duty can be done from duty, or on the equally questionable *definition* of the term 'duty' to the same effect. Most critics nevertheless think that this must indeed be Kant's view. We have already seen (in chapter 5.2. above) that – irrespective of whether it is Kant's view or not – the reliance on such an assumption or stipulative definition is unsatisfactory. To be sure, there are some interpreters who think it is possible to avoid attributing this view to Kant.[85]

2. Reading 'recognised as contrary to duty' as 'recognised by the agent as wrong' (The subjective reading)

The case for the alternative interpretation of the passage is made most explicitly by Curzer, who thinks it unnecessary and unsophisticated to commit Kant "to the claim that it is impossible to act from duty without acting in accord with duty"[86]. His main charge against the standard reading is that it "assumes that Kant ignores the common occurrence of morally motivated but misguided acts"[87]. He therefore rejects it in favour of a "more charitable interpretation of the passage [which] would assume Kant to be aware of [the] fact [that] [p]eople often perform morally wrong acts because they mistakenly believe them to be morally required"[88]. According to this reading,

> we need not take [Kant] to be omitting all acts recognized *by the reader* to be contrary to duty. Luckily, there is another way to read the quoted passage. I take Kant to be omitting all acts recognized *by the agent* to be contrary to duty. The question does not arise as to whether these acts might be done from duty, since *the agent believes* they conflict with duty. On this reading, Kant is not committed to the claim that it is impossible to act from duty without acting in accord with duty. Kant leaves open the possibility that [...] an agent could perform a morally wrong act from duty if the agent thinks the act is morally right. Acts which are not in accord with duty may be done not only from inclination, but also from duty. This interpretation attributes to Kant a more sophisticated view, and allows [for] morally motivated, immoral acts [...].[89]

---

[85] For a detailed discussion of the merits of the objective reading see chapters 8 and 9 below, where I will argue that it should be rejected for various reasons, including the fact that it lacks explanatory force.

[86] Howard Curzer, 'From Duty, Moral Worth, Good Will', *Dialogue* 36 (1997), p. 291.

[87] Ibid., p. 290.

[88] Ibid.

[89] Ibid. Cf. also Köhl, who cites our *Groundwork* passage (GMS IV:397.11-14) and comments: "This

The suggestion then is to read 'recognised as contrary to duty' as 'recognised *by the agent* as wrong'.[90] To put this interpretation and the main motivation which underpins it into the context of the terms introduced in chapter 1: Driven by a concern with "the common occurrence of morally motivated but misguided acts"[91] – the existence of which Kant ignores or overlooks according to the standard interpretation – proponents of the subjective reading want to interpret Kant as allowing for the possibility of such acts ('actions from mistaken moral judgements', as I labelled them earlier), and therefore find it necessary to make room in Kant's account for the possibility of performing a morally wrong act from duty. I suggested earlier that this line of thought rests on the Logical Compatibles View of Kant's twofold distinction of duty, as well as on the Identity Thesis which claims that 'morally motivated but misguided acts' *are*, in Kantian terms, 'actions from duty but contrary to it'. I will address both of these theses in some detail in chapters 8 and 9 below, where I will argue that the suggestion to read 'contrary to duty' as '*subjectively* wrong', while it is an improvement on the standard interpretation, is problematic insofar as it is motivated by a concern with action from mistaken moral judgement and dependent on these two questionable assumptions.

---

thesis is only true if we read it in such a way that 'all actions which are already recognised *by the agent* [...] as contrary to duty' are excluded from the start as candidates for morally good actions. It is wrong, in my opinion, if it requires that morally good actions cannot *in fact* be contrary to duty. Not the *objective*, but only the *subjective* accordance with duty of an action is a necessary condition of its morality [...]." See Harald Köhl, *Kants Gesinnungsethik* (Berlin: de Gruyter, 1990), p. 74 (my transl.). He proceeds to cite Hardwig's paper (op. cit.) in support of his own criticism of Henson's claim (cf. op. cit., p. 40) that the question whether an action has moral worth can be asked only about an act which is in fact, i.e. objectively, required by duty: Köhl thinks that "Kant himself was possibly of the opinion that an action 'from duty' must be *objectively* in accordance with duty, that is, apart from the agent's opinion" (Köhl, op. cit., p. 73, my transl.), but intimates – and this is the relevant point in the present context – that if Kant did hold this view he would be mistaken, since "there is [such a thing as] an erroneous moral judgement together with the motive to do that which is erroneously thought to be in conformity with the moral law because it conforms with the moral law, i.e. from duty". (ibid., my transl.) Unfortunately, Köhl does not consider the issue in great detail and the claim is not supported by any argument.

[90] Henceforth I shall refer to this interpretation of Kant's expression 'recognised as contrary to duty' as 'the subjective reading'.

[91] Curzer, op.cit., p. 290.

3. The respective weaknesses of these interpretations and a problem with both readings

For the moment we should simply note as a preliminary result that both interpretations have obvious weaknesses: The suggestion to read 'recognised as contrary to duty' as '*objectively* wrong' is dubious – as I have already argued (see chapter 7.1 above) – because it has Kant relying either on a stipulative *definition* of the term 'duty' or on a questionable *assumption* (in both cases to the effect that no action which violates duty can be done from duty). On the other hand, the suggestion to read 'recognised as contrary to duty' as '*subjectively* wrong' in an attempt to make room for the possibility of action from mistaken moral judgement is problematic to the extent that it relies on the Identity Thesis and the Logical Compatibles View, which are equally questionable assumptions, as I will proceed to argue in chapters 8 and 9.

In addition, however, there is one textual consideration which applies equally to both interpretations and which would suggest that neither is entirely satisfactory, quite apart from any other weaknesses. As we saw earlier, in his search for the motive of duty Kant employs a 'method of isolation'[92] which proceeds in the first instance by identifying various classes or types of action. Some types of action are dismissed ('set aside' or 'passed over') more or less immediately as not relevant to this search, while others are discussed, with examples, in considerable detail. In following such a strategy, Kant – unless, of course, we take him to have forgotten to consider certain types of action[93] – is *including* in the search all those classes of action which he does not explicitly *exclude*. But as I pointed out earlier in relation to Sedgwick's suggestion that Kant only sets aside *some* actions which are recognised as contrary to duty, namely those which are *easily* recognised as such: not only does Kant not make such a distinction, but in the discussion that follows Kant's statement, he never considers or even mentions *any* actions which fail to conform to the moral law. That is to say, given his strategy at that stage of the discussion, we must take him to have excluded all such actions from further consideration.

But bearing this last point in mind in relation to our two interpretations of Kant's expression, a similar argument can be made in the present context: On the suggestion that 'recognised as contrary to duty' should be read as '*objectively* wrong', Kant would

---

[92] Paton's term. See op. cit., p. 47.

[93] A possibility which I will consider shortly.

intimate that in his search for actions which might be done from the motive of duty he will pass over one *type* of act: He will set aside all those actions which are recognised *by the observer* to be contrary to duty (while the agent might erroneously believe them to be required), that is, all actions which are *objectively* wrong (while they are perhaps *subjectively believed to be* right by the agent). But by indicating that he will only pass over this one type of action – albeit *all* the members of that class – Kant would be implying that he will *not* pass over some other type of action, presumably all those the contrariness of which is already recognised by the agent, i.e. those which are *subjectively* wrong. In doing so he would imply that he is *including* this class of action in the following investigation into the motive of duty and thus raise an expectation in his readers that a discussion of their possible moral worth is imminent. However, as we know, following his brief statement at GMS IV:397.11-14, Kant does not mention *any* actions contrary to duty again.

Of course the same argument applies, *mutatis mutandis*, to the alternative suggestion that 'recognised as contrary to duty' should be read as '*subjectively* wrong'. Here the actions which Kant passes over would again be limited to one type of act, in this case to those which *the agent* recognises as contrary to duty, and Kant would thereby be indicating to his readers that he will *not* pass over, that is to say, that he intends to *include* in the search for the moral motive, the class of actions which are *objectively* wrong. However, as we know he does *not* include them. That Curzer, for one, clearly thinks along these lines is obvious when he says that "we need not take [Kant] to be omitting all acts recognized *by the reader* to be contrary to duty"[94], rather, we should "take Kant to be omitting all acts recognized *by the agent* to be contrary to duty"[95]. Curzer thinks that Kant does *not* pass over actions which are objectively wrong yet done from duty, and he takes this to mean that Kant "leaves open the possibility"[96] of such acts. Likewise, Baron maintains that Kant "does not claim that honest mistakes are impossible, and it would be uncharitable to attribute this view to him if we do not have to"[97]. She, likewise, suggests that reading Kant's 'already recognised as contrary to duty' as 'recognised *by the agent* to be contrary to duty' means that "there is no need to read Kant as ruling out this possibility [i.e. the possibility of action from duty but not

---

[94] Curzer, op. cit., p. 291.

[95] Ibid.

[96] Ibid.

[97] See her 'Acting from Duty (GMS, 397-401)', in: Christoph Horn & D. Schönecker (eds.): *Groundwork for the Metaphysics of Morals* (Berlin: de Gruyter, 2006), p. 73.

in accordance with it]"[98]. But to say that Kant does *not* rule out objectively wrong acts, but *does* exclude subjectively wrong ones, is to say that he *includes* the former but *not* the latter in the search for morally motivated actions. But this is patently false since, as we have seen, he does not include either.

The context of Kant's statement is important: The classes of action which Kant does *not* set aside does not include, as Curzer and Baron appear to imply, types of action which Kant – we do not quite know why – decides to ignore in the remainder of the *Groundwork* despite having just identified them as relevant to the investigation, thus 'leaving open the possibility' that they might nevertheless be bearers of moral worth – a possibility which perhaps might be investigated another day. Rather, the actions which Kant does *not* pass over are precisely the types of action which his method of isolation singles out as the primary candidates for moral worth and which therefore are at the centre of the discussion that follows. Neither objectively wrong acts nor subjectively wrong acts are such candidates.

There might be a way to defend the alternative, subjective reading against my criticism. Although they certainly do not make it clear that this is indeed their view, the following might be what adherents of the subjective reading have in mind when they suggest that Kant's strategy of setting aside some types of action and considering others 'leaves room for the possibility' of action from duty but not in accordance with duty. I have taken Kant's investigation into all possibly moral worth-bearing actions to be comprehensive, at least in the sense that – whatever other faults it might have – Kant has not simply overlooked an entire class of action which might have moral worth. That is to say, I have taken him to *implicitly include* in the search all types which he does not *explicitly exclude*, thus ruling out the possibility that there are yet more types which he ignores entirely. According to this interpretation the class of types of action which he considers is comprised of those types he omits (sets aside) and those he investigates, and together they constitute all the types of action one could possibly expect Kant to consider in the context, that is to say, they exhaust all the relevant possibilities. This implies, crucially, that there are no types of action which Kant could possibly have omitted (overlooked). But, of course, one might reject this premise and hold that there is always the possibility that, in addition to the classes Kant includes and the ones he rules out, there are others which he just does not mention. On this view, when Kant *excludes* from the investigation all actions which are recognised by the *agent* as contrary to duty, he does not thereby *implicitly include* those actions which are

[98] Ibid.

recognised by the *observer* as contrary to duty. Rather, he does not even mention them, either because he decides to ignore them entirely, or because he is unaware of their existence, or because he simply forgets to include them in the discussion. Kant's failure either to omit (set aside) explicitly or to investigate action from duty but not in accordance with duty is therefore indicative of his omission (oversight) of such action, which thus 'leaves open the possibility', which *we* might now investigate, that such action has moral worth.

A full reply to this suggestion will have to wait until chapters 8 and 9, where I outline my interpretation of the passage. The short answer is that I will deny the claim, made by the adherents of the subjective reading, that Kant fails to set aside action from duty but not in accordance with duty. I suggest that if we wish to be charitable to Kant and avoid attributing views to him which we do not have to, then the way to do this is not by claiming that his ignorance of and consequent failure to exclude explicitly action from duty but not in accordance with duty 'leaves open the possibility' of such action. Rather, I think it is preferable to show that Kant, for very good reasons, does not identify a class of 'action from duty but not in accordance with duty' *as such*, which would explain why he cannot explicitly set it aside *as such*. Since, as we know, Kant neither draws the relevant distinction between subjectively wrong and objectively wrong actions in the first place, nor returns to either type of action in the following discussion of the motive of duty, and, finally, given the argumentative strategy he adheres to in employing his method of isolation, the only sensible conclusion to draw from this is that we must take him to have excluded *all* such actions from further consideration.

We are now in a position to note a number of intermediate results by way of a summary of the last two chapters: First, I argued (in chapter 6.2.) that when Kant rules out "all actions that are already recognised as contrary to duty" at GMS IV:397.11-14, the text is absolutely clear both that he *takes himself* to be setting aside *all* impermissible actions, and that he *does* indeed do so. Against the suggestion that 'already' should be read as 'easily' and that Kant only rules out those wrong actions which are *obviously* so, I have maintained that he does indeed rule out *all*, not merely *some* such actions.
Second, I have attempted to show in the present chapter that it would be a mistake to read into Kant's statement a distinction, even an implicit one, between 'actions recognised by *the observer* as contrary to duty' on the one hand, and 'actions recognised by *the agent* as contrary to duty' on the other, and then to read him as setting aside only one but not the other, since he mentions *no* actions recognised as contrary to duty

again and so clearly passes over *all* such actions. Thus both the standard 'objective' reading of 'actions recognised as contrary to duty' and the alternative 'subjective' reading, *insofar as they attribute such a distinction to Kant*, are, at the very least, misleading. It should become apparent shortly why it is crucial to recognise that Kant does not even draw this distinction. For the moment it suffices to conclude that both these results confirm that in his quick dismissal of action contrary to duty Kant does indeed dismiss *all* such action.

Chapter 8
REJECTION OF THE LOGICAL COMPATIBLES VIEW

Where does all this leave Kant's statement at GMS IV:397.11-14? In the last chapter I argued that in addition to the problem which both the subjective and the objective interpretations share, i.e. their reliance on a distinction which Kant does not make as such, as well as the exegetical problems arising from it, they each have their own weaknesses which make them unattractive: The 'objective' reading relies on the assumption (or the definition of the concept of 'duty' to the effect) that moral worth presupposes rightness, while the 'subjective' reading rests on the assumption of the Logical Compatibles View and the Identity Thesis.

In the next two chapters I will argue that the objective reading does not stand up to scrutiny and hence should be abandoned, while the subjective reading – or rather, a heavily qualified version of it – should be accepted, but only if it can meet certain conditions: In particular, a coherent defence of the subjective reading must either provide a cogent argument for, rather than assume, the Logical Compatibles View and the Identity Thesis on which it rests, or, alternatively, it must show why they should be rejected – and how they can be rejected. Secondly, it must not attribute to Kant a distinction between the agent's and the observer's standpoint which he does not make, since, as I have outlined in the last chapter, this would imply that he sets aside only some, not all, actions contrary to duty. And finally it must, in the first place, be preferable as an interpretation of the text, that is, provide a clear explanatory advantage over the alternative, objective reading.

1. Three textual considerations in support of a qualified subjective reading

Beginning with the latter, we must now turn our attention to several features of Kant's statement at GMS IV:397, 11-14 which I have until now ignored. Kant's words, again, are:

> I here pass over all actions that are already recognised as contrary to duty, even though they might be useful in this or that intent; for with these there is not even the question whether they might have occurred from duty, since they even conflict with it. (GMS IV:397.11-14)

First consideration: In all the interpretative issues which I have discussed so far, the main problem was clarifying exactly *which* actions Kant omits from further consideration in the investigation into the motive of duty. But there is also the question – perhaps the more interesting and the more important question – *why* he does so. The second half of Kant's statement appears to provide an answer: the actions which he sets aside can be ignored for the purposes of the investigation, we are told, because "with these there is not even the question whether they might have occurred from duty, since they even conflict with it". (GMS IV:397.13f.) The aim of this second half of the statement is obviously to clarify and justify the move made in the fist half, that is, the exclusion from the investigation of "all actions that are already recognised as contrary to duty" (ibid.). In reply to the question *why* these actions are omitted from the investigation, the reader is now informed that with such actions there is no question that "they might have occurred from duty" (ibid.). Clearly, if it is indeed out of the question, i.e. in some sense impossible, that such actions could be done from the motive of duty, then this would sufficiently explain and justify ruling them out. So Kant's argument hinges on his ability to show that the question concerning the motive does indeed not even arise when it comes to action which is already recognised as contrary to duty. However, once again one might ask *why* there is no "question whether they might have occurred from duty" (ibid.), and to this query Kant replies that the question whether such actions might have been done from duty does not arise "since they even conflict with it" (ibid.). So as far as the argumentative structure of the passage is concerned, and with regard to "all actions which are already recognised as contrary to duty" (ibid.), Kant takes himself to be arguing that their being in conflict with duty explains why the question whether they might be done from duty cannot even arise, and the absence of this question in turn explains why he intends to pass over all such actions.

Whether, all things considered, Kant's reasoning is sound is, of course, an important question (to which I will return in chapter 9), but for the moment I want to pursue a different line of inquiry. Since my first aim is to explore whether the subjective or the objective reading is preferable as an interpretation of the text, my first question will be whether, on the assumption that Kant's reasoning *is* sound, either of the two readings can account for and explain the stringency of his argument. On the basis of which presuppositions, if any, would Kant's apparently faulty reasoning be consistent? (Or, to misuse familiar Kantian terms: What are the conditions for the possibility of the validity of his argument?) In particular, the crucial point is whether

on either reading Kant's central claim can be made coherent or can even be justified, that is, the claim that the question concerning the motive does not even arise in relation to actions which are already recognised as contrary to duty, and that it cannot arise precisely for this reason: *because* "they even conflict with it" (ibid.).

Applying the objective reading to Kant's argument and paraphrasing Kant's words somewhat, we get the following: *Regarding actions which are in fact (i.e. objectively) wrong, the question whether they are done from duty never arises, because they in fact (objectively) conflict with duty.* Or, simplified further: *Actions which are wrong cannot be done from duty because they* are *wrong.*

The subjective reading yields: *Regarding actions which the agent believes to be wrong, the question whether they are done from duty never arises, because the agent believes that they conflict with duty.* Or, again: *Actions which the agent believes to be wrong cannot be done from duty because they are* believed *to be wrong.*

Now it seems clear that the objective reading does not in any sense clarify why or how the fact that an action is in fact impermissible should explain that it cannot be done from duty, while on the subjective reading it is already easier to see how the fact that an action is *thought to be* impermissible might explain why it cannot be done from duty. The subjective reading appears to be the better rendition of the argument by virtue of the fact that it offers an account of the force of the supposed justification contained in it. This is even more conspicuous when we remember that, as I pointed out before, according to Kant no merely permissible action can be done from duty, and that to perform an obligatory action from duty is to act *from the notion that the action is morally required.* Thus adjusting both readings accordingly we get:

Objective reading: *Actions which are wrong cannot be done from the notion that they are required because they* are *wrong.*

Subjective reading: *Actions which the agent believes to be wrong cannot be done from the notion that they are required because they are* believed *to be wrong.*

It is now evident that on the objective reading favoured by the majority of scholars the putative justification for the absence of the question concerning the motive of duty ('because they *are* wrong') is neither a justification nor an explanation; rather, it simply restates the main assumption in this context, namely that no action from duty can be contrary to it, and it is therefore superfluous to the argument. But without a justification of the claim that the question concerning the motive cannot even arise we are left only with the assertion that wrong actions cannot be done from duty, which is precisely the assumption in question, and thus the entire argument begs the question. It still remains a mystery what it is about actions which are objectively contrary to duty

that means that the question whether they could have been done from duty cannot arise.

By contrast, on the subjective reading the argumentative force of Kant's statement is laid bare: If an agent believes an action to be contrary to duty – that is, if the agent not only does not believe the action to be required, but in fact believes it to be impermissible – and if to act from duty is to act from the notion that the action is required, then it is obviously impossible for the agent to perform the action from duty. Hence the question concerning the motive of duty cannot even arise. On this account, the justification given ('because they are believed to be wrong') does indeed justify ignoring the question concerning the motive because believing an action to be impermissible rules out the possibility that one performs this action from the notion that it is required. Thus the subjective reading offers a perfectly straightforward and coherent explanation of the way in which an agent's belief that an action conflicts with duty both precludes there being any question of it being done from the motive of duty, and justifies passing over all such action.

There are in addition some further and unrelated textual clues which could be taken as supporting a subjective reading, according to which Kant passes over all actions believed *by the agent* to be contrary to duty.

Second consideration: I will begin with a less significant point, which must be mentioned here for the sake of completeness, if for no other reason, since it concerns the one remaining part of Kant's statement at GMS IV:397.11-14 which I have not yet considered. When Kant begins his discussion of the concept of duty by announcing that he will "pass over all actions that are already recognised as contrary to duty, *even though they might be useful in this or that intent*" (ibid., emph. added), we might take him to be making the concession that while he will pass over these actions on account of their contrariness to the moral law, he does so despite recognising their usefulness in other respects. In other words, we would take Kant to apologise to the reader for the fact that he is now concerned exclusively with moral questions and will disregard actions which are useful in many non-moral respects, such as from prudential considerations. On this interpretation of the text the admission that some immoral actions are in fact useful in non-moral respects is taken to refer to Kant's passing over these actions. Thus Kant begins by announcing that he will ignore immoral actions *despite* acknowledging their usefulness.[99] But this might seem odd in the context: that

[99] Korsgaard, for one, reads the passage (GMS IV:397.11-14) along these lines: "I take Kant to be saying

morality is his exclusive concern should by now be abundantly clear to even the last of his readers and it would seem strange that Kant should feel the need to reiterate this point at this stage in the text.

There is, however, a different way to read the text by taking "even though they might be useful in this or that intent" (ibid.) to refer not to Kant's passing over these actions, but to their being "recognised as contrary to duty" (ibid.). On this reading Kant is saying, firstly, that there are some actions which, "*even though* they might be useful in this or that intent" (ibid., emph. added), are nevertheless "*recognised* as contrary to duty" (ibid., emph. added), and secondly, that he will pass over *all* such actions. I think it is clearly preferable to read the text as stating that whoever recognises an action as impermissible might do so despite considering the action to be useful in other respects, rather than reading it as making the obvious point that Kant, in the context of analysing the concept of duty, is justified in ignoring immoral actions but feels the need to acknowledge their *usefulness* for other purposes.

Now the one remaining question is how the preferable interpretation might also serve as supporting evidence for an agent-centred, subjective reading of the entire passage at GMS IV:397.11-14, which I believe it can. If we read Kant's statement, along the lines which I have suggested, as the claim that some "actions [...] are already *recognised* as contrary to duty, *even though they might be useful* in this or that intent" (ibid., emph. added), then the point is that an action's usefulness for other purposes, inasmuch as it promises an advantage, can be an obstacle to the *recognition* of its moral impermissibility. It can be an obstacle, albeit one which is by no means insurmountable, because immoral actions are said to be recognised as such *despite* their usefulness for other purposes, so in these cases at least the obstacle has been overcome. The question now is whether it is more likely to be an obstacle for an observer or for an agent. My suggestion is that Kant has the latter in mind: A dispassionate observer who analyses an action in terms of both its usefulness and its morality would, *qua* observer, not be any less inclined or less likely to recognise an action's immorality on account of its usefulness. The moral deliberator and agent, on the other hand, *qua* agent engaged in evaluating possible courses of action open to *him*, clearly faces the danger that his awareness of an action's usefulness and the advantages it offers for his own prudential purposes might enter into his reasoning

---

that any value these actions may have must come from their utility." See her 'From Duty and for the Sake of the Noble: Kant and Aristotle on Morally Good Action', in: Stephen Engstrom & J. Whiting (eds.), *Aristotle, Kant, and the Stoics – Rethinking Happiness and Duty* (Cambridge: Cambridge University Press, 1996, pp. 205f.

about its moral permissibility. Thus this consideration is further evidence that when Kant says that "actions [...] are already *recognised as contrary to duty*, even though they might be *useful* in this or that intent" (ibid., emph. added), he attributes the recognition of the action's contrariness to duty (as well as the recognition of its usefulness, as we have seen) to the agent.

Third consideration: I now turn to another textual clue which links the two classes of actions which Kant takes himself to be excluding from the investigation. I will on this occasion state the case for the subjective reading by arguing that a particular feature of the text can be accounted for only on that reading. Immediately following the statement at GMS IV:397.11-14, Kant goes on by saying that he will "also set aside those actions which are *really* [wirklich] in accordance with duty, to which, however, people have immediately no inclination [...]." (GMS IV:397.14-16, my transl., emph. added) If, on the subjective reading, one were to read Kant's 'recognised as' in the previous sentence as 'recognised (perhaps mistakenly) by the agent as', one might be inclined to juxtapose this with his use of 'really' here. Both of these terms would then qualify the deontic status of the types of acts in question. In the first case, 'recognised (by the agent) as' qualifies the action's contrariness to the moral law, while in the second, 'really' qualifies its accordance with it. Kant then would be contrasting two kinds of actions which, for *equally obvious*[100], but otherwise *very different* reasons, can both be safely excluded from the search for the moral motive.

He first identifies actions which can be set aside merely by virtue of their *perceived* deontic status, i.e. because they are *believed* by the agent to be contrary to duty. In these cases knowing the actions' actual rightness or wrongness is unnecessary. It is *obvious* that they cannot be done from duty – whether they are in fact contrary to it or not – because the agent already believes them to be contrary to it. (Even an agent who is mistaken in this belief and, performing the action despite his belief that it is impermissible, unknowingly acts in accordance with duty, could not act from duty, because to act from duty is to act from the notion that the action is required, and *ex hypothesi* the agent here acts rightly *despite* his belief that the action is impermissible.) Thus Kant first sets aside a class of action simply *because* the agent, rightly or wrongly, thinks the action is impermissible. This type of action is ruled out merely on the basis that its *perceived* deontic status is contrariness to duty, and Kant takes this to imply, for the reasons outlined, that it *obviously* cannot be done from duty.

---

[100] Cf. Timmermann, op. cit., p. 30 on this point.

Secondly, Kant argues, we can ignore in the investigation into the motive of duty even some actions the *real* deontic status of which is 'in accordance with duty'. These actions can be set aside *despite* the fact that they *really are* in accordance with duty. Of course, actions which are in accordance with duty are the primary target of the following investigation into the motive of duty, so it might come as a surprise that he sets some of these aside. But Kant here identifies a subclass of these actions, namely those towards which people have no immediate inclination, but which they nevertheless perform from an ulterior motive, typically self-interest. Whether or not such actions are in addition believed to be in accordance with duty by the agent is irrelevant, because, as the example of the prudent merchant shows, Kant thinks that they are obviously done from self-interest and thus not from duty. So Kant thinks he can exclude such actions because they, too, are *obviously* irrelevant, but for a very different reason: "For there it is *easy to distinguish* whether the action in conformity with duty is done from duty or from a self-seeking purpose." (GMS IV:397.17-19, my transl. & emph.) Thus the second type of action which Kant rules out from the investigation are actions which, despite the fact that they *really are* in accordance with duty, are easily seen to be done from self-interest, i.e. are also *obviously* not done from duty.[101]

Of course, one can juxtapose 'really' with 'recognised as' only if one reads the latter, with the subjective interpretation of the passage, as 'recognised by the agent as'. On an objective reading of the statement there will be no such contrast. In that case, however, Kant's inclusion of the adverb 'würklich' would be superfluous and somewhat odd. Thus one might argue conversely that, inasmuch as one takes Kant's inclusion of 'really' to be meaningful or even of some significance, its presence suggests a juxtaposition with 'recognised as' and thereby supports the subjective reading according to which the recognition of an action's rightness or wrongness is the agent's.

I must reiterate that my aim is a limited one here: Kant thinks of two types of agents as *obviously* lacking a good will: those who already believe their action to be

---

[101] Again, this reading appears to account for certain features of the text in a way in which standard interpretations do not. Thus, Korsgaard thinks that Kant passes over all actions recognised as contrary to duty "for the same reason" as he sets aside "actions that are in accordance with duty but that are not chosen for their own sakes" (the prudent merchant): In both cases "any value these actions may have must come from their utility", while "a morally good action", by contrast, "must be chosen for its own sake". (Korsgaard, op.cit., pp. 205f.) This reading is not incompatible with what I have said, but it does not explicate certain nuances of the text.

contrary to duty and those who have a strong ulterior motive of self-interest to perform an action which also happens to coincide with duty. Whether he is right in thinking that these types of agent lack a good will, or even that they obviously lack a good will, is not the issue here. My point is merely that in describing these types of agent, Kant contrasts two very different reasons for ruling out the two classes of actions, or two types of agent, and the juxtaposition of these different reasons requires that 'already recognised as contrary to duty' be read as 'recognised *by the agent* as contrary to duty'.[102]

To conclude: I have argued that there are at least three specific textual considerations which each support a subjective reading of the passage and thus serve to question the traditional objective reading. Furthermore, the first and most important of the three points suggests that there is in fact a tension between a subjective reading and the notion that Kant's account can accommodate the possibility of action from duty but not in accordance with duty, a possibility which according to proponents of the most common version of the subjective reading is first opened up by such a reading. It will be the business of the remainder of this chapter to determine exactly the grounds of this conflict and to propose a solution to it.

---

[102] To my knowledge the specific textual points I raise in the second and third considerations above, i.e. concerning the juxtaposition of 'really' with 'recognised as' and regarding the significance of the clause 'even though they might be useful in this or that intent', have not been picked up in the literature. They are, admittedly, relatively minor considerations. They do not, nor are they intended to, support my larger claim that Kant must be taken as arguing from a first-person perspective, or my rejection of the two theses. This larger claim is supported by the first of the three textual considerations above and most importantly by the arguments adduced in chapters 8.2. and 8.3., as well as in chapter 9, below. I do think, however, that the second and third considerations constitute further objections to the objective reading. I see their role as that of auxiliary arguments within the overall context of my defence of a qualified subjective reading. Ultimately nothing hinges on them, but even from a purely exegetical point of view the fact that they bring out another dimension of Kant's statement at GMS IV:397.11-19 might be thought to be of interest.

2. Kant's adoption of the agent's standpoint

Having tested the two readings as to their ability to explain adequately Kant's reasoning, and having found that various textual considerations suggest that the subjective reading offers a clear explanatory advantage over the alternative, we must now consider the second of the three difficulties which a coherent defence of the subjective reading must address.[103] How is it possible to read Kant as arguing that the agent's belief that an action conflicts with duty prevents the motivational question becoming an issue and thus justifies passing over all such action, without thereby attributing to Kant a distinction between the agent's and the observer's standpoint, that is, between an action's objective and its subjective wrongness – a distinction which, as we saw earlier, he simply does not make?

My suggestion is that the only alternative to ascribing a distinction between viewpoints to him is to read Kant as simply *adopting* one of the viewpoints, namely the stance of the moral deliberator, decision-maker, and agent. At least in the relevant passage in the *Groundwork* he considers the issues from the first-person perspective of the agent. It is important that Kant does not – contrary to what the adherents of the view that he differentiates between the observer's and the agent's perspective suggest – *distinguish* this viewpoint from others, such as the observer's. He does not turn the possibility of adopting one standpoint or another into an issue or a subject of the discussion. Rather, his argument is simply presented from this point of view, making it a general underlying feature of his approach to the issues.

However, one might wonder whether it is not contradictory to read Kant as arguing from the agent's first-person perspective (thus setting aside only actions which the agent recognises as wrong), and simultaneously to claim that he sets aside *all* impermissible actions. In other words, how can the subjective reading be reconciled with Kant's quick dismissal of *all* action contrary to duty? If, as I have argued, his argument *presupposes* the first-person perspective, then the consequences for actions recognised as wrong from either of the two perspectives under consideration are twofold: Actions which the agent already recognises as contrary to duty (subjectively wrong acts) are ruled out of the investigation into the motive of duty by a simple and *explicit* line of reasoning: it is impossible for an agent to act from duty, that is, from the notion that the action is required, when he has recognised the action as wrong. On the

[103] See the three conditions outlined in the introduction to chapter 8 above.

other hand, actions which the observer, but not the agent, recognises as contrary to duty (objectively wrong acts) are indeed never mentioned by Kant but nevertheless excluded *implicitly*, that is, by implication of the viewpoint Kant adopts in his discussion. The very possibility of recognising an action as objectively wrong, and of contrasting it as such to a merely subjectively wrong act, presupposes the adoption of a third-person perspective. But since Kant adopts a first-person perspective, the possibility of such action is precluded.

It should be noted that the reading proposed here differs significantly from the subjective reading dominant in the literature as proposed by Baron, Curzer, et al. In fact, it seems that the very concern which gave rise to their subjective reading in the first place, namely the possibility of morally motivated but objectively wrong actions, can be shown to be precluded by such a reading, properly qualified: Because the agent's first-person perspective is an underlying feature of Kant's argument, rather than a topic of his discussion, the subjective reading, rightly understood, far from leaving open the possibility of morally motivated but objectively wrong actions, actually rules out such a possibility. That is, once the full implications of the subjective reading are spelled out, actions from duty but not in accordance with duty are seen to be impossible.

3. The order of deliberation

I have so far in this chapter defended the view that the textual evidence suggests an agent-centred reading of Kant's statement at GMS IV:397,11-14, and more specifically an interpretation which explains his passing over *all* actions recognised as contrary to duty by attributing to him the adoption of the standpoint of the moral deliberator, decision-maker, and agent. The question now is how the kind of qualified subjective reading which I propose relates to the Logical Compatibles View and the Identity Thesis. In my outline of the various ways in which the subjective reading needs to be qualified, I suggested that a coherent defence of it must also do either of two things: it must either justify the assumption of the Logical Compatibles View and the Identity Thesis on which it rests, or be prepared to reject them, in which case it must show why and how they can be rejected. It will come as no surprise that I favour the latter approach.

We have already seen that, contrary to what the proponents of the unqualified subjective reading suggest, from the viewpoint of the agent actions from duty but not in accordance with duty would in fact seem to be impossible. In order to see why

exactly this is so, we must now consider one further item of textual evidence which, I believe, lends further support to the agent-centred reading. Although neither the Logical Compatibles View and the Identity Thesis nor their respective criticisms can be separated neatly, my primary concern here is with the rejection of the Logical Compatibles View, while the Identity Thesis will be the subject of the next chapter.

There is one final feature of Kant's statement which we have not yet fully explained, namely his claim that the question concerning the motive of duty does not even arise in relation to all actions which are *already* recognised as contrary to duty by the agent. It is clear that 'already' in some sense qualifies the recognition of the impermissibility of the action. In chapter 6 above I rejected the suggestion that we should read 'already' as 'easily', which would limit the class of actions which Kant sets aside to those which are *obviously* impermissible. As I outlined in the introduction, according to Kant there are two questions which we may ask in relation to every action, one concerning its rightness and another concerning its moral worth. In the statement at GMS IV:397,11-14, Kant explicitly mentions the latter of these as "the question whether [the action] might have occurred *from duty*" (ibid., my emph.), while he refers to one possible answer to the former question, whether or not the action is in accordance with duty, when he speaks of the class of "all actions that are already recognised as contrary to duty" (ibid.). Now Kant's claim is that the question concerning the motive *does not arise* when an action is "*already* recognised as contrary to duty" (ibid.). In my defence of a qualified subjective reading above, I explained how and why an agent's recognition of an action's impermissibility rules out his performing the action from duty. But while my argument there hinged on the fact that only the agent's *belief* that an action is impermissible (as opposed to an action's *actual* impermissibility) can explain the claim that the question concerning the motive does not even arise, Kant's use of the term 'already' adds a further dimension to that explanation. It suggests that the question concerning an action's rightness arises *first*, and the question concerning its motive *second*. In the context of actions contrary to duty this means at the very least that when the first question has *already* been answered in a certain way, i.e. when actions are "*already* recognised as contrary to duty" (ibid.), then the second question simply does not arise: "with these *there is not even the question* whether they might have occurred from duty, since they even conflict with it." (Ibid., my emph.) As Kant's discussion of the various examples shows, the second question concerning the motive does very much arise when it comes to actions in conformity with duty: the point – or at least one important point – of that discussion is that such actions can be done from self-interest, or from immediate inclination, or from duty.

Thus rightly understood the subjective reading, and in particular Kant's insistence that impermissible actions are "*already* recognised as contrary to duty", that is, *before* we might even consider their moral worth, suggests a certain *order of deliberation*. From the agent's standpoint, the question whether a proposed course of action is in accordance with the moral law or conflicts with it typically arises *before* a decision is made and acted on. This implies, as we have seen, specifically that an agent who already recognises an action as contrary to duty cannot perform that action from the notion that it is required. But more generally, Kant's statement suggests that in his view deliberation about the morally relevant facts and principles, deciding on an appropriate course of action, acting (or failing to act) on that decision, and finally reflecting about one's motives proceeds in a certain order, and that for the agent the question concerning the motive only arises once the rightness or wrongness of the action has been established.

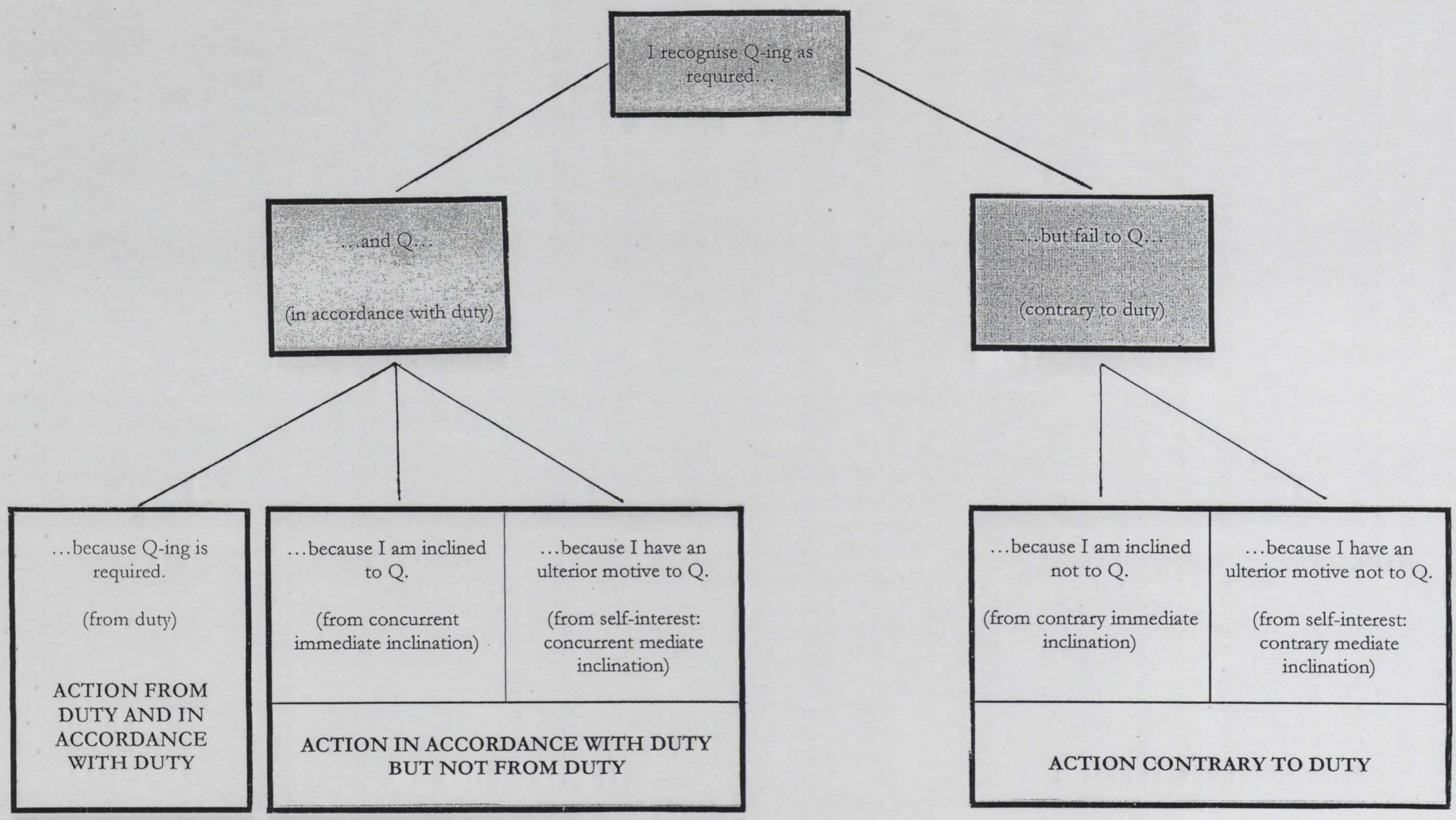
It is now clear how an agent-centred account of Kant's argument, implying, as it does, an order of deliberation and a dependence of the criterion of moral worth on the criterion of rightness, stands in stark contrast to the notion, insisted on by the proponents of the Logical Compatibles View, that Kant's twofold distinction between an action's rightness and its moral worth is one made by a third-person observer using two independent criteria, thus yielding four possible classes of action. If the question concerning the moral worth of the action arises only *after* the rightness of the action has been established, and only in relation to those actions which have been recognised as *conforming* to the criterion of rightness, then Kant's two criteria are *not* logically independent and compatible in such a way as to yield four possible classes of action.[104] Since, for the various reasons outlined in the last three chapters, the qualified agent-

[104] As I outlined in chapter 1, in identifying two attributes as logical compatibles, commentators insist or imply that their logical form is a double disjunctive such that either of the two attributes could be affirmed or denied of an action, without this predetermining the affirmation or denial of the other. From such a case there would indeed necessarily result four logical possibilities. On the other hand it is also clear that not any two features or attributes are logically compatible. So whether Kant's attributes of action, 'from duty' and 'in accord with duty', are logical compatibles, should, at least initially, be an open question. But Hardwig, along with most commentators and without explaining why he thinks this is so, simply presupposes that they obviously are: "Like any distinction between logical compatibles, this one yields four logically possible classes of action." (Op. cit., p. 68) This unwarranted assumption is decisive for much of the rest of the debate, for it first introduces the need to account for Kant's apparent oversight of action from duty but not in accord with duty.

centred reading which I defend constitutes the most consistent and generous interpretation of the text, and since such a reading is incompatible with the Logical Compatibles View, I propose that the latter must be rejected. In its place I suggest an account of Kant's statement which takes into consideration the implied first-person stance and order of deliberation.

Thus, the criterion of rightness, the agent's acting (or failure to act) on or at least in accordance with his recognition, and the criterion of moral worth, when combined from the agent's first-person standpoint and in the order in which the agent is confronted with them, give rise to only three possible types of action (diagram 2).

THE AGENT'S PERSPECTIVE



(diagram 2)

In contrast to the objective reading and the unqualified subjective reading, this account explains the various features of Kant's argument which remained a mystery on the basis of the traditional accounts: first of all it lays to rest the notion, expressed by adherents of both readings, that there is some sort of lacuna in Kant's account. Contrary to appearances, there is no missing type of action and thus no oversight on Kant's part. There are no four possible classes of action in the first place, simply because on the agent-centred first-person standpoint which Kant's argument implies, the criteria of rightness and moral worth are *not* logical compatibles. This fully accounts for the fact that the possibility of acion from duty but not in accordance with duty does not, and cannot, arise.

We are now also in a position to respond to a number of the textual difficulties raised by commentators. Firstly, it seems that there are no grounds for the initial complaint that "Kant never considers or even mentions [...] action from duty but not in accord with duty. One would have thought that he would mention this logical possibility, even if only in order to discount it as not really possible."[105] It is of course true that Kant does not consider action from duty but not in accord with duty, i.e. action under this particular description. But he does not do so precisely because he *does* discount the very possibility of *distinguishing* between a moral and a non-moral motive when the action is contrary to duty. Consequently, he passes over *all* actions recognised as contrary to duty. For Kant, then, all actions which are not in accordance with duty are of no interest whatsoever and he puts them aside. So in response to the common complaint we might say: Not only does Kant not consider actions from duty but not in accordance with duty, he does not consider actions which are neither in accordance with duty nor from duty either. He furthermore explains – if only very briefly – *why* he excludes them from the investigation, and he dismisses the need and the possibility to make the very distinction between different motivational grounds when an action is contrary to duty. Thus, with the *Groundwork* passage in mind and considered from the agent's perspective, the question concerning the motive simply does not arise once an action has been recognised as contrary to duty. Unless one were to include in Kant's framework a type of action which has the agent recognising his duty, but refrain from doing it *because* (not *despite*) it is his duty – i.e. from the motive of *devilishness*[106] – all

[105] Hardwig, op. cit., p. 68.

[106] I will not consider such a possibility in the present context, but Curzer suggests that Kant should allow for such a motive: "Kant should give up the claim that [...] there are only two types of motives. He should postulate a third motive which I shall call the 'motive of devilishness', and assign negative

action contrary to duty must be motivated by mediate or immediate inclination. In other words, from the agent's perspective actions which the agent already recognises as contrary to duty, if they are nevertheless performed, are always done from inclination. Secondly, this account makes the alternative accounts obsolete. The other approaches to explaining Kant's apparent oversight, i.e. the attempt to attribute to him an unfounded assumption that no action from duty can be contrary to it which I discussed in chapter 5.1., or the stipulative definition of the concept of duty to the same effect (chapter 5.2.), as well as the appeal to Kant's apparent denial of a problem of knowledge in morality (chapter 5.3.), and finally the suggestion, mentioned in chapter 1, that perhaps the *Metaphysics of Morals* might be the more appropriate place for Kant to consider actions contrary to duty but from duty, are all at best comparatively weak. If nothing else, they are also misleading, because in their reliance on the Identity Thesis they suggest a link between the Logical Compatibles View and the problem of action from mistaken moral judgement,[107] which further confuses the

---

moral worth to acts done from this motive. What is the motive of devilishness? Just as people motivated by duty perform right acts because they are right, people motivated by devilishness perform wrong acts because they are wrong." Cf. op. cit., pp. 309 ff.

[107] To cite only one example, this can be seen in the case of Hardwig. Finding the textual evidence inconclusive, he argues that the solution to the puzzle of the missing type of action is to show that, quite independently of the question whether Kant *did* in fact think that there are no mistaken moral judgements and hence no actions from duty but not in accordance with duty, Kant *must* have held this view because he wants to uphold both the *principle of autonomy* and the *principle of universality*. If *both* of these are to be upheld, Hardwig thinks, then there cannot be action from duty but not in accord with duty. "Given this view that the autonomy of the moral legislator and judge is necessary for moral obligation, we are in a position to see why Kant cannot admit that there are actions from duty but not in accord with duty. For [...] he also maintains that moral legislation and judgement must be objectively or universally valid - valid for all rational beings - and a mistaken moral principle of judgement obviously would not be objectively or universally valid, simply because it is mistaken. To admit that there can be actions from duty but not in accord with duty is to admit that autonomous moral judgements need not be objectively or universally valid and also that objectively or rationally valid moral judgements are not always reflected in a given agent's autonomous judgements." (Op. cit., p. 72) So Hardwig's *Kant* would have to conclude that there are no mistaken moral judgements, and this would explain why he fails to even mention the possibility of action from duty but not in accord with duty. But according to *Hardwig*, the conclusion should be quite the reverse: Mistaken moral judgements obviously do exist, and their existence reveals "a conflict or tension between the requirement that moral judgements be universally valid and the requirement that they be autonomous" (ibid., p. 75). Moreover, he concludes that this is "a fundamental problem for many theories of ethics" (ibid.), which leaves them in the following dilemma: "Because actions from duty but not in accord with duty exist, ethical theories cannot consistently maintain that moral principles and judgements must be both

issues, as I will attempt to show in the next chapter.

I conclude that as far as the analysis of the text is concerned, Kant's argument is sound on the basis of the first-person stance which he adopts and the order of deliberation which this implies. But one might wish to explore the issue apart from and beyond immediate textual concerns and ask what, if any, the reasons might be for adopting this stance in the first place. Are there underlying philosophical reasons for adopting the agent's viewpoint in the context of the investigation into the motive of duty? This question should be of concern to both the critic and the Kantian. From the latter's point of view, the aim of the inquiry would be to consider whether there are reasons to think of the subjective reading as not only the strongest reading of the dialectic as it is presented by Kant, but also as expressing something substantial and true. The critic, on the other hand, will want to examine whether Kant's reasoning is sound all things considered, not merely on the assumption that he adopts this particular stance, an assumption which one might not wish to share. In other words, the crucial question in both cases is: On what grounds, if any, might Kant be thought to be *right* to adopt this stance? This question is the subject of the next chapter.

---

autonomous and objectively or universally valid. And yet we seem to need to say that they must be both." (Ibid.)

Atwell's criticism of this argument is scathing: "Contrary to Hardwig, who holds that for Kant there can be no mistaken moral judgement (since they are 'autonomous' – whatever that might mean), Kant himself writes: 'I can indeed err at times in my objective judgement as to whether something is a duty or not […]' [MM VI:401]. The simple fact is that the principle of autonomy and the principle of universality, as criteria of rightness, are one and the same […]". See his *Ends and Principles in Kant's Moral Thought* (Dordrecht: Martinus Nijhoff, 1986), p. 144, fn.8. I agree with Atwell's rejection of Hardwig's argument, because the latter falsely attributes to Kant the view that one cannot be mistaken in one's autonomous judgements. But more importantly, I have argued that we must resist the notion, expressed most explicitly by Hardwig and Kerstein, but shared by others, that the absence of action from duty but not in accordance with duty from Kant's account has *anything* to do with his belief that knowing one's duty is easy, because there simply is no connection between the problem of knowledge in morality and the question concerning the possibility of action from duty but not in accordance with duty.

Chapter 9

REJECTION OF THE IDENTITY THESIS

One might think that even if the agent's stance accounts for Kant's reasoning, he is plainly wrong to adopt it. *Prima facie*, one obvious way to support this line of thought is by appeal to the idea that actions from mistaken moral judgement – or actions from duty but not in accordance with duty, in Kant's terms – clearly exist, or perhaps even that they are a common occurrence. Either way, this objection insists, a moral theory worth its salt must be able to account for such actions, and if Kant adopts an argumentative stance which precludes their possibility, then so much the worse for his account. An explanation of his reasoning might clarify why he does not consider action from duty yet contrary to it, but as far as the issue of action from mistaken moral judgement is concerned this just proves the inadequacy of Kantian moral theory.

To put this suggestion into the context of the terms of the debate, one might insist that if Kant's adoption of a first-person standpoint cannot accommodate the possibility of action from duty but not in accordance with duty, then that merely shows that this possibility will need to be examined from a third-person perspective. In this chapter I will critically examine this proposal. My strategy will be to argue that the very notion of 'action from duty but not in accordance with duty' is incoherent, and that Kant is therefore right to rule out such action, and that he is correct to do so for reasons entirely independent from his grounds for adopting the agent's standpoint. However, if, as I believe, the suggestion under consideration is correct insofar as the idea that there are actions from mistaken moral judgement is *not* incoherent, then this suggests that the view which identifies one with the other must be flawed, and consequently I will proceed to reject the Identity Thesis. This opens the way for a new analysis of action from mistaken moral judgement in Kantian terms and a consideration of whether, and to what extent, Kant's moral theory might accommodate such action, which will be the subject of chapter 10.

As we saw earlier, the Identity Thesis originates from the common idea that actions which are in fact morally impermissible can be motivated by an agent's sincerely held but mistaken belief that they are morally required, and it makes the claim that such actions are most appropriately captured in Kantian terms as 'actions from duty but not in accordance with duty'. In the last chapter my primary aim was to show that Kant's

argument at GMS IV:397.11-14 presupposes a first-person perspective and why on such a stance the possibility of actions from duty but not in accordance with duty is ruled out. I will now attempt to show that such action is impossible quite apart from any considerations concerning the standpoint from which it might arise. In order to do so I will need to reject the Identity Thesis and show that the description 'action from duty but not in accordance with duty' does not constitute an appropriate analysis in Kantian terms of action from erroneous moral judgement.[108]

1. The incoherence of the very notion of 'action from duty but not in accordance with duty'

It is worth reconsidering how the description 'action from duty but not in accordance with duty' came about in the first place. Such actions are said to be 'from duty' because they "are motivated by *the desire to do* one's duty"[109], because the agent "*intends to do* his duty"[110] and does "what he *takes to be* his duty [...] because it *is seen to be* his duty"[111]. Such an agent acts from the notion that the action is required, that is, from duty. But because the action in question does not correspond to the actual requirements of morality, the agent must be *mistaken* about what his duty is. "Because he is mistaken, he is in fact not acting in accord with duty"[112]. Now it is clear that the very description of 'action from duty but not in accordance with duty' involves not only, like the other three types of action, two separate judgements, one regarding the

---

[108] While the various considerations in favour of the agent-centred interpretation of Kant's statement (chapter 8), by virtue of the first-person perspective and the agent's order of deliberation implied by it, undermine the assumption expressed by the Logical Compatibles View, the current argument, if successful, works against both this view and the Identity Thesis. If the very notion of action from duty but not in accordance with duty can be shown to be incoherent, then, contrary to the Identity Thesis, action from mistaken moral judgement cannot *be* action from duty but not in accordance with it – unless, that is, one were willing to argue that the idea of action from mistaken moral judgement is itself incoherent. And if the very notion of action from duty but not in accordance with duty is incoherent, then this in itself, quite apart from the adoption of any particular argumentative stance, would suggest that the Logical Compatibles View must be mistaken in its insistence that Kant's twofold distinction should give rise to four classes of action. Thus my arguments in this chapter are explicitly directed against the Identity Thesis, but can be taken to apply to both theses.

[109] See Hardwig, op. cit., p. 69 (emphasis added).

[110] Ibid. (emphasis added).

[111] Ibid. (emphasis added).

[112] Ibid. (emphasis added).

action's rightness and one concerning its moral worth, but also, *unlike* the other three types of action, two different standpoints from which the two judgements are made, respectively. That this should be so should come as no surprise and might seem to constitute the very nature of such actions and therefore the very point of considering them in the first place: action from mistaken moral judgement must by its very nature involve, on the one hand, a judgement that a certain action is morally required which serves as the ground for that action, and, on the other, the recognition of the judgement as erroneous. Now it is clear that the (mistaken) judgement that an action is morally required and the judgement that this judgement is indeed mistaken cannot be made by the same person (in the same capacity and at the same time), whether we consider the person as agent or as observer. Rather, the action is described as 'from duty' on the basis of the *agent's* own judgement that the action is required, and his acting on this belief. But the judgement that the action is in fact '*not* in accordance with duty' – a judgement which first identifies the original judgement as *mistaken* – is made *contrary* to the agent's belief, and therefore cannot be the agent's *qua* agent.[113]

There is, consequently, an asymmetry between the description of action from mistaken moral judgement as 'action from duty but not in accordance with duty', on the one hand, and the descriptions of the other three types of action on the other. It is this asymmetry which seems to me to be sufficiently noteworthy to raise our suspicion concerning the appropriateness of the description. If we consider the three types of action arising from Kant's twofold distinction, viz. 'action from duty and in accordance with duty', 'action in accordance with duty but not from duty', and 'action neither in accord with nor from duty', it is clear that each instance of the term 'duty' in these expressions stands simply and unambiguously for whatever duty is in the circumstances. Nor does it matter whether one favours a subjective reading of the term 'duty' as referring to 'that which the agent recognises as his duty', or an objective interpretation according to which 'duty' stands for 'that which the moral law in fact requires'. In either case the term 'duty' refers unambiguously: On an objective reading of the term 'duty', i.e. taking it to refer to 'that which the moral law *in fact* requires', it is impossible for an agent to act from duty but not in accordance with duty, because it is impossible for an agent to act from the judgement that a certain action is required when, *ex hypothesi*, that judgement is correct and the action is indeed required, but yet to fail to act in accordance with it. Equally, on a subjective reading of 'duty' which

[113] Of course, while corrective judgement cannot be passed by the same person in the same capacity at the same time, it may be passed by the same person at a later time (the agent adopting the stance of observer of his former self) or by another person at the same time. The point is that either way two different standpoints are involved.

takes the term to stand for 'that which *the agent* recognises to be required', it is impossible for an agent to act from duty but not in accordance with duty, because it is impossible for an agent to act from the notion that the act is required, when, *ex hypothesi*, this implies that it *is* required, and yet fail to act in accordance with it. This confirms that the notion of 'action from duty but not in accordance with duty', unlike the descriptions of the other three types of action, requires that the term 'duty' does not stand unambiguously for either 'that which the moral law *in fact* requires' or 'that which *the agent* recognises to be required', but rather that it refers to the former when the action's rightness is concerned and to the latter with regard to its motive. And if we remember that the term 'action from duty but not in accordance with duty' is intended to serve as a description of action from *mistaken* moral judgements, then it is clear that it *must* do so, since otherwise it could not successfully pick out action involving a mistake. That is to say, the very nature of action from mistaken moral judgement implies that its description involves judgements made from several standpoints.

I have, of course, argued that there are a number of good reasons for preferring the subjective over the objective reading. But that is not the issue here; the point, rather, is simply that on either reading the term duty is unambiguous in three of the four classes of action, i.e. the classes of action which Kant mentions in the *Groundwork*, while it is ambiguous in one, namely in the case of the description of 'action from duty but not in accordance with duty' which he is accused of ignoring for no good reason. My suggestion, then, is that there is – and that the very nature of action from mistaken judgement implies that there *must* be – an asymmetry in this respect between the latter and the former, and that this very asymmetry might offer a reason why Kant is right to exclude the possibility of action from duty but not in accordance with duty from his discussion of the concept of duty in the *Groundwork*.

The first conclusion to be drawn from this is that on the basis of either a first-person or a third-person standpoint, the notion of 'action from duty but not in accordance with duty' is *incoherent*. That is to say, there are no 'actions from duty but not in accordance with duty' *simpliciter*. This is so whether one interprets Kant as arguing from an agent's or an observer's standpoint, because it is logically and psychologically impossible to even conceive of such action from either the agent's or the observer's standpoint alone.[114] It is impossible because such action by definition

[114] John Atwell, in a study which predates most of the other contributions to this debate and which has largely been ignored in this context by later commentators, makes this point in relation to the subjective reading only. He states that since Kant in his discussion of the motive of duty in the *Groundwork*

involves the *combination* of the two standpoints: the notion of 'action from duty but not in accordance with duty' implies an observer who, with the benefit of hindsight and from a third-person standpoint, issues his all-things-considered second-order judgement that another agent's (or his own former) first-order judgement that an action is morally required is, in fact, mistaken. We will see in chapter 10 how such second-order judgements might be accommodated by Kant's moral theory. But for the moment, and as far as the *Groundwork* passage is concerned: if nothing else, it seems far-fetched to interpret Kant's twofold distinction of the concept of duty as an extravagant device allowing for elaborate combinations of standpoints, especially when in all but one class of action such extravagance is entirely superfluous.

The second point is that the term 'action from duty but not in accordance with duty' is inadequate as a rendition of the common notion of action from erroneous moral judgement. This would suggest a rejection of the Identity Thesis. We saw that on either the subjective or the objective reading of 'duty', the term is unambiguous in three of the four classes of action, and ambiguous in one, namely in the description of 'action from duty but not in accordance with duty'. Thus our very ability to render action from mistaken moral judgement into Kantian terms rests on an ambiguous, equivocal sense of the term 'duty', at least if done in the way suggested by the Identity Thesis.

2. An alternative account of action from mistaken moral judgement

I suggest that the only way to avoid this ambiguity is to make it explicit: An agent who acts from a mistaken moral judgement acts *from* what he believed to be his duty but

---

"adopts the standpoint of an agent or decision-maker [...], there is no possibility of 'action from duty but not in accord with duty', since this would require an agent's thinking to himself: 'I recognize that my action does not accord with duty, yet I shall nevertheless do this action from the motive of duty.' And that thought is probably logically (to say nothing of psychologically) impossible." See his *Ends and Principles in Kant's Moral Thought* (Dordrecht: Martinus Nijhoff, 1986), p. 29. Atwell does not consider why Kant adopts this standpoint, nor how this might be extrapolated from the text, but he is surely right to insist that if we take Kant to argue from a first-person standpoint of the agent, then the very notion of 'action from duty but not in accordance with duty' is indeed logically and psychologically questionable – a point which has been overlooked or ignored by subsequent contributors to the debate. My point is that this notion is incoherent *per se*, not merely on the subjective reading.

performs an action which we, the observers, or perhaps he himself at a later point in time, recognise *not* to have been his duty. The first use of 'duty', concerning the criterion of moral worth, refers to 'what the agent mistakenly took to be his duty', whereas the second, regarding the criterion of rightness, refers to 'what we observers with the benefit of hindsight now believe to have been the agent's duty'. This spells out the exact nature of the ambiguity of the term 'duty' in the description 'action from duty but not in accordance with duty'. But in so doing we have found a more exact way of describing in Kantian terms the common notion of action from mistaken moral judgement. We have, therefore, in effect rejected as inadequate the idea that action from erroneous moral judgement is captured by the notion of 'action from duty but not in accordance with duty'. That is, we have rejected the Identity Thesis. The very notion of 'action from duty but not in accordance with duty' is inadequate because it is a complex expression which rests on an equivocal use of the term duty. An analysis of the complexity of the expression will clear up the ambiguity and yield an alternative, more adequate interpretation of mistaken moral judgements in Kantian terms. Thus we get: 'Action *from* what the agent believed to be his duty but which we, the observers, or the agent in retrospect, recognise *not* to have been his duty', or more simply: 'Action from *apparent duty* but not in accord with *real duty*', or again: 'Action from *subjective duty* but not in accordance with *objective duty*'. But given the complexity of this description and the complexity of the state of affairs it describes, and the different standpoints it involves, it is now obvious why a class of action of this kind does not fit into the fourth category of action according to the classification suggested by the combination of the Logical Compatibles View and the Identity Thesis (see diagram 1 in chapter 4 above).

We can conclude that the Identity Thesis must be rejected for three reasons: First, the notion of 'action from duty but not in accordance with duty' is *incoherent* because regardless of whether one interprets the term 'duty' subjectively as 'that which the agent recognises as morally required', or objectively as 'that which the moral law in fact prescribes', it is psychologically impossible to perform such action and logically impossible to even conceive of it. Secondly, the term 'action from duty but not in accordance with duty' is inadequate as a rendition of the common notion of action from mistaken moral judgement because it rests on an equivocal use of the term duty. Finally, when this ambiguity is laid bare it becomes clear that such action by definition involves a *combination* of the agent's and an observer's respective standpoints. To the agent's original judgement that an action is required, his acting or failing to act in accordance with that recognition, and his reflections on the motive from which he

might have done so, there is now added a higher-order level of reflection about the original judgement, which, crucially, allows for the possibility that the original judgement was mistaken. This second-order judgement might be made about another agent's moral judgement and act, or about one's own past judgements and actions. The point is that in either case one judges as mistaken from a third-person standpoint a moral judgement which was made from a first-person standpoint. Whether or to what extent the possibility of such second-order judgements requires an insight into the criteria for the rightness of the action which was not available to the agent, or additional information, or the benefit of maturity or at least hindsight – these are important questions worth exploring, but I cannot consider them in the present context. The crucial point here is merely that to identify a moral judgement as mistaken is to adopt a third-person perspective from which one can issue a second-order judgement that a particular first-order judgement is mistaken.[115]

To illustrate my conclusions with some examples of action from mistaken moral judgement, let us consider first the case of an agent's first-order judgement that an action is morally required. Let us assume further that the agent, because in acting in accordance with his judgement he would have to overcome his contrary inclinations, has reason to believe that he (as far as it is ever possible to know one's motives) is acting from duty; or the case of an agent who is aware that he is, from inclination, acting contrary to what he judges to be morally required of him in the context and thus believes (again, as far as it is ever possible to have reason to believe) that he is not acting from duty; or again, an agent who, in accordance with duty, performs an action which he judges to be morally required and towards which he is also already naturally inclined, thus believing that he is doing his duty but not acting from duty. In any such case, our judgement – or, for that matter, the agent's own later judgement – that his original moral judgement as to the rightness of the action was in fact mistaken would introduce a second-order level of reflection into the picture which does indeed allow for the possibility of a mistake, but which also clearly goes beyond the framework of Kant's twofold distinction of the concept of duty. Again, even apart from any other

---

[115] Observing that in "our ordinary manner of speaking, we say that something was the right act for the wrong reason", Steven Sverdlik remarks that "[i]t is odd that the expression 'the wrong act for the right reason' is not common. It is certainly intelligible, though." See his 'Motive and Rightness', *Ethics* 106 (1996), p. 327. If my analysis of action from duty but not in accordance with duty is correct, the conclusions I have drawn should be sufficient to address the puzzlement expressed here and by showing that the expression, contrary to appearances, is *not* intelligible.

considerations, this shows why it would be wrong to expect a treatment of actions from mistaken moral judgements within that framework, and that it would be equally foolish to expect it at all, if one is looking for such an account in terms of 'action from duty but not in accordance with duty'.

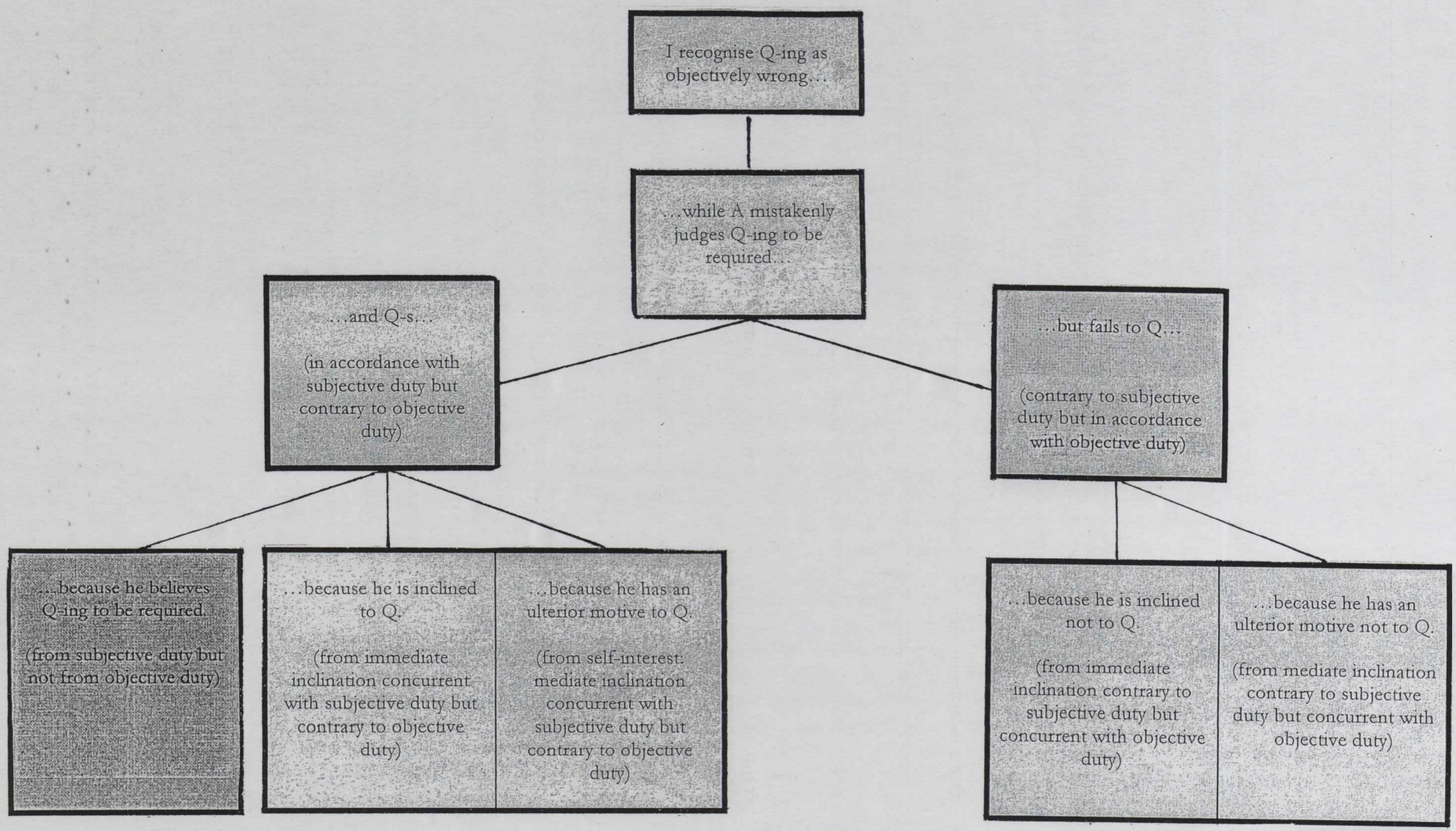
But furthermore, as the three examples above suggest, there is no reason why the application of a second-order judgement to a first-order judgement, such as to label the latter mistaken, should only be made about a first-order judgement of the type described in the first example, which is the type that according to the Logical Compatibles View and the Identity Thesis Kant failed to consider, i.e. an action which the agent judges to be morally required and performs because he recognises it as such. Rather, such a second-order judgement can in fact be made about *any* action involving *any* type of first-order judgement; e.g., about the action which an agent judges to be required and performs from inclination, or about the action which the agent does from inclination despite judging it to be contrary to duty.

3. Conclusion

Thus we find that the correct description of action from mistaken moral judgement, i.e. 'action from what the agent believed to be his duty but which we, the observers, or the agent in retrospect, recognise not to have been his duty' designates only one of many types of action involving an observer's second-order judgement. This type of action not only cannot fit into the typology of actions which Kant's twofold distinction yields, but it is also not the only type that fails to conform to that framework. Rather, it is merely one example of an entirely different genus of act-description, i.e. one involving both a first-order agent's judgement, and a second-order observer's judgement. Thus the observer's third-person perspective, which gives rise to second-order judgements concerning agent's first-order judgements, their acting (or failing to act) on or at least in accordance with this recognition, and their possible motives, yields the following scheme (diagram 3), which includes our 'missing' type of action.[116]

[116] Note that, as noted before, 'A' can stand either for another agent or for the observer's 'former self'. In both cases an observer's stance is adopted towards A's deliberation, decision, action, and motivation.

ACTION FROM MISTAKEN MORAL JUDGEMENT: THE OBSERVER'S PERSPECTIVE



(diagram 3)

Having explained and justified why action from duty but not in accordance with duty is impossible, and having determined the appropriate description in Kantian terms of action from mistaken moral judgement, we have now come to the conclusion of the main business of this study.

If my analyses are correct and a coherent application of the conceptual framework developed in Kant's twofold distinction can only be guaranteed if it is limited to the agent's first-person perspective, there are, as far as I can see, two possible ways forward: If we insist that moral praise and blame, and with these the notions of 'acting from duty' and 'moral worth', are essentially an observer's response to conduct and thus imply a third-person perspective, then my account might serve as an argument for the internal incoherence of Kant's twofold distinction and the need for a complete separation of, on the one hand, first-person considerations concerning normative criteria for the rightness of actions, and on the other, third-person evaluations of motivation and moral worth.

Alternatively, one might consider the issue in the following way: Since I have rejected the Identity Thesis and the Logical Compatibles View, Kant's rejection of the possibility of action from duty but not in accordance with duty does not by itself have any substantial implications regarding the possibility, imputability, and possible moral worth of actions from mistaken moral judgement, which pose an interesting problem in their own right. What my conclusions do imply in that case, however, is that such a debate cannot meaningfully be conducted in terms of Kant's agent-centred twofold distinction of the concept of duty. In the following chapter I will briefly consider the parameters within which this second option might be pursued and examine one specific approach to the issue, namely the appeal to Kant's control and effort conditions.

Chapter 10
MORAL WORTH
AND ACTION FROM MISTAKEN MORAL JUDGEMENT

1. Outline of the problem

Having separated the question of action from mistaken moral judgement from the problem of the possibility of action from duty but not in accordance with duty, we are now in a position to consider the former in its own right. The problem can be put as follows: Does Kant acknowledge the possibility of actions from erroneous moral judgement, and if so, what is their status (in terms of moral worth)? Can an action which is objectively impermissible have moral worth when the agent mistakenly believes it to be required, and if so, under what conditions? In this chapter I will consider Samuel Kerstein's suggestion that Kant's requirement that actions must be objectively right in order to have moral worth should be revised in such a way as to allow that some actions which are objectively contrary to duty might yet have moral worth. My aim here is not to come to a conclusion one way or another on this question, but to outline the parameters of the debate and to bring into focus the various ways in which Kant's control and effort conditions (see chapter 3.2.) might be appealed to by both sides of the debate.

Kerstein suggests that while Kant's account of duty in the *Groundwork* would clearly not allow for such a possibility, certain remarks in his later writings suggest that he was moving towards the idea that some actions which are objectively contrary to duty might have moral worth. Initially, it might seem that this is impossible for Kant, for he holds that for an action to have moral worth it must be done from duty, that is, from the notion that it is required, and it must also fulfil two further conditions:

(1) The action must be morally required.

(2) The agent must summon all the means in his power to realize the end of his action.[117]

Let me, first of all, set aside two types of action involving error of judgement which are, at least in the present context, and from a Kantian perspective,

[117] These are two of Kerstein's four Kantian conditions for moral worth. The remaining two concern the motive of duty and are not relevant in the present context. See his *Kant's Search for the Supreme Principle of Morality* (Cambridge: Cambridge University Press, 2002).

uncontroversial.

(A) Actions involving a morally irrelevant and blameless mistake: If a person poisons a friend in the justified but mistaken belief that he is giving him Aspirin, then, insofar as the consequences of the action are entirely beyond the control of the agent and he was motivated by the belief that his action was morally required, his action is clearly praiseworthy, i.e. it has moral worth. This case is *uncontroversial* because the error is *clearly* not the agent's: whatever factors led to the undesirable consequences, they were not within the agent's control.

(B) Actions involving a morally relevant and blameworthy mistake: These are cases where the agent acts contrary to what is morally required because he makes a morally relevant mistake, such as neglecting to deliberate sufficiently about what his duties are, or failing to take into consideration another's legitimate interests. In these cases the agent clearly is blameworthy to the extent that he failed to do something which he should and could have done, and to that extent the action does not have moral worth. This case is *uncontroversial* because it is *clear* that the errors which led the agent to act contrary to duty are morally relevant and within the agent's control.

It is worth noting that what settles the question in both cases is an appeal to Kant's control condition and effort condition, respectively: The control condition, '*an action's moral worth cannot be diminished by what is beyond the agent's control*' applies in the first case (A), and the effort condition '*moral worth is diminished by lack of constraint and effort on the agent's part*' in the case of (B). Now the crucial and controversial question is whether Kant does or should allow that there can be action of the following kind.

(C) Action involving a morally relevant but not blameworthy mistake: Such action would have to meet the following criteria, and any action that fulfils these criteria has moral worth: The action is contrary to duty; the agent (at least at the time of the action) sincerely but mistakenly believes the action to be required; and the agent has done everything in his power to determine his duty, and yet failed, i.e. the mistake involved must be due to factors beyond the agent's control.

It is highly *controversial* whether there are any actions of this type, and if so, what they might look like, because it is not at all clear how an error on the agent's part could be described both as not in the agent's power to avoid (thus rendering the agent blameless) and yet morally relevant. *Prima facie*, it seems that Kant should and would clearly reject the possibility of such action, for Kant holds that 'ought implies can', or, as we might explicate this in the terms of the current problem: *if x is relevant to the moral evaluation of a person, then x must either be available to that person, or, if x is not available to them, the person is responsible for x not being available,* or, somewhat simplified: *Whatever is*

relevant to the moral evaluation of a person must in principle be available to that person. Now, it seems clear that this sort of principle is, again, a version of Kant's control and effort condition: Whatever is unavailable to an agent is outside of his control, what an agent cannot control he is not morally accountable for, and *moral worth is not enhanced or diminished by what lies beyond an agent's control.* Thus it appears that in an attempt to reject the very possibility of type C actions, the obvious strategy for the Kantian would be an appeal to the control condition.

2. Kerstein's asymmetry thesis

Samuel Kerstein claims that Kant should relinquish the requirement that for an action to have moral worth it must be in accordance with duty. "Kant should acknowledge that actions contrary to duty [...] can have moral worth."[118] His argument for this claim is that there is in Kant's view of moral worth a "disturbing asymmetry"[119] between two kinds of failure: While failure to judge correctly whether one's action is morally permissible precludes it from having moral worth, failure to attain the end of one's action does not.[120] Yet both kinds can be such that the failure is not a failure of will (i.e. a succumbing to inclination for which the agent can be held responsible), but a matter of one's circumstances, upbringing, or cognitive abilities, for which one is not responsible.[121] That is, both kinds of failure can be due to circumstances beyond the agent's control and should therefore, in the spirit of Kantianism, be immaterial to an action's moral worth.[122]

The main example Kerstein cites is that of Colonel Mikavitch[123]: well-educated and morally reflective, she has embraced the Categorical Imperative as the supreme principle of morality. Unfortunately, she finds herself in a situation where she will be captured and, under torture, reveal the location of innocent people who her captors are intent on killing, unless she kills herself. After careful consideration, against her own inclination, and with great courage she decides that taking poison in order to take her own life is the only way she can save innocent lives. So the colonel does everything

---

[118] Ibid., p. 119.

[119] Ibid.

[120] Ibid.

[121] Cf. ibid., p. 121.

[122] Cf. ibid., p. 119.

[123] Cf. ibid., p. 121f.

that could be asked of a conscientious moral deliberator and agent. She does not succumb to inclination, nor does she ignore any of the relevant facts. Hence, Kerstein argues, her action should be considered to have moral worth and be an expression of a good will, *irrespective* of whether or not we agree that taking her own life was in fact what was required of her. But "Kant would insist that the colonel's suicide could have moral worth only if it were in accordance with duty".[124] Kerstein then considers whether the colonel's action was *in fact* in accordance with Kantian duty and concludes that it is both difficult to answer this question conclusively and unclear whether Kant himself would have condemned the colonel. However, this is not the point he is making. Rather,

> the point this case is designed to illustrate is simple. According to Kant, if the colonel's action was not morally permissible, it would follow that it could not have had moral worth. Suppose that – after due consultation with the world's greatest Kantian casuists – we discovered that, measured by Kant's principle [i.e. the Categorical Imperative], the colonel's suicide was morally impermissible. On my view, this discovery would not, in itself, warrant the conclusion that her action failed to have moral worth. Intuitively speaking, as far as moral worth goes, it just would not matter. The colonel did her best to determine whether her course of action was morally permissible. If she did not succeed, *her failure stemmed, it seems, not from lack of sincere effort but rather from the limits of her cognitive capacities.* Kantian casuistry is hard. Much as it is often beyond her control whether the world cooperates and she succeeds in her efforts to promote the happiness of others, so is it beyond her control whether she succeeds in discerning whether a given action meets the standard of Kantian moral permissibility.[125]

Acknowledging that this kind of example might "have an artificial ring"[126] to it, Kerstein nevertheless insists that there are

> plenty of people who throughout their lives have done their best to determine what is right but, measured by the standard of the Categorical Imperative, have failed. In the determination of an action's possible moral worth, Kant discounts the effects an agent's action actually has in the world, apparently on the grounds that these effects are beyond her control. Nevertheless, at least in the *Groundwork*, he does not discount mistaken moral judgements, even though, when an agent makes her best effort to get it right, these also seem to be beyond her control. It is this *questionable asymmetry* that the [case is] designed to bring into focus.[127]

---

[124] Ibid., p. 122.

[125] ibid., p. 122f., emph. added.

[126] Ibid., p. 124., emph. added.

[127] Ibid.

3. Two objections and Kerstein's replies

Kerstein considers the following objection to his point: The asymmetry thesis rests on the notion that it is possible for an agent, in the attempt to determine what is right, to try his utmost and yet fail.[128] But Kant rejects this notion: he thinks that what is morally required is plain and obvious to the common understanding (cf. GMS IV:403f.). Agents do not fail to act as the moral law requires unless swayed by their inclinations. Ordinary reason not only always has (a version of) the Categorical Imperative in view, it also "knows very well how to distinguish in every case that comes up what is good and what is evil, what is in conformity with duty or contrary to duty" (GMS IV:403f.). The objection thus appeals to *the alleged transparency of morality* and the absence of a problem of knowledge in morality which we encountered in chapter 5.3. Kerstein's reply to this objection is that even if we grant that ordinary reason always has (something like) the Categorical Imperative in view, "it is a strain to deny that even with the best intention and effort, we might fail to apply this standard correctly. After all, haven't we behind us two hundred years of scholarly disagreement on how to apply the Categorical Imperative test?"[129]

In anticipation of the further objection that allowing the colonel's action to have moral worth would rationally compel us to make the unwelcome admission that *any* action which is done from a mistaken belief that it is required – including the most *odious actions*[130] – must have moral worth, Kerstein offers a revised account of moral worth. His suggestion is that Kant should drop the first condition (the action's conformity with duty) and put in its place the weaker condition that the agent, having made a *sincere effort* at determining her duty, must believe that the action is in conformity with duty. So for an action to have moral worth it must meet the following revised version of the two conditions:

(1*) The agent must make a *genuine effort* to determine what her duty is.[131]

(2) The agent must summon all the means in his power to realize the end of his action.

Kerstein thinks most odious actions would fail to meet at least some of these

---

[128] Cf. ibid., p. 127.

[129] Ibid., p. 128

[130] See the examples cited in chapter 4.3. above.

[131] Ibid., p. 130

conditions, but admits that he "cannot prove it to be impossible [that some odious action would fulfil all conditions], thereby giving [the] action moral worth".[132] We cannot "avoid the painful conclusion that being faithful to the spirit of Kantianism – and, I believe, of ordinary views regarding moral worth – requires us to admit the possibility (though by no means the likelihood) that some terribly wrong actions have moral worth".[133]

4. Mistaken moral judgements, moral worth, and the appeal to Kant's control and effort conditions

In addition to the above arguments, Kerstein finds further support for his suggestion that the conditions of moral worth in the *Groundwork* should be revised in Kant's own later writings on the subject, especially in the *Metaphysics of Morals*. Here, Kerstein argues, "Kant makes an acknowledgement that [...] would have been welcome in the *Groundwork*: without being led astray by his inclinations, an agent can make an error in determining what his duty is", and Kant "also suggests that when [he] does, the agent is not morally blamable."[134] In his discussion of conscience in the *Metaphysics of Morals* Kant makes the following points:

> [C]onscience is practical reason holding the human being's duty before him for his acquittal or condemnation in every case that comes under a law. [...]
>
> [A]n erring conscience is an absurdity. *For while I can indeed be mistaken at times in my objective judgement as to whether something is a duty or not*, I cannot be mistaken in my subjective judgement as to whether I have submitted it to my practical reason (here in its role as judge) for such a judgement. [...]
>
> *But if someone is aware that he has acted in accordance with his conscience, then as far as guilt or innocence is concerned nothing more can be required of him.* It is incumbent upon him only to enlighten his *understanding* in the matter of what is or is not duty.
>
> The duty here is only to cultivate one's conscience, to sharpen one's attentiveness to the voice of the inner judge and to use every means to obtain a hearing for it (hence the duty is only indirect). (MS VI:400f., emph. added)

Kerstein thinks that Kant is here moving towards a position that acknowledges that "[w]hen an agent is heeding his conscience, that is, doing what he believes the

---

132 Ibid.

133 Ibid., p. 132.

134 Ibid., p. 125.

Categorical Imperative to prescribe, he is not morally at fault for acting contrary to duty".[135] And, most significantly in the present context of our discussion, he claims that Kant is led to adopt this position by his acceptance of the control condition:

> Recall that in his discussion of moral worth in the *Groundwork* Kant appeals to our intuition that an agent's action is not to be disqualified from having moral worth by anything we take to be outside of his control. By claiming in the *Metaphysics of Morals* that a person is not always morally blamable for acting contrary to duty, he is, in effect, admitting that whether an agent acts contrary to duty can be determined by factors outside of his control – for example, how adept he is at applying the Categorical Imperative. But now suppose that an individual not only acts in accordance with his conscience but does so for its own sake. He obeys his conscience simply because it is the right thing to do. It seems clear not only that this person might end up acting contrary to duty but that her doing so could stem from conditions over which she has no real power. In this case, the very intuition to which Kant appeals in the *Groundwork* would direct him not to disqualify the agent's action from having moral worth. By acknowledging that when he abides by his conscience, an agent can blamelessly abide by his duty, *Kant sets himself on a path toward the view that morally impermissible actions can have moral worth.* [...]
>
> I suspect that he would concur that even if Colonel Mikavitch acted contrary to duty, her action could have moral worth. According to the example, she had the Categorical Imperative in view, and if she acted contrary to it, it was owing solely to a mistake in applying it. [... In summary, in the *Metaphysics of Morals*] *Kant moves toward recognizing the moral worth of some morally impermissible actions,* namely those which stem from errors in applying the Categorical Imperative.[136]

Now, to return to our original point: Kerstein appeals to Kant's control condition in order to argue that there *are* actions contrary to duty which have moral worth (C), and – unless I am mistaken in attributing this position to him – Kant would appeal to the same principle in an effort to *reject* their very possibility. It is not clear that there is an *obvious* flaw in either argument, but it seems clear that both cannot be right. I suspect that there is either some non-obvious flaw in one of the arguments, or an ambiguity in the way in which Kant and Kerstein respectively would appeal to the control condition.

5. Concluding remark

It was not the ambition of this chapter to provide a solution to the problem of the possible moral worth of action from mistaken moral judgement. However, I hope to

---

[135] Ibid., p. 125.

[136] Ibid., p. 126, emph. added.

have suggested a starting-point and outlined some of the parameters for a fruitful debate in this area of Kantian moral theory, in particular with respect to the role Kant's control and effort conditions might play in this context. The main aim of this study, on the other hand, was to establish the impossibility of conducting such a debate in terms of the notion of 'action from duty but not in accordance with duty' – and this I do consider to have shown conclusively.

BIBLIOGRAPHY

COMMENTARIES ON KANT'S *GROUNDWORK*

Guyer, Paul: *Kant's Groundwork for the Metaphysics of Morals. A Reader's Guide*, London: Continuum, 2007.

Horn, Christoph, C. Mieht, N. Scarano: *Immanuel Kant: Grundlegung zur Metaphysik der Sitten - Kommentar*, Frankfurt am Main: Suhrkamp, 2007

Kaulbach, Friedrich: *Immanuel Kants 'Grundlegung zur Metaphysik der Sitten' – Interpretation und Kommentar*, Darmstadt: Wissenschaftliche Buchgesellschaft, 1996.

Schönecker, Dieter & A. Wood: Kant's *Grundlegung zur Metaphysik der Sitten: Ein Einführender Kommentar*, Paderborn: Schöningh, 2007 (3rd ed.).

Sedgwick, Sally: *Kant's Groundwork for the Metaphysics of Morals. An Introduction,* Cambridge: Cambridge University Press, 2008.

Timmermann, Jens: *Kant's Groundwork for the Metaphysics of Morals. A Commentary*, Cambridge: Cambridge University Press, 2007.

Wolff, Robert P.: *The Autonomy of Reason. A Commentary on Kant's Groundwork of the Metaphysic of Morals*, New York: Harper & Row, 1973.

STUDIES OF KANT'S MORAL PHILOSOPHY (MONOGRAPHS)

Acton, H.B.: *Kant's Moral Philosophy*, London: Macmillan, 1970.

Allison, Henry E.: *Kant's Theory of Freedom*, Cambridge: Cambridge University Press, 1990.

Atwell, John E.: *Ends and Principles in Kant's Moral Thought*, Dordrecht: Martinus Nijhoff, 1986.

Bache, Kurt: *Kants Prinzip der Autonomie im Verhältnis zur Idee des Reiches der Zwecke*, Kant-Studien Ergänzungshefte: No. 12, Würzburg: Arnulf Liebing, 1909.

Baron, Marcia W.: *Kantian Ethics almost without Apology*, Ithaca & London: Cornell University Press, 1995.

Bittner, Rüdiger: *Moralisches Gebot oder Autonomie*, Freiburg/München: Alber, 1983.

David Cummiskey: *Kantian Consequentialism*, Oxford: Oxford University Press, 1995.

Fairbanks, Sandra Jane: *Kantian Moral Theory and the Destruction of the Self*, Boulder, Colorado: Westview Press, 2000.

Falaturi, Abdoldjavad: *Zur Interpretation der Kantischen Ethik im Lichte der Achtung*, Diss. 1965.

Grenberg, Jeanine: *Kant and the Ethics of Humility: A Story of Dependence, Corruption, and Virtue*, Cambridge: Cambridge University Press, 2005.

Guyer, Paul: *Kant and the Experience of Freedom. Essays on Aesthetics and Morality*, Cambridge, 1993.

Herman, Barbara: *The Practice of Moral Judgement*, Cambridge/ Mass. & London, 1993.

Heubült, Willem: *Die Gewissenslehre Kants in ihrer Endform von 1797 – Eine Anthroponomie*, Bonn: Bouvier, 1980.

Hill, Thomas E. jr.: *Dignity and Practical Reason in Kant's Moral Theory*, Ithaca & London, 1992.

——: *Human Welfare and Moral Worth: Kantian Perspectives*, Oxford: Clarendon Press, 2002.

Kaulbach, Friedrich: *Das Prinzip Handlung in der Philosophie Kants*, Berlin: de Gruyter, 1978.

Kerstein, Samuel J.: *Kant's Search for the Supreme Principle of Morality*, Cambridge: Cambridge University Press, 2002.

Köhl, Harald: *Kants Gesinnungsethik*, Berlin: de Gruyter, 1990.

Korsgaard, Christine M.: *Creating the Kingdom of Ends*, Cambridge: Cambridge University Press, 1996.

Moore, A.W.: *Noble in Reason, Infinite in Faculty. Themes and Variations in Kant's Moral and Religious Philosophy*, London: Routledge, 2003.

Neiman, Susan: *The Unity of Reason*, Oxford 1994.

Nisters, Thomas: *Kants Kategorischer Imperativ als Leitfaden humaner Praxis*, Freiburg/ München: Alber, 1989.

Paton, H.J.: *The Categorical Imperative: A Study in Kant's Moral Philosophy*, London: Hutchinson, 1947.

——: *The Moral Law*, London: Hutchinson, 3rd ed. 1956.

Reath, Andrews: *Agency and Autonomy in Kant's Moral Theory*, Oxford: Clarendon Press, 2006.

Reiner, Hans: *Pflicht und Neigung*, Meisenheim/Glan: Anton Hain, 1951.

Ross, W.D.: *Kant's Ethical Theory. A Commentary on the "Groundwork"*, Oxford: Clarendon Press, 1954.

Ross, W.D.: *The Right and the Good*, Oxford: Clarendon Press, 1930.

Schiller, Friedrich: *Über Anmut und Würde* (1793), in Bittner & Cramer (eds.): *Materialien zu Kants 'Kritik der praktischen Vernunft'*, Frankfurt am Main: Suhrkamp, 1975, p211-220.

Sternberg, Kurt: *Beiträge zur Interpretation der kritischen Ethik*, Kant-Studien Ergänzungshefte: No. 25, Würzburg: Arnulf Liebing, 1912.

Stratton-Lake, Philip: *Kant, Duty and Moral Worth*, London: Routledge, 2000.

Sullivan, Roger J.: *Immanuel Kant's Moral Theory*, Cambridge: Cambridge University Press, 1989.

——: *An Introduction to Kant's Ethics*, Cambridge: Cambridge University Press, 1994.

Tugendhat, Ernst: *Vorlesungen über Ethik*, Frankfurt am Main: Suhrkamp, 1993.

Wood, Allen W.: *Kant's Ethical Thought*, Cambridge: Cambridge University Press, 1999.

——: *Kantian Ethics*, Cambridge: Cambridge University Press, 2008.

STUDIES OF KANT'S MORAL PHILOSOPHY (COLLECTIONS)

Beck, Lewis White (ed.): *Proceedings of the Third International Kant Congress* (Rochester 1970), Dordrecht: Reidel, 1972.

Bittner, Rüdiger & K. Cramer (eds.): *Materialien zu Kants 'Kritik der praktischen Vernunft,* Frankfurt: Suhrkamp, 1975.

Chadwick, Ruth (ed.): *Immanuel Kant: Critical Assessments, Vol. III*, London, 1992.

Gerhardt, Volker, R.-P. Horstmann, & R. Schumacher: *Kant und die Berliner Aufklärung: Akten des IX. Internationalen Kant-Kongresses, Band 3*, Berlin: Walter de Gruyter, 2001.

Guyer, Paul: (ed.) *Kant's Groundwork of the Metaphysics of Morals: Critical Essays*, Oxford: Rowman & Littlefield, 1998.

Höffe, Otfried (ed.): *Grundlegung zur Metaphysik der Sitten - Ein kooperativer Kommentar*, Frankfurt am Main: Vittorio Klostermann, 1989.

Klemme, Heiner F. & M. Kuehn (eds.): *Immanuel Kant: Critical Essays, Vol. II: Practical Philosophy*, Aldershot: Ashgate, 1999.

Pasternack, Lawrence (ed.): *Immanuel Kant: Groundwork of the Metaphysic of Morals in focus*, London: Routledge, 2002.

Timmons, Mark (ed.): *Kant's Metaphysics of Morals: Interpretative Essays*, Oxford: Oxford University Press, 2002.

Wood, Allen W. (ed.): *Groundwork for the Metaphysics of Morals: Immanuel Kant*, New Haven & London: Yale University Press, 2002.

ESSAYS ON KANT'S MORAL PHILOSOPHY

Allison, Henry E.: 'Morality and freedom: Kant's reciprocity thesis', in: *Philosophical Review* 95 (1986), reprinted in: *Immanuel Kant: Critical Assessments* (Vol. III), ed. Ruth Chadwick, London, 1992, pp.68-75.

Beversluis, John: 'Kant on moral striving', Kant-Studien 65 (1974).

Baron, Marcia: 'Kantian Ethics', in: *Three Methods of Ethics: A Debate*, Oxford, 1997.

Baron, Marcia W.: 'Acting from Duty (GMS, 397-401)', in: Christoph Horn & D. Schönecker (eds.): *Groundwork for the Metaphysics of Morals* (Berlin: de Gruyter, 2006), pp. 72-92.

Broadie, Alexander & E.M. Pybus: 'Kant's concept of respect', Kant-Studien 66 (1975) pp. 58-64.

Curzer, Howard J.: 'From Duty, Moral Worth, Good Will', *Dialogue* 36 (1997), 287-322.

Dietrichson, Paul: 'What does Kant mean by 'acting from duty'?', Kant-Studien 53 (1961/62), reprinted in: *Kant – A Collection of Critical Essays*, ed. R.P. Wolff, London: Macmillan, 1967.

Funke, Gerhard: '"Achtung fürs moralische Gesetz" und Rigorismus/Impersonalismus-Problem', Akten des 4. Internationalen Kant Kongresses Mainz 1974, ed. G. Funke, Sonderheft, Kant-Studien 65 (1974).

Goyard-Fabre, Simone: 'L'exemplarité méthodologique du concept de légalité', in: Otfried Höffe (ed.): *Grundlegung zur Metaphysik der Sitten – Ein kooperativer Kommentar*, Frankfurt am Main: Vittorio Klostermann, 1989.

Hardwig, John: 'Action from duty but not in accord with duty', in: *Ethics* 93 (1983), reprinted in: *Immanuel Kant, Critical Assessments* (Vol. III), ed. Ruth Chadwick, London 1992, pp.68-75.

Henson, Richard G.: 'What Kant might have said: moral worth and the overdetermination of dutiful action', Philosophical Review 88 (1979).

Korsgaard, Christine M.: 'Motivation, metaphysics, and the value of the self: a reply to Ginsborg, Guyer, and Schneewind', Ethics 109 (10/1998), pp.49-66.

Korsgaard, Christine M.: 'Kant's analysis of obligation: the argument of Groundwork I.', in: Guyer (ed.): *Kant's Groundwork of the Metaphysics of Morals: Critical Essays*, Oxford: Rowman & Littlefield, 1998.

Laska, Peter: 'The problem of an ultimate determining ground in Kant's theory of the will', in L.W. Beck (ed.): *Proceedings of the Third International Kant Congress* (Rochester 1970), Dordrecht: Reidel, 1972.

Laska, Peter: 'Kant on moral worth: a reply to Murphy', Kant-Studien 59 (1968).

Louden, R.B.: 'Kant's virtue ethics', *Philosophy* 61 (1986), reprinted in: Klemme & Kuehn (eds.) *Immanuel Kant: Critical Essays*, Vol. II.

MacBeath, A. Murray: 'Kant on moral feeling', Kant-Sudien 64 (1973).

Marc-Wogau, Konrad: 'Amoralisch, moralisch und unmoralisch nach Kants Ethik', in: *Lehrstücke der praktischen Philosophie und Ästhetik*, ed. K. Bärthlein u. G. Wolandt, Basel/ Stuttgart 1977.

Moritz, Manfred: 'Pflicht und Moralität: Eine Antinomie in Kants Ethik', Kant-Studien 56 (1965/66)

Murphy, Jeffrie G.: "Kant's Concept of a Right Action" in: *Kant Studies Today*, ed. L.W. Beck, LaSalle 1969.

Potter, Nelson: 'The Argument of Kant's Groundwork, chapter I.', in Paul Guyer (ed.): Kant's Groundwork of the Metaphysics of Morals: Critical Essays, Oxford: Rowman & Littlefield, 1998.

Soloweiczik, R.: 'Kants Bestimmung der Moralität', Kant-Studien 5 (1901).

Sorrell, Tom: 'Kant's good will and our good nature', in: Guyer (ed.) *Kant's Groundwork of the Metaphysics of Morals: Critical Essays*, Oxford: Rowman & Littlefield, 1998.

Sverdlik, Steven: 'Kant, Nonaccidentalness and the Availability of Moral Worth', *The Journal of Ethics* 5 (2001), 293-313.

Walker, R.C.S.: 'The rational imperative: Kant against Hume', *Proceedings of the British Academy* LXXIV (1988), reprinted in: *Immanuel Kant, Critical Assessments* (Vol.III), ed. Ruth Chadwick, London 1992, pp.311-329.

OTHER SECONDARY LITERATURE ON KANT

Guyer, Paul: *Kant*, London: Routledge, 2006.

Höffe, Otfried: *Immanuel Kant*, München: C.H. Beck, 1992.

——: *Immanuel Kant*, Albany: State University of New York Press, 1994.

Kemp, John: *The Philosophy of Kant*, Oxford: Oxford University Press, 1968.

Körner, Stephan: *Kant*, London: Penguin, 1955.

Reinhold, Karl Leonhard: *Letters on the Kantian Philosophy*, ed. Karl Ameriks, transl. James Hebbeler, Cambridge: Cambridge University Press, 2005.

Wood, Allen W.: *Kant*, Oxford: Blackwell, 2005.

COMPARATIVE STUDIES

Athanassoulis, Nafsika: *Morality, Moral Luck and Responsibility*, Basingstoke: Palgrave Macmillan, 2005.

Baron, Marcia, P. Pettit & M. Slote: *Three Methods of Ethics*, Oxford: Blackwell, 1988.

Engstrom, Stephen & Jennifer Whiting (eds.): *Aristotle, Kant, and the Stoics – Rethinking Happiness and Duty*, Cambridge: Cambridge University Press, 1996.

Hursthouse, Rosalind: *On Virtue Ethics*, Oxford: Oxford University Press, 1999.

Pendlebury, Gary: *Action and Ethics in Aristotle and Hegel: Escaping the Malign Influence of Kant*, Aldershot: Ashgate, 2006.

Kerner, George C.: *Three Philosophical Moralists: Mill, Kant, Sartre*, Oxford: Clarendon Press, 1990.

Sherman, Nancy: *Making a Necessity of Virtue: Aristotle and Kant on Virtue*, Cambridge: Cambridge University Press, 1997.

ARISTOTLE'S ETHICS

Ackrill, J.L.: *Aristotle's Ethics*, London: Faber & Faber, 1973.

Allan, D.J.: *The Philosophy of Aristotle*, Oxford: Oxford University Press, 1970.

Annas, Julia: *The Morality of Happiness*, Oxford: Oxford University Press, 1993.

Aristotle: *The Complete Works of Aristotle*, revised Oxford translation, ed. Jonathan Barnes, Princeton: Princeton University Press, 1984.

——: *Nicomachean Ethics*, transl. Terence Irwin, Indianapolis/Cambridge: Hackett, 2nd ed. 1999.

——: *The Ethics of Aristotle: The Nicomachean Ethics*, transl. J.A.K. Thomson, ed. Hugh Tredennick & Jonathan Barnes, Harmondsworth: Penguin, 2nd ed. 1976.

Engberg-Pedersen, Troels: *Aristotle's Theory of Moral Insight*, Oxford: Clarendon Press, 1983

Hardie, W.F.R.: *Aristotle's Ethical Theory,* Oxford: Clarendon Press, 1980.

Hughes, Gerard J.: *Aristotle on Ethics*, London: Routledge, 2001.

Meyer, Susan Sauvé: *Aristotle on Moral Responsibility – Character and Cause*, Oxford: Blackwell, 1993.

Pakaluk, Michael: *Aristotle's Nicomachean Ethics: An Introduction*, Cambridge: Cambridge University Press, 2005.

Stewart, J.A.: *Notes on the Nicomachean Ethics of Aristotle,* vol. I, Oxford: Clarendon Press, 1892.

Warne, Christopher: *Aristotle's Nicomachean Ethics: Reader's Guide*, London: Continuum, 2006.

Lloyd, G.E.R.: *Aristotle: The Growth and Structure of his Thought*, Cambridge: Cambridge University Press, 1968.

Ross, David: *Aristotle*, London: Methuen, 5th ed. 1949.

MILL'S UTILITARIANISM

August, Eugene: *John Stuart Mill: A Mind at Large*, London: Vision Press, 1975.

Berger, Fred R.: *The Moral and Political Philosophy of John Stuart Mill*, Berkeley: University of California Press, 1984.

Capaldi, Nicholas: *John Stuart Mill: A Biography*, Cambridge: Cambridge University Press, 2004.

Harrod, R.F.: 'Utilitarianism revised', *Mind* 45 (1936), 137-156.

Lyons, David (ed.): *Mill's Utilitarianism: Critical Essays*, Oxford: Rowman & Littlefield, 1997.

McCloskey, H.J.: *John Stuart Mill: A Critical Study*, London, Macmillan, 1971.

Mill, John Stuart: *Utilitarianism,* Longmans, Green & Co., London, 13th ed. 1897.

——: *Utilitarianism*, ed. Roger Crisp, Oxford: Oxford University Press, 1998.

Ryan, Alan: *John Stuart Mill*, London: Routledge, 1974.

Smart, J.J.C.: 'An outline of a system of utilitarian ethics', in J.J.C. Smart & B. Williams: *Utilitarianism: For and Against*, Cambridge: Cambridge University Press, 1973.

Skorupski, John: *John Stuart Mill*, London: Routledge, 1989.

West, Henry R.: *An Introduction to Mill's Utilitarian Ethics*, Cambridge: Cambridge University Press, 2004.

——: *Mill's Utilitarianism: A Reader's Guide*, London: Continuum, 2007.

Williams, Bernard: 'A critique of utilitarianism', in J.J.C. Smart & B. Williams: *Utilitarianism: For and Against*, Cambridge: Cambridge University Press, 1973.

OTHER SECONDARY LITERATURE

Annas, Julia: *Platonic Ethics, Old and New*, Ithaca & London: Cornell University Press, 1999

Audi, Robert: *Moral Knowledge and Ethical Character*, Oxford: Oxford University Press, 1997.

Davidson, Donald: *Essays on Actions & Events*, Oxford: Clarendon Press, 1980.

Duncan, A.R.C.: *Practical Reason and Morality*, London: Thomas Nelson & Sons, 1957.

Findlay, J.N.: *Values and Intentions*, London: George Allen & Unwin, 1961.

Gosepath, Stefan (ed.): *Motive, Gründe, Zwecke – Theorien praktischer Rationalität*, Frankfurt am Main: Fischer, 1999.

Hume, David: *A Treatise of Human Nature*, ed. L.A. Selby-Bigge, Oxford, 2nd ed. 1978.

——: *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, ed. By L.A. Selby-Bigge, Oxford, 3rd ed. 1975.

Irwin, Terence: *Plato's Ethics*, Oxford: Oxford University Press, 1995.

Kemp, J.: *Reason, Action and Morality*, London: Routledge & Kegan Paul, 1964.

Korsgaard, Christine M. et al./ ed. Onora O'Neill: *The Sources of Normativity*, Cambridge: Cambridge University Press, 1996.

Nagel, Thomas: *The Possibility of Altruism*, Oxford: Clarendon Press, 1970.

——: 'Moral luck', in: *Mortal Questions*, Cambridge, 1979.

Plato: *The Collected Dialogues*, ed. Edith Hamilton & Huntington Cairns, Princeton: Princeton University Press, 1961.

Ramsey, Hayden: *Beyond Virtue: Integrity and Morality*, London: Macmillan, 1997.

Rist, John Michael: *Stoic Philosophy*, Cambridge: Cambridge University Press, 1969.

Sidgwick, Henry: *Outlines of the History of Ethics for English Readers*, London: Macmillan, 1931.

Stroud, Barry, *Hume*, London: Routledge & Kegan Paul, 1977.

Williams, Bernard: *Ethics and the Limits of Philosophy*, Cambridge, Massachusetts: Harvard University Press, 1985.

——: *Moral Luck*, Cambridge, 1981.