Why Kant Could not Have Been a Utilitarian

JENS TIMMERMANN
University of St Andrews

In 1993, Richard Hare argued that, contrary to received opinion, Kant *could* have been a utilitarian. In this article, I argue that Hare was wrong. Kant's theory would not have been utilitarian or consequentialist even if his practical recommendations coincided with utilitarian commands: Kant's theory of value is essentially anti-utilitarian; there is no place for rational contradiction as the source of moral imperatives in utilitarianism; Kant would reject the move to separate levels of moral thinking: first-order moral judgement makes use of the principle of morality; and, relatedly, he would resist the common utilitarian distinction between actions and their motives because any correct description of an action must refer to motivation. The article concludes with the thought that any consequentialist theory based on pre-given ends (teleology) lacks the philosophical resources to distinguish between willing something as a means and as an end, leaving means only, and destroying transparency.

The thesis that Kant was not a utilitarian is comparatively uncontentious in contemporary moral philosophy. Even Richard Hare took it for granted when little more than a decade ago he put the question whether Kant *could* have been a utilitarian to the philosophical community.[1] However, as so often, matters get much more complicated, and more contentious, when one addresses the question *why* this thesis is true. Hare argues that the philosopher Immanuel Kant (1724–1804) was not a utilitarian because, as a result of his 'rigorous puritanical upbringing' (p. 2), some of his moral judgements were incompatible with utilitarianism. He mentions Kant's views on lying, suicide and capital punishment,[2] in Hare's opinion isolated moral judgements that count for little compared to the more philosophical 'formal' side of Kantian ethics. Hare says that Kant's 'formal theory can certainly be interpreted in a way that allows him – perhaps even requires him – to be one kind of utilitarian' (p. 2). I wish to accept Hare's invitation to explore the question further (p. 2). As I am sure he would agree, the task should be considered an exercise in moral philosophy rather than mere historical

[1] 'Could Kant Have been a Utilitarian?', *Utilitas* 5 (1993), p. 2. Cf. also Peter Rohs, 'Warum Kant kein Utilitarist war', *Zum moralischen Denken*, ed. C. Fehige and G. Meggle (Frankfurt, 1995), and R. M. Hare's reply, 'Warum war Kant kein Utilitarist?', in the same collection. Plain page numbers in brackets refer to Hare's 1993 paper; references to Kant's works feature the volume number of the Prussian Academy edition in Roman numerals. All other references are to be found in the footnotes.

[2] The last item on the list is somewhat surprising. A revised Kantian theory focusing on the dignity of the individual seems in a better position to take a principled stance against the death penalty than any variant of utilitarianism.

© 2005 Cambridge University Press
Utilitas Vol. 17, No. 3, November 2005
doi:10.1017/S0953820805001639 Printed in the United Kingdom

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

scholarship, as it is likely to make us reflect on the essential aims, methods and criteria of ethical theory.

Yet I fear that this is where our agreement ends. It is certainly true that under the influence of his upbringing, his cultural environment or personal provocation[3] Kant made some rather ill-advised moral judgements incapable of being sustained even in the light of his own theory. However, I shall argue that the ethical framework that remains if one disregards such errors of judgement is still essentially anti-utilitarian. The foundations of Kant's taxonomy of value as well as his moral psychology and method of ethics are incompatible with any recognisably consequentialist ethical theory.[4] There are seemingly utilitarian thoughts in Kant's ethical writings. Hare identifies some of them, and I shall add some more. But occasional 'utilitarian' tendencies – such as a generally favourable attitude towards universal happiness – do not make a utilitarian, least of all if they are part of an unequivocally anti-consequentialist theoretical framework.

In short, R. M. Hare thought that Kant's moral philosophy might be utilitarian despite the odd non-utilitarian belief. The exact opposite is true. Kant's theory is staunchly anti-utilitarian, and generally anti-consequentialist, some subordinate arguments of utilitarian semblance notwithstanding.[5] Kant never loses sight of his basic convictions, there is not a single genuine utilitarian 'lapse'. Thus the view that Kant and utilitarianism do not go together, the very position Hare attacks, is essentially correct.[6] This result is thoroughly unspectacular. Yet I hope that re-establishing the traditional view will reveal some philosophically valuable insights.

[3] As perhaps in the case of his controversy with Constant over lying – cf. my 'The Dutiful Lie: Kantian Approaches to Moral Dilemmas', *Kant und die Berliner Aufklärung: Proceedings of the 9th International Kant Congress*, ed. V. Gerhardt, R.-P. Horstmann and R. Schumacher (Berlin and New York, 2001).

[4] In what follows, I shall confine my discussion to R. M. Hare's article, but some of the arguments also apply to David Cummiskey's *Kantian Consequentialism* (New York and Oxford, 1996), e.g. the Kantian rejection of distinguishing between levels of moral thinking, or action and motivation (cf. sections IV and V of this article). However, many other contentious claims – such as Cummiskey's consequentialist reconstruction of the second variant formulation of the categorical imperative – require careful attention far beyond the scope of this article.

[5] The issue is complicated by the fact that Hare allows his own version of utilitarianism to be softened by universalization in a semi-Kantian way. There is therefore the further question of whether Hare could have been a Kantian (or indeed a utilitarian).

[6] Whether intuitionists, Rawlsians or proponents of other rival ethical theories criticized by Hare for illegitimately claiming Kantian ancestry (p. 3) can count on Kant's support is quite another matter.

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

I. HAPPINESS AND IMPARTIALITY

No one in his sane mind would wish to deny that happiness can be a wonderful thing. Kantians deem it to be universally desired by all human beings, and worthy of general promotion. As Hare notes, in the *Doctrine of Virtue* all moral duties are discussed under the general heading of striving for 'one's own perfection and the happiness of others', which neatly sums up what for Kant a moral life is all about (VI:385 ff.).

We shall return to the duty of personal perfection later. As to the second, altruistic aspect of morality, Kant expressly commands that we adopt the ends of our fellow human beings as our own in the course of discussing the 'formula of humanity as an end in itself', the second variant formulation of the categorical imperative offered in the *Groundwork for the Metaphysics of Morals* (IV:430). The upshot of the fourth case illustrating the categorical imperative's first variant, the 'formula of universal law of nature' is essentially the same: we ought to assist others in their pursuits, we ought to make them happy (IV:423). This is the main reason why Hare thinks that Kant might have been a utilitarian: both theories appear to favour general happiness. Utilitarians and Kantians alike 'can prescribe that we should do what will conduce to satisfying people's rational preferences or wills-for-ends – ends of which happiness is the sum' (p. 4).

Is this similarity really as significant as Hare considers it to be? There is reason to be cautious. The idea that helping others can be morally laudable is hardly a distinguishing feature of Kantian or even utilitarian ethical theory. Any normative moral theory – with the possible exception of blatant egoism, if normative egoism should count as a moral theory – accords some genuine moral weight to furthering the interests of other persons. Not even a strong shared commitment to impartiality would justify Hare's drawing on the resemblance to support his main interpretative thesis. Comparing ethical theories with respect to their altruistic tendencies, we must address the questions of *why* and *to what extent* happiness is considered to be worthy of general promotion. Differences between Kant's ethical theory and utilitarianism will then quickly emerge.

As to the first question, the crucial difference between Kant's ethical theory and utilitarianism is that for Kant the sphere of morally relevant patients coincides with that of moral agents. Morality is essentially something that holds amongst equals, and in moral action we as well as others count because of the common capacity to act autonomously. Utilitarians, by contrast, include beings in the sphere of morally relevant patients by virtue of their sentience, or their capacity to be happy and to suffer. It is therefore hardly surprising that even *when*

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

utilitarians and Kantians coincide in their judgement that in a certain situation there is an obligation to help (to prevent suffering, to promote happiness etc.) their reasons are rather different. In terms of the second variant of the categorical imperative, Kant thinks that the special status of human beings as rational agents affords their projects a certain kind of weight that we cannot rationally and consistently ignore in favour of our own projects. Making other people's ends your own is a crucial aspect of Kant's command to treat them as ends in themselves, as persons, not as mere tools or means to one's own purposes. In terms of the basic 'formula of universal law', if in times of need we expect the help of others, as we surely must,[7] there is a certain inconsistency in not being prepared to help others when they happen to require your help. These argumentative strategies are hardly utilitarian in spirit.

Although the vast majority of ethical theories encourage the helping of others to some degree, they differ with regard to how impartial one's efforts ought to be. Indeed, the differences between Kantians and utilitarians become even more obvious when, secondly, we consider the extent to which they ask us to make others happy. Utilitarianism is strongly egalitarian in that no one as such deserves to be happier than any other person. Agents must act accordingly. Hare deems Kant to be in agreement with utilitarianism in this respect. He equates Kant's exposition of one's duty of love for one's neighbour (VI:450) with Bentham's injunction, that everybody is to count for one and nobody for more than one: '[I]f we treat equal preferences as of equal weight, utilitarianism is the result. But that is precisely what Kant is telling us to do in this passage' (p. 5).

However, in Kantian ethics the relation between happiness and moral goodness is more complicated than the quotation suggests. As morality revolves around rational agency rather than sentience, happiness is not considered morally good. It is the wrong kind of thing for that.[8] Moreover, happiness – though universally desired – need not

[7] There are two reasons for this. First of all, Kant thinks that you cannot just rid yourself of need at will: you *want* help. Second, the person who refuses to help others commits himself rationally to will the means of relief by virtue of his own selfishness if *he* were in need. A maxim of cold-hearted egoism cannot become a universal law.

[8] If, in contrast to utilitarian ethics, Kantians insist that it is misleading to call happiness *morally* good, this should not be dismissed as a case of terminological pedantry: it indicates three points of acute philosophical disagreement. First, for Kant happiness is not the ground of the moral goodness of actions, not even of altruistic actions. Put somewhat differently, everybody has an obligation to make others happy, but the bare fact that your action would make a person happy does not as such count in favour of your action even if your intention to make him or her happy is justified. Second, and relatedly, not all acts and attitudes that promote other people's ends are morally good or praiseworthy. Nor need all acts and attitudes that thwart the ends of others be even *pro tanto* morally bad. Other people's concerns and preferences are, to be sure, morally relevant. They enter the moral equation. But they do not as such count as 'reasons for

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

be desirable at all. Rather, happiness is good, all things considered, if and only if the happy person is morally good.[9] Happiness is not a means to moral goodness, or vice versa, but moral goodness confers objective goodness on happiness.[10] What does this mean? Kant says that we disapprove of morally evil people who are happy, because they do not deserve the happiness they enjoy. As he argues on the first page of the *Groundwork*, only morally good people are worthy of happiness:

> [A]n impartial rational spectator can take no delight in seeing the uninterrupted prosperity of a being graced with no feature of a pure and good will, so that a good will seems to constitute the indispensable condition even of worthiness to be happy. (IV:393)

action'. Third, happiness or desire satisfaction cannot be morally good because it is the wrong kind of thing to be a candidate for a distinctly moral kind of goodness. Moral theory concerns persons as rational agents, their character and behaviour. Possible candidates for moral goodness must take account of that. The utilitarian theory of value is completely independent of human agency. The satisfaction or frustration of desires can come about without human agency, perhaps even, as in the case of animals, in the absence of human agents. This does not, of course, mean that desire satisfaction and pain are morally irrelevant; but it does entail that it would be a category mistake to call them *morally* good or bad. From the first-person point of view, of course, happiness is not only morally relevant to an agent's life, it is also immediately important. Kant frequently emphasizes that everybody naturally, inescapably and quite legitimately wishes to be happy (cf. e. g. *Groundwork* IV:414 f.). Human agents are even rationally committed to furthering their own ends. Happiness is good non-instrumentally. It is good in itself, its lack of *moral* goodness notwithstanding. Echoing the ancient definition of *eudaimonia,*we might say that in Kantian ethics the value of one's own happiness does not rely on anything else by virtue of being a means to it, whereas many other things are means to happiness; but contrary to the ancient definition, it is no longer true that everything is done for the sake of happiness. Moral action is not.

[9] Note that Kant did not, as the Stoics did, identify happiness with virtue, or virtue with the only good. (On this point cf. T. H. Irwin, 'Kant's Criticisms of Eudaemonism', *Aristotle, Kant, and the Stoics. Rethinking Happiness and Duty*, ed. S. Engstrom and J. Whiting (Cambridge, 1996).) He is well aware of the fact that moral goodness may leave agents utterly miserable, and that they ought to do the moral thing even if misery is the likely outcome. Instead, Kant puts morality and happiness in lexical order, making the impersonal (rational) goodness of the one dependent on the absolute goodness of the other. However he does agree with the Stoics in one respect: preference satisfaction or pleasure is not morally good, pain is not morally bad; even though they differ as to whether happiness and pain can be at all good or bad under any circumstances. He also agrees with them in that he considers moral goodness to be infinitely preferable to pleasure, moral badness infinitely more worthy of rejection than pain. This is the gist of the example of the Stoic who has a toothache, quoted approvingly by Kant in the *Critique of Practical Reason*: 'One may always laugh at the Stoic who in the most intense pains of gout cried out: Pain, however thou tormentest me I will still never admit that thou art something evil (*kakon*, *malum*)!; nevertheless he was correct. He felt that the pain was an ill, and his cry betrayed that; but he had no cause whatever to grant that any evil attached to him because of it, for the pain did not in the least diminish the worth of his person but only the worth of his condition. A single lie of which he had been aware would have had to strike down his spirits' (V:60).

[10] Cf. e.g. Christine Korsgaard's well-known article 'Two distinctions in goodness', *Creating the Kingdom of Ends* (Cambridge, 1996), and Thomas E. Hill Jr., 'Happiness and Human Flourishing', *Human Welfare and Moral Worth* (Oxford 2002), pp. 169–73.

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

The result is that different persons may differ with regard to their desert, depending on whether they are morally good agents or not. To choose a comparatively uncontroversial example, if you face the choice of helping either Andrew or Beatrice and Andrew is a thoroughly decent person whereas Beatrice is not, it would quite plausibly be a good thing for you to assist Andrew in his pursuits whereas it would not be such a good idea for you to help Beatrice. This is not because Beatrice's project is morally bad or irrational, nor because making Beatrice happy would in the long run help her to make others miserable. Beatrice simply does not deserve your assistance.[11] Andrew counts for one, Beatrice does not. The thought that one cannot universally will that bad people be undeservedly happy is an intuition that utilitarian ethical theory will find difficult to accommodate.[12]

Yet Kantian partiality is not even confined to potential beneficiaries that differ with respect to the moral worthiness of their projects or character. Amongst persons of equal moral standing there may yet be meaningful distinctions that justify differential treatment. Take Andrew and Charlotte, both of which are, to the best of your knowledge, equally decent and well-intentioned Kantian agents. Charlotte, however, did you a great favour two years ago. Should you be impartial in your help, provided you can help only one of the two? Of course not, despite their equal moral status. You *owe* Charlotte a favour, as a matter of gratitude. Again, helping Charlotte may not maximize overall happiness, and even if it does, *that* is not the reason why you ought to help her. Charlotte deserves your help more than Andrew because she has entered a special relationship with you. Similarly, Kant says at one point that the general duty of practical love must be made more specific by certain rules, such as the rule that you ought to honour your parents.[13] Complete impartiality would perhaps be appropriate in a world of innocent equals, of persons without a history, without fault or merit. Unlike Kant, utilitarianism seems to ignore the fact that we do not live in such a world.

[11] Cf. the Methodology of the *Doctrine of Virtue*, VI:481. The teacher suggests to his pupil that one ought not to give soft cushions to a 'lazy fellow', make sure that a 'drunkard is never short of wine', 'give a violent man audacity and strong fists so that he could crush other people', etc.

[12] At the same time, the person who actually deserves happiness may not in fact be happy. The world Kant lived in was not just in its distribution of happiness. Nor is ours. Thus the ancient tie between moral excellence and the happy life of the agent is finally severed. Even the morally good person cannot reasonably expect to be happy, but he can hope to be worthy of happiness. In other words the goodness of happiness is conditional upon the moral character of the agent.

[13] *Metaphysics of Morals*, VI:390. Interestingly, he also suggests that duties towards one's parents can be reduced to duties of gratitude. The mere fact of owing one's existence to someone is said not to carry any kind of obligation with it because it is the condition of all obligation.

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

Kant's ethical theory is therefore much less impartial than utilitarianism; or rather it provides an immediate justification for some kinds of partiality. Relatedly, there is no suggestion that we ought to aim at mathematical maximization of happiness. This does not necessarily mean that Kantian altruism is any less demanding than utilitarianism, although it is difficult to determine exactly how demanding 'practical love' should be. Kantian moral agents most certainly have to be seriously committed to furthering other people's welfare if and when appropriate;[14] but beyond that I fear that, in a spirit close to virtue ethics, Kant would merely suggest that as beneficence is a wide or imperfect duty there are no hard and fast rules, and that a good person will make the right judgement.[15]

This, in rough outline, is the core of Kant's taxonomy of value; and there is not a jot of utilitarianism in it. Even where the two theories coincide in their judgement as to what we ought to do, the difference is not about *whether* to intend the consequences of an action, but about *why* we ought to intend them. Kant's theory would thus be anti-utilitarian even if his moral recommendations coincided with those of utilitarianism. Utilitarianism is not just the theory that tells us to maximize happiness, but rather the theory that recommends maximization of happiness *because happiness is the good*. Kantians deny this, no matter what shape their first-order commands take.[16]

Kant's notion of desert is particularly un-utilitarian. Utilitarians cannot accord any direct ethical weight to the moral standing of the 'recipient' of happiness. It cannot – so runs the familiar complaint – as

[14] Like the Humean virtue of justice, the extent to which one ought to help others is bound to depend on how needy people are. There is some indication that Kantian morality is pretty demanding when it comes to helping others. One of the 'casuistical questions' concluding the section on beneficence in the *Metaphysics of Morals* reads as follows: 'How far should one expend one's resources in practising beneficence?' The answer, if vague, is striking: 'Surely not to the extent that one would oneself finally come to need the beneficence of others' (VI:454). Going further would violate the 'indirect' duty not to neglect one's own well-being (cf. IV:399, VI:388) and also violate the spirit of universalizability.

[15] Once one has adopted the end of making others happy, there is an 'indirect' duty to take the means to realize this end that does not, however, as an 'indirect' duty, have any direct moral weight. As with other ends, one needs to be acquainted with the requisite rules of skill (technical imperatives), i.e. one must do one's duty skilfully. 'Utilitarian' considerations may enter at this subordinate level; but the duty of beneficence itself is grounded in entirely non-utilitarian considerations.

[16] Cf. the Kantian rejection of levels of moral thought below (Section IV). Another way of making this point is that any theory taking the goodness of states of affairs to be promoted by action as basic would, according to Kant, be heteronomous. It would no longer be the person's rational self that lays down the standards of ethics, no matter what these standards command at the first-order level. First-order commands are secondary when the status of an ethical theory is concerned. Even a system of strict Kantian 'deontology' would bear the stigma of heteronomy if divine fiat, rather than rational autonomy, turned out to be the source of the normativity of such moral commands.

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

such matter whether you make a hardened criminal happy or a moral saint, or whether you pay due attention to debts of gratitude. It may matter indirectly because of the effect the happiness of others has on the well-being of people in general; but for the Kantian, that is not good enough.

II. PERFECT AND IMPERFECT DUTY

One reason in particular accounts for the fact that Hare overlooks so many important differences between Kantian moral philosophy and utilitarian ethical theories: his strong focus on the *Groundwork*'s fourth illustration, which explains why we have a duty to further other people's ends. Of all the examples that illustrate the workings of the categorical imperative, it is admittedly closest to a utilitarian injunction to make everyone happy, in letter if not in spirit: the duty still rests on considerations of dignity, rationality and autonomy, not on the goodness of achieving pleasure, happiness or desire satisfaction. But there are three more categories of duty, for instance perfect duties to others, a category that Hare almost completely neglects. This is unfortunate: a closer examination of how perfect duty is generated by the categorical imperative would have revealed a most anti-utilitarian tendency in Kantian ethics. There are two reasons for this.

First of all, the strict Kantian duty not to make false promises is not to be construed along consequentialist lines. Hare suggests that in the case of a false promise Kant 'uses a form of argument usually now called utilitarian generalization' and glosses this strategy as follows: Kant (it is said) asks ' "How would things stand if my maxim became a universal law?" and answers that promises would become "empty shams" ' (p. 7). Hare immediately criticizes this method as not very strong 'because one might will as a universal law that people should break promises in precisely one's own present situation, when one can get away with it and the institution of promising would survive'.[17] But the argument is stronger than Hare thinks.

Kant does not employ a strategy called 'utilitarian generalization', and this strategy is not equivalent to the categorical imperative. Crucially, the 'formula of universal law' does not turn on the mere desirability of the consequences of one's actions when employed to

[17] This is not an accurate description of the agent's purpose – he simply wishes to make a false promise and get away with it. His maxim, to be universalized, is not the specific one of 'making a false promise in this or that situation', but rather the one cited by Kant: to extract himself out of a difficulty by making a false promise. Any other situation would have prompted a similar *ad hoc* modification. The agent who considers himself to be acting on the more specific, allegedly universalizable maxim is deceiving himself about his true principles of action. More about action descriptions in Kantian ethical theory and in consequentialist theories in Section V on 'Motives and Consequences' below.

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

generate perfect duties and is therefore not subject to Hare's objection. The question of 'how things would stand' is only an intermediate step. This kind of consideration is not meant to justify a maxim. Instead, all depends on whether the agent could still act on the personal principle he is tempted to act on – to make a false promise when convenient – if his own attitude were shared by everyone else. (This personal principle or attitude is called a 'maxim'.) In the case of false promises the answer must be that he cannot. If the attitude he is tempted to act on were universal – practised everywhere on earth and at all times in the history of humanity – his opposite number would not trust him. He could not achieve the very thing he wants to achieve with his lying promise. There is a 'contradiction in conception', as it has been called (cf. *Groundwork*, IV:421 f.); and it is this kind of contradiction that grounds perfect duties to others.[18] The idea that rational contradiction is the ground of moral badness is utterly alien to utilitarianism and any other brand of consequentialism. Again, Kantian ethics would not be utilitarian even if the two theories were practically co-extensive in their first-order recommendations.[19]

Second, there is the idea that the imperfect duty to help other human beings in their pursuits is restricted by perfect duties such as the duty not to make false promises. If in order to help someone you have to steal, murder or lie, helping the person in question is not morally good. Note that we reach this conclusion not by weighing reasons, or pros and cons. The fact that it would enable you to help someone is not to any extent a redeeming feature of an immoral act. A morally bad action is not good in one respect and bad in another. An action contrary to the *categorical* imperative is straightforwardly and unambiguously bad. I fear that the primacy of strict duty over wide duty is something utilitarians will not be happy with, philosophically, not only practically. Given that Kant has a philosophical argument for the primacy thesis[20]

[18] This argument (i) goes well beyond the thought that (ii) you would not wish to be subject to an action or attitude if it was universal, a thought recognized by Hare in *Freedom and Reason* (Oxford, 1963); which in turn goes beyond the idea that (iii) the world would not be nice, or that people might be unhappy, if everybody acted in a certain way or shared a certain attitude. Only the third strategy is genuinely utilitarian.

[19] Cf. Elizabeth Pybus, 'Utilitarianism: A Basic Flaw?', *Philosophy* 55 (1989), p. 554. The 'flaw' detected by Pybus consists in the assumption that if it is good to further happiness it must therefore be good *because* it is good to further happiness, i.e. it must be good for utilitarian reasons. Those who seek to interpret Kant in a utilitarian spirit appear to fall prey to this fallacy.

[20] As I argue in *Immanuel Kant: Grundlegung zur Metaphysik der Sitten* (Göttingen, 2004), p. 122, strict duty always overrides and silences apparent wide duty because it involves *both* a 'contradiction in conception' as well as a 'contradiction in the will', whereas wide duty involves merely a contradiction in the will.

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

it would be a dubious move to dismiss it as an error of judgement, or blame it on his 'inbred rigorism'.[21]

There is one constraint on impartiality that even Hare acknowledges to be Kantian. Commenting on a quotation from the *Metaphysics of Morals*,[22] Hare grants that 'immoral or anti-social ends' are not worthy of consideration for general promotion, and likens Kant's position to that of Harsanyi.[23] Kant might, Hare thinks, be said to be a 'rational will utilitarian' (p. 4). However, this concession is still inadequate. What has gone wrong?

First of all, there is the familiar idea that whether an agent deserves happiness, and hence our help, does not just depend on his present projects, but on his moral merit as a whole. This goes well beyond Harsanyi's proposal. Second, the 'immoral and anti-social ends' that we are to rule out in our deliberations on whose preferences to satisfy would presumably be spelled out in consequentialist terms;[24] but in that case, Kant would not agree, as his criteria for immoral ends that do not deserve support are not consequentialist. It should also be noted that it is Hare who glosses Kant's original 'immoral' (*unsittlich*) with the term 'anti-social', which considerably narrows the scope of illegitimate ends. Duties to the self, for instance, are excluded.

III. MORAL PERFECTION

Practical love is important, but for Kant and other traditional moral philosophers altruism is but one dimension of morality. There are duties to the self as well as duties to others. In other, less narrowly Kantian words: there are genuinely ethical norms solely concerning the agent's own life. Kant and Aristotle would be on the same side of the philosophical divide,[25] utilitarianism on the other.[26] In a much-disputed passage, Kant even says that duties to the self are the foundation of

[21] In addition, a closer investigation of these matters would have shown that Kant's argument against false promises – never mind the primacy of strict duty – is hardly utilitarian in spirit. The reason why I should not lie to a person about the chances of my repaying a sum of money to obtain a loan is not that this would make him unhappy or be contrary to his preferences and ends. It would be wrong even if he desired to be deceived because this desire is in itself immoral. We shall return to this line of argument. Few utilitarians will find it congenial.

[22] 'The duty of love for one's neighbour can . . . also be expressed as the duty of making others' *ends* my own (provided only that these are not immoral)' (VI:450).

[23] John C. Harsanyi, 'Problems with Act-Utilitarianism and with Malevolent Preferences', *Hare and Critics*, ed. D. Seanor and N. Fotion (Oxford, 1988).

[24] Cf. Bernard Williams, *Ethics and the Limits of Philosophy* (London, 1985), p. 87.

[25] The sphere of the moral includes the classical 'moral' virtue of moderation and the development of one's talents, not for prudential reasons but rather because of rational constraints that – as in the case of conventional universalization – go beyond prudence.

[26] Even though utilitarians from J. S. Mill onwards have either been critical of duties to the self or ignored them altogether, the only reason why utilitarianism as an ethical

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

all duty, and prior to duties to other persons. Without duties to the self, he (enigmatically) argues, there would be no duties at all (VI:417–18). Yet Hare obviously ranks duties to the self among Kant's embarrassing errors of judgement. In his view, in the first and third illustrations of the categorical imperative, the prohibition of self-murder and laziness, Kant 'is going astray through trying (in order to buttress his inbred convictions) to use arguments from universalizability outside their proper field, which is duties to other people' (p. 6).

It is obvious that Hare regards the first and third examples as cases of universalization. But duties to the self do not rest on universalization in the run-of-the-mill sense of modern ethical theory. They do not involve persons other than the agent. Nor does the application of the categorical imperative that generates them. The duty to develop one's talents does not turn on the thought that the world would be a horrible place if everyone neglected the development of his or her talents in like manner, but rather on the thought that there is a certain immoral inconsistency in an attitude that for the sake of present pleasure undermines the realization of one's rational ends and aims. Without going into the intricate details of how the formula of universal law yields duties to the self we should note that what is at stake is solely the possible timelessness of one's principles of action. We have already seen that timelessness is an aspect of the categorical imperative that plays a role even in the case of duties to others. As far as duties to the self are concerned, it is the only relevant consideration. Also, the second variant formulation asks you to use neither yourself nor others as mere means – which Kant does not make use of universalization in any sense. This suggests that the Kantian notion of a duty to the self is more robust than Hare would allow. Removing them from the Kantian system leaves a large gap, not least because they are meant to be fundamental to morality as a whole.

There are many ways in which duties to the self can said to be 'prior' to other duties, but two of them in particular highlight the un-utilitarian nature of Kantian ethics. First, let us return to the second variant of the categorical imperative, the 'formula of humanity as an end in itself'. It is the criterion of strict or 'necessary' duty that we are not supposed to use either ourselves or others as *mere* means to the ends we happen to pursue: as *things*. One of the more pragmatic methods of testing the permissibility of one's proposed course of action with regard to other people is to test whether the other person can 'possibly agree to my way of behaving towards him', whether he can 'himself contain the end of this action' (IV:430). Crucially, this kind of consent

theory cannot accommodate duties to the self is that as its end is maximum happiness it is arguably at odds with the idea of specific duties to *any* individual, oneself or others.

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

is not to be construed along the lines of actual consent. The question is whether the other person can *possibly* consent; and to answer this question, he has to take the duties towards his own person into account. There is much more to be said about this, but I cannot delve into the depths of moral metaphysics here. We should, however, note that Kant's theory of duties to the self might offer a way out of the dilemma of the 'fanatic' that plagued Hare in *Freedom and Reason*.[27] There are rational first-personal constraints beyond prudence, construed as the art of best achieving one's actual ends. An agent's assent to an action can be morally wrong even if no one's life is concerned but his own. (Note also that the familiar objection that the idea of a duty to the self makes no sense because the agent could always release himself from such an obligation is question begging. It presupposes *that there are no duties to the self*.)

Second, moral perfection consists subjectively in the 'purity of one's disposition to duty' and objectively in 'fulfilling one's duties to oneself' (VI:446). This definition from the *Metaphysics of Morals* would be utterly puzzling if moral perfection consisted largely in practical love. It is still puzzling if it is supposed to consist in practical love to some extent. Yet the definition points to a shift of focus Kant asks us to make from a purely ethical point of view. If you look at ethics from within, from the point of view of the agent, his principles, motives, etc., *all* duties also involve duties to the self. The reason is that even in the case of duties to others such as the command not to lie to them or to help them, there is a duty to *oneself* to go beyond the mere performance of the act. The act itself is all the other person can demand (perfect duty) or expect (imperfect duty), to do the right thing for the right reason is something you owe to yourself. This idea is deeply un-utilitarian.[28]

In conclusion, there is no danger that moral perfection is 'chasing its own tail' (p. 9) in trying to be anything over and above practical love, as Hare contends. It is true that perfection, the kind of perfection we are asked to pursue in addition to furthering the happiness of others, is moral perfection, which, as Hare states, consists in the acquisition of virtue. It is also true that 'part of this virtue will consist in the disposition to fulfil the duties to others laid down on the utilitarian side of the table' (p. 9). Hare then wonders what the other part could be, and

[27] Ironically, the strict utilitarian who wishes to maximize utility no matter what the consequences for his own life would be seems dangerously close to Hare's fanatic. A utilitarian who starts taking considerations of universality into account is already on his way to Kantianism. The question arises whether – occupying the middle ground – Hare could have been a Kantian; or indeed a utilitarian.

[28] We shall return to the topic of moral motivation in section V below.

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

has difficulties coming up with a convincing answer. Yet the sphere of the moral by no means coincides with altruism or 'practical love'.[29]

It should lastly be noted that the concept of moral perfection makes Kant's ideas about human flourishing, as we may (again) call it today, much more interesting. His technical conception of 'happiness' (*Glückseligkeit*) is for the most part quite impoverished. It is defined in experientialist terms, as the fulfilment of the sum-total of one's desires, the 'complete well-being and satisfaction with one's condition' (*Groundwork*, IV:393).[30] Kant occasionally sounds less subjectivist than this might suggest. He then seems to rely on human nature to fix and rank the desires we have. Yet with even less faith in natural teleology than Kant, those who wish to take more of an 'objective list' approach would be well advised to turn to moral duties to the self for a richer conception of human well-being.[31]

[29] *Pace* Hare, moral perfection is not an empty ideal, which is precisely the criticism Kant levels against the notion as employed by his perfectionist contemporaries. As Hare notes, Kant says that the ontological concept of perfection 'shows an inevitable tendency to go round in a circle and is unable to avoid covertly presupposing the morality it has to explain' in the *Groundwork* (IV:443) and elsewhere. Indeed if an *ontological* concept of perfection is considered to be *philosophically fundamental*, a vicious circle is bound to arise. Hare is entirely right: 'There would be nothing *else* in the duty to make ourselves perfect, except the duty to make ourselves disposed to make ourselves perfect. It would still not have been determined what the perfection, or the performance of the duty to promote it, would consist in' (p. 9). There is no danger of this in Kant, who does not try to derive concrete moral duties from the ideal of perfection. Rather, Kant first defines moral goodness in terms of strict and wide duties to the self as well as to others and *then* states that perfection consists in conforming to this independent standard.

[30] There is also an 'indirect' duty to care for one's own happiness (*Groundwork*, IV:399). Strictly speaking an indirect duty – as opposed to a direct strict or wide duty – is not a duty at all. An indirect duty is somewhat utilitarian in spirit. Kant thinks that we have an indirect obligation to further our own happiness. That we do not have a direct duty seems clear because of the fact that we care for our happiness anyway; and there cannot be a duty to do what one wants to do as a matter of course. That's why perfect beings (God?) do not have duties – they follow the moral law as a matter of course; a divergence or conflict between their very own ends and moral requirement is a priori impossible. However, human beings are not perfectly rational, not about morality and not about their own lives, so there is room for an indirect duty to care about one's happiness, it seems. The reason why happiness, in this case, is morally relevant is not that it is just one's own happiness. That would be selfish, and hence not a moral consideration at all. The idea seems to be, rather, that people who do not take care of their own happiness are likely to get into all sorts of difficulties and conflicts which make it more likely that they will overstep the boundaries of what is morally right. If you are poor, you will be tempted – perhaps even psychologically compelled – to steal, etc. How far-reaching this indirect duty is, remains unclear. It is an indirect duty because furthering your happiness is no more than a means to keeping you away from moral danger. The direct duty is that of moral care.

[31] Cf. Thomas Hill, 'Happiness and Human Flourishing', pp. 195–200, for some possible explanations of the fact that Kant takes so little interest in the ancient concept of 'human flourishing'. It seems most likely to me that he was simply not very aware of it because he assumed that Aristotle and the Stoics were talking about 'happiness' in the Kantian

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

IV. LEVELS OF MORAL THINKING

Utilitarians like Hare might object that utilitarian ethical theory can account for our intuitions about desert, special relationships and even duties to the self with recourse to separate levels of ethical thought.[32] Distinguishing such levels is not an essential feature of utilitarianism, but confining utilitarian principles largely to the level of reflection is certainly not uncommon. Moreover, as Kant's first-order precepts are hardly utilitarian in spirit, those who wish to argue that he *could* have been a utilitarian naturally watch out for utilitarian principles at a higher-order philosophical level. *If* Kant believed in separate levels of moral thinking, Hare's claim that he could have been some sort of utilitarian might still be saved. Kantian ethics may turn into 'utilitarianism at one remove' (p. 10). However, there is overwhelming evidence that Kant would vehemently resist such a division.

Kant leaves little doubt that the categorical imperative is immediately operative in ordinary first-order deliberation. After the categorical imperative has been stated and illustrated for the very first time – and not yet under this technical name[33] – in *Groundwork* I, Kant says that:

> it would be easy to show how common human reason, with this compass in hand, 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, if, without in the least teaching it anything new, we only, as did Socrates, make it attentive to its own principle. (IV:404)

Kant makes two related claims. First, he declares that the principle just discovered can easily be put into practice by everyone. You do not have to be a moral philosopher to be moral. He sometimes even suggests that philosophical sophistication is detrimental to sound moral judgement. Second, Kant affirms that the categorical imperative is at work in ordinary moral judgement. It is the task of the 'analytic' part of moral philosophy, conducted in the first two sections of the *Groundwork*, to reveal the a priori principle of judgement that is implicitly operative

sense. Some elements of *eudaimonia* are subsumed under imperfect duties to the self, thus making the flourishing of a rational creature an essentially moral matter.

[32] Cf. Hare, *Moral Thinking: Its Levels, Method and Point* (Oxford, 1981).

[33] The language of technical, pragmatic and categorical imperatives is introduced only in the more philosophical second section, cf. IV:412 ff.

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

in everyone's moral thinking,[34] that of ordinary people (section I) and Kant's rivals in moral philosophy alike (section II).

In the *Critique of Practical Reason*, he gladly grants his critics that he has not established a 'new principle' but only a 'new formula'. He ridicules those who seek to establish completely new moral guidelines:

> But who would even want to introduce a new principle of all morality and, as it were, first invent it? Just as if, before him, the world had been ignorant of what duty is or in thoroughgoing error about it. (V:8, fn.)

Philosophy must precisely identify the principle implicit in everyday judgement; and then find suitable ways of teaching it, of making its representation more clear and vivid in people's minds.

There are thus no split levels of moral thinking in Kantian ethics. On the contrary, Kant would wholeheartedly endorse Bernard Williams's criticism, cited by Hare, 'that by dividing moral thinking into two levels' one demotes 'our ordinary intuitive convictions and prima facie principles into a merely instrumental role' (p. 10).[35] Kant wants persons to do the right thing for the right reason. He wants all agents to be in immediate contact with moral value in each and every moral

[34] The *Groundwork* begins with what Kant takes to be the common answer to the ancient question of the *summum bonum*: the greatest and only unconditional good is a good will. A note on Kant's ethical method. Hare is right in emphasizing that Kant was not content with ordinary moral intuition or judgement (pp. 2–3) but rather wanted to get to the bottom of it, to discover the 'metaphysical' principle implicit in ordinary judgement which might then be used as a more reliable guideline; but Hare misconstrues the nature of judgement as well as Kant's attitude towards it. We have seen already that ordinary moral judgement is not very far removed from the philosophical principle of ethics; and even though Kant would have 'rightly excluded from this part of his enquiry any empirical data, whether about what actually goes on in people's minds or about anything else', this is true only if the data are truly empirical: 'antecedently held substantial moral judgements' (p. 12), if examined from the proper perspective of moral philosophy, are not purely empirical. The basis is a priori. It is not true that 'the only source of these could be something that goes on in people's minds, that is, intuitions' if intuitions are taken to be empirical facts. As the foundation of moral judgement is rational, Kant rightly rejects moral sense theories as empirical, but a rejection of common sense ethics does not entail a rejection of ordinary moral judgement or intuition. The very first paragraph of the second section of the *Groundwork* explicitly rejects the allegation that an analysis of commonly held moral views is an empirical investigation (IV:406). Ordinary people apprehend much more than the concepts of morality. They have an a priori grasp, however clouded by self-interest, of the a priori principle of ethics. It is one thing to reject a theory as an adequate theory of morality, and quite another thing to reject 'data' which may or may not be 'empirical'. Moral judgements need not be empirical if moral sense theories are rejected. Kant even tries to tie his ethical theory back to ordinary intuition in the 'deduction' of morality at the highly abstract level of moral thought exhibited in the third section of the *Groundwork*.

[35] Bernard Williams, 'The Structure of Hare's Theory', *Hare and Critics*, pp. 189–93. Williams proposes genealogy as a non-reductive, non-instrumentalist method of revealing the value of our moral practices in his *Truth and Truthfulness* (Princeton, 2002), esp. pp. 90–100.

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

judgement, which is incompatible with splitting levels of moral thought. Transparency is vital for Kantian autonomy.

It is an interesting question why some ethical theories are less at odds with our common pre-philosophical moral views and values than others. When the proverbial woman in the street learns about classical utilitarian value theory she may well be disconcerted; but when she is told an Aristotelian story about a flourishing human life of excellent activity in accordance with its characteristic 'function', she is more likely to be reassured. Kantian moral philosophy is similarly reassuring, and similarly unlikely to erode commonly held ethical convictions.

What makes such reactions adequate, if indeed they are? The answer should presumably be construed along the following lines. Unlike moral commands in indirect versions of utilitarianism, the categorical imperative is a guide to moral action that reveals the source of moral goodness. As a result of its 'formal' character, it is *not* reductive but what perhaps should be called 'abstractive'. It does not presuppose morally valuable ends of action, as the principle of utility does, but rather helps the agent to choose ends in the light of his own rational principles. If the 'material' principle of utility were the correct principle of morality at any level, one's actions would be no more than a *means* to the pre-given end of universal happiness. They could not in themselves be morally valuable; and, the Kantian argument continues, their role could easily be taken over by other factors, even by chance. None of this is true even of the duty that admittedly resembles utilitarianism, Kantian 'practical love'. The categorical imperative as a purely formal principle explains why it is *part* of moral goodness, why it is in itself morally good.

Another way of putting the reason why Kant would have rejected Hare's division of levels of moral thought is that, for Kant, the personal principles are of primary moral importance, whereas the consequences take second place at best. The two distinct levels of moral thought suggest consequentialism, if not vice versa. Why was Kant not a consequentialist?

V. MOTIVES AND CONSEQUENCES

The question of motivation divides Kantians and utilitarians; and what is more, it marks one of the defining controversies between consequentialism on the one hand and non-consequentialist moral theories, such as virtue theory and Kantian ethics, on the other. Aristotelians and Kantians agree that for an action to be morally good it must be done with moral goodness firmly in mind – even though they would find it hard to agree on an account of what distinguishes

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

fine human actions from base ones. Kant would say that that a morally good action must be performed for the sake of duty rather than just in accordance with it. Utilitarians would reject this kind of motivational rigorism and emphasize the importance of actual or expected consequences for sentient beings instead. This is yet another reason why even in the light of – as Hare calls it – his 'formal' theory Kant could not have been a utilitarian. A rigorist theory of moral motivation is not an additional feature that adherents of the categorical imperative as the supreme principle of morality can take or leave. It is one of Kant's most important insights.[36]

This is revealed by the close connection between Kant's rebuttal of levels of moral thinking and his insistence that an agent must act from a sense of duty, for the sake of the action itself, if his action is to be morally good. He insists that the source of value – agreement with reason – must the basis of concrete first-order moral judgements and that this agreement must motivate the agent if his action is to possess moral worth. The categorical imperative is the test or criterion, whereas the contradiction – or, rather, the lack thereof – is the ground of moral goodness. Utilitarians try to distinguish between the rightness of the act and the goodness of the agent, taking ethics to be concerned primarily with the former. Kant would reject this very distinction: the latter is to be the *ground* of the former. The action is morally good, rather than in mere conformity with duty, if and only if it springs from the moral goodness of the agent, for which there are formal moral standards – in other words, if and only if the action is appropriately motivated.[37]

[36] The categorical nature of the moral imperative entails motivational rigorism, and vice versa. To establish this is the project of the first section of the *Groundwork for the Metaphysics of Morals*.

[37] One motive that is certainly improper by Kantian moral standards is self-interest. One should not do an act conforming to duty because of any advantage that it may bring; and unlike his ancient predecessors, Kant thinks that it is hopeless to pretend that self-interest and morality can be brought to coincide. Moral action does not always pay, and even when it does, that is not what makes it special. So, unlike psychological hedonism, Kantian moral psychology firmly denies that we act morally to feel good about ourselves. However, it does not follow that we do not do so. Kant does in fact think that acting morally makes us feel good, but he asserts that this feeling is not the motive of morally worthy action. It is just a product of it. You act morally because you see that a certain thing is the appropriate thing to do, the thing required. No more than the reason-generated motive of respect is required to move you, and respect does not proceed from your motivational set prior to your moral judgement, from the ends you are naturally committed to. Nonetheless, being conscious of having done the deed for the right kind of reason produces a feeling of moral contentment which is an essential part of a good human life, even though it is by no means sufficient. Presumably, and paradoxically, this feeling could not be a motive to begin with. 'Doing your duty' to feel good about it is not a possible motive: actions done for morally inadequate motives will not lead to moral contentment. It occurs if and only if the right thing is done for the right reason.

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

Let us have a look at a simple example from both the Kantian and the utilitarian perspectives (the slightly distasteful kind of example moral philosophers delight in). Beatrice notices a child drowning in the pond that she can easily save. She is in a great hurry. Yet it occurs to her that saving the child would get her picture in the papers, and that she would probably receive some kind of reward. She is attracted by this prospect and saves the child with the expectation of fame and public recognition in mind. In Kantian terminology, she acts *in accordance* with duty but not *from* duty, nor indeed from any other motive that ethical theory would deem to be morally good.

It is important to identify the precise point at which Kantians and utilitarians disagree – there are surprisingly many points of agreement. Both agree that other things being equal Beatrice morally ought to save the child; both agree that she is not a morally virtuous person; and they both agree that saving the child for immoral reasons is in some sense preferable to letting him drown. (It is certainly preferable from the child's perspective, or that of the child's parents.) The difference seems to be the following. Utilitarians would say that the action is right or morally good, even though the agent is not, whereas Kantians would reject this distinction as philosophically indefensible: the action was admittedly in accordance with duty but, like the agent, not morally good.

The utilitarian distinction between actions and motives is intuitively plausible. Is there any good reason for Kantians to reject it? One obvious thought in their favour is that any kind of motivation other than the thought of moral appropriateness makes the action's coinciding with duty an accidental feature of the situation. For Kant, moral action must come about *reliably* as a consequence of a person's principles for his action to have moral worth, and there are two ways in which things can go awry if they accord undue weight to inclination. First, there is the familiar question whether a particular non-moral incentive reliably produces actions according with morality. Sympathy often points in the direction of what is morally good; but the two do not necessarily coincide. It is quite possible to assist those who do not deserve it merely from sympathy. There is, secondly, the reliability of the agent's maxims. The agent must, as a matter of personal principle, reliably be moved to act on the right motive. In the first case, it is strictly speaking the situation that is unreliable. Only doing the right thing for the right reason will reliably lead to action in accordance with morality. In our example, Beatrice saves the child for the wrong reasons, and it was because of the circumstances that she saved the child. The circumstances suggested saving the child to her – but Kant and common sense would suggest that we judge actions for what they are, not for the accidental circumstances they occur in, or indeed the consequences they (are expected to) have. The child was simply morally lucky that

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

Beatrice's craving for public honour counted in favour of her saving him.

There is another philosophical lesson to be learnt. Notoriously, actions can be described in many different ways.[38] What Beatrice does can be described variously as her saving the child's life, as her pulling him out of the water, as her satisfying her vain desires, as her getting her feet wet, and so on. However, as far as moral judgement is concerned, some descriptions will be more relevant than others. Kantian ethics provides us with the philosophical resources to solve the problem of action description because it judges them on the strength of their source or origin, rather than their consequences. The fact that Beatrice would have to get her feet wet is no more than an accidental feature of the situation. It is less than a means, it is something she knowingly, if grudgingly, lets happen. Stepping into the water, on the other hand, should probably count as a *means* to saving the child. Equally, as the example is set up, saving the child is a mere means to public fame and recognition: you have to step into the water to save the child, getting your feet wet; you have to save the child to get your picture in the papers; and that is what, in the end, she is after. (It is probably both a part of and a means to personal fame.) Everything else is at best an accidental means to the end. Kant would therefore argue that the one and only morally adequate description of her action[39] is that she is deliberately doing whatever she does for personal gain, the end of which fame is a part. But there is nothing morally laudable about *that*. What seems to be morally laudable is in fact only a means, or even a mere side effect of a selfish action. Only the person who does the act that accords with morality *because* he or she regards it to be the morally right thing to do actually *does* what he or she ought to do. In Kantian terms, only actions done from a proper sense of duty possess moral worth. The agent's attitude and the action that follows cannot be separated.[40]

[38] Cf. G. E. M. Anscombe, 'Modern Moral Philosophy', *Virtue Ethics*, ed. R. Crisp and M. Slote (Oxford, 1997), p. 28. The paper was first published in 1958.

[39] The Kantian notion of an action also helps to counter the allegation that there is always a description under which an act can be universalized: the relevant item to be universalized is the agent's maxim, which always includes the end or attitude that motivates the action. All other descriptions are irrelevant. The idea that the problem of relevant descriptions can be countered with recourse to Kant's theory of maxims is central to Onora Nell's (O'Neill's) *Acting on Principle: An Essay on Kantian Ethics* (New York, 1975).

[40] If one accepts motivational rigorism, there is still the question of how to formulate the reason of an action done 'from duty', a category of moral philosophy rather than ordinary discourse. A Kantian visiting a friend in hospital would be unlikely to tell her that he is there to discharge his duty. (Nor, for that matter, is a sentimentalist likely to say that he visits her owing to a feeling of sympathy.) If asked, the Kantian would – truthfully – tell his friend that he has come because she is his friend, because he thought she needed

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

This, it seems to me, is a convincing argument against the consequentialist distinction between agents and their actions.[41] There is a deeper reason for this. Because in consequentialist theories like utilitarianism the ends of good action are heteronomously determined by a teleological theory of value,[42] utilitarianism lacks the philosophical resources to distinguish between willing an action as a mere means and as an end. For Kantians, who argue that one ought to choose the ends of moral action on the basis of purely formal considerations, this distinction *morally* makes all the difference; whereas for utilitarians 'morally good' action is no more than a means to promoting a pre-given end: happiness, however it may be spelled out in detail.

his company or the like. The reasons for considering her needs morally relevant would of course be Kantian. If pressed for a further answer, he might say that he thought coming to visit her was simply the right thing to do, in technical terms: his duty.

[41] It may still be the *intention* of the consequences, or the ends pursued. To show that ends are not all there is to ethics because we need *formal* constraints is much more complicated; but it can be done. There is an additional layer of problems if the act is regarded as a means to maximization.

[42] Two (related) remarks on terminology. (i) Consequentialism I take to be the style of ethical theory that makes the moral quality of an action depend solely on the consequences of an action, intended or real; it does not reside in the action itself (cf. Anscombe, 'Modern Moral Philosophy', p. 37, and Richard Hare, 'A Utilitarian Approach to Ethics', *Objective Prescriptions and Other Essays* (Oxford, 1999), p. 151). By contrast, one could call the utilitarian theory of value 'teleological' in that it determines the ends that are of value. Like Smart and Williams (J. J. C. Smart and Bernard Williams, *Utilitarianism For and Against* (Cambridge, 1963), pp. 13, 83), Hare has difficulties seeing how anyone, Kant included, could not be a consequentialist. He says that what is wrong about lying is the consequence of someone's being deceived and then moves on to say that it is the 'intended' consequences that make acts wrong (p. 15). Yet it is obvious that Kant was not a consequentialist in this or any other sense. It is the rational contradiction in one's will that makes an action wrong, a certain rational inconsistency, not the actual or intended consequences of the action. The latter enter the moral argument, but they do not 'do the trick'. Moreover, it seems a mistake to say that 'intended' consequences are a species of consequences, as Hare does. Intended consequences are intentions, not consequences, and consequentialists who ground their theory in intentions are well on their way to abandoning consequentialism. (ii) The term 'deontology', used by Hare to characterize Kant's theory (p. 2), is not very helpful if it is taken merely to put an emphasis on duty because it obliterates the deeper meta-ethical differences between 'deontology' on the one hand and 'consequentialism' or 'teleology' on the other. For example, the categorical imperative does not admit of degrees of moral goodness or badness, as consequentialist theories do. Their respective styles of argument are entirely different. And, ironically, the term 'deontology' was made popular by the father of utilitarianism, Jeremy Bentham, in 1826. The *Oxford English Dictionary* quotes Bentham: 'Ethics has received the more expressive name of *Deontology*'. It is defined as 'the science of duty', 'that branch of knowledge which deals with moral obligation', or 'ethics'. (Cf. Robert B. Louden, 'Toward a Genealogy of "Deontology"', *Journal of the History of Philosophy* 34 (1996).) In this sense, there is a utilitarian deontology. I doubt whether there is any other philosophically useful sense of the word. We should probably banish it from our active vocabulary.

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

VI. CONCLUSION

If this analysis of Kantian moral psychology and ethical theory is correct, what shall we make of Hare's opening quotation? Does Kant not, in Norman Kemp Smith's words, say that 'the happiness of all mankind' is 'the supreme end' in the Methodology of the *Critique of Pure Reason*? He does,[43] but we must always remember that in Kantian ethics the ends that people pursue are not the be-all and end-all of morality. Personal happiness is naturally the most immediately important end for any human agent. It is an 'inclusive' good constituted by the satisfaction of all of one's desires and inclinations. Moreover, everyone ought to further the happiness of others. But the ends we are inclined to pursue, considered individually or in sum, still do not determine the moral appropriateness of people's actions and their principles. There are formal restrictions. Paradoxically (as Kant is well aware), these restrictions can also be spelled out in terms of a radically different kind of ends: human creatures, independently existing ends in themselves that demand respect, not production or promotion. These are the ends for the sake of which we perform moral actions.

In sum, Kant was not and *salva identitate* could not have been a utilitarian because the satisfaction of preferences or the promotion of happiness has neither the weight nor the theoretical status that utilitarianism accords to it.[44] He insists that the faculty of reason

[43] Kant sounds somewhat less enthusiastic in the original German. He says the new metaphysics, conceived of negatively as the censor of philosophy rather than the fount of transcendent truths, should make sure that we do not move away 'von dem Hauptzwecke, der allgemeinen Glückseligkeit', i.e. 'from the chief end, universal happiness' (B 879).

[44] The connections between happiness and morality in Kantian ethics are quite intricate. Over and above the basic idea that everybody naturally and legitimately wishes to be happy – which is common ground between Kantians and most utilitarians – there are six points of contact. There is, first of all, a direct if imperfect duty to make others happy; second, an indirect duty not to neglect one's own happiness; third, the fact that moral action contributes towards the agent's contentment, even though it cannot be the motive for moral action and does by no means always outweigh the losses incurred; fourth, we take delight in other people's moral actions; and fifth, acting on moral principles makes the agent worthy of happiness: the conjunction of the two factors is said to be the highest (comprehensive) good. A sixth and final point should come as no surprise. If all agents were moral everyone would be maximally happy as well. But that does not mean that happiness should be the only end our actions aim at, least of all the ground of moral value. The last point follows from the preceding one – especially Kant's emphasis on perfection, i.e. being worthy of being happy. If everyone adopted the motto as a personal principle, they would be worthy of happiness and others would be justified in making them happy. Moreover, everyone would take delight in his own moral actions and the moral actions of others. The world would be a wonderful place. But still, making the world a wonderful place is not the one thing that justifies universal compliance with morality. It is its highly welcome effect. It is therefore not impossible that second-level utilitarianism should command all rational agents to be first-order Kantians for the sake of maximizing overall utility. However, in that case utilitarianism would probably be redundant. Cf. Collins's *Lectures on Ethics*, XXVII:285 f., which emphasizes the importance of meritorious duty:

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press

grounds moral goodness both in ordinary moral thought and in moral philosophy, and that agents must be motivated accordingly. Making other people happy, responding to their preferences, even their morally permissible ones, is no doubt in many cases a morally laudable thing to do – but it is not morally good by virtue of the value of happiness, which is conditional, nor is it the only kind of moral goodness.[45]

jt28@st-andrews.ac.uk

'[I]f only all men together were unanimously willing jointly to promote their happiness, we might make a paradise in Novaya Zemlya' (the Russian Arctic archipelago later used as a nuclear test site); and Refl. 6899. On the role of happiness in Kantian ethics generally cf. Thomas Hill, 'Happiness and Human Flourishing'.

[45] I should like to thank audiences in St Andrews, Helsinki and Cracow for their helpful suggestions, as well as Robert Louden and Elizabeth Pybus for comments on earlier versions of this article.

https://doi.org/10.1017/S0953820805001639 Published online by Cambridge University Press