Hume, Kant and the Search for a Modern Moral Philosophy

Ingvar Johansson
Umeå University[1]

(This paper has, slightly revised, been published in Swedish in the Danish journal *Philosophia* **27,** 3-4, 1999, pp. 5-43; the note numbers have been adjusted in order to fit the published version.)

The modern problem of ethics is essentially based on three opinions. (1) We cannot ground norms, values, or virtues in an external authority like God; not even if he exists, which, at least, we do not know. (2) Norms, values, and virtues are man-made. (3) Knowledge or beliefs about matters of fact cannot make us act; in particular, such beliefs cannot turn us into moral persons. Is it really possible, the question is, to regard these three claims as true and at the same time believe that there is something worthy of the name ethics or some close synonym?

Sometimes philosophical problems benefit from being seen, for a while and with hindsight, through the minds of those who first grappled with them. Looking back at the history of philosophy before our own rapidly dying century, there seems to be a fair amount of consensus in the Western World who the great thinkers are. A top ten list would probably consist of the two great Greeks, Plato and Aristotle; the three bright British empiricists, Locke, Berkeley, and Hume; the three clever Continental rationalists, Descartes, Spinoza, and Leibniz; and, at last, the two good (or marvelously bad) Germans, Kant and Hegel. Among these, only Hume and Kant accept something like the triple-claim above and the problem it raises.[2] Moreover, both of them thought they could solve the problem, although, of course they did it in different ways. I myself find their proposed solutions as wanting, but, nonetheless, I think that both of them have found *necessary* ingredients of a modern moral philosophy.

I think that we have lessons to learn from Hume's philosophy primarily (but not only) when we take the spectator's view of morals, and lessons to learn from Kant's philosophy primarily (but not only) when we take the agent's view of morals.[3] Hume has no real place for agency in his ontology, whereas Kant puts agency (i.e. 'causality of freedom') outside the spatiotemporal world. He placed it in a purely intelligible world, the transcendent realm of *noumena*. But a real modern ethics should have agency within the spatiotemporal world. It is fruitful to start in Hume and Kant but we should of course not, like this paper, rest content here.

---

[1] There are two persons at the Department of Philosophy and Philosophy of Science in Umeå which I want to thank: Per Nilsson (for extensive Kant-discussions without which my interpretation of Kant would never have emerged) and Gunnar Andersson (for the Catullus quotation and for the insight that respect for logic is grounded in respect for truth; see section 7).

[2] I guess that some readers want to say that, surely, Hume accepts the triple-claim but that Kant, who tried to ground ethics in reason, did not accept the third claim. I explain my interpretation of Kant at the beginning of section 4 below. Reason for Kant, is *practical* reason, and that is something else than knowledge.

[3] I do not know who was the first to make an observation like this, but Wilhelm Dilthey wrote in the early 1880s that ethics is both ”present as judgement of behavior by an observer and as an element in motivation ... ethics has until now taken one or the other form as its basis - Kant and Fichte taking the ethical as living power in motivation and the principal English moralists along with Herbart taking it as a power responding to actions of others from the outside - ”; Dilthey, *Introduction to the Human Sciences*, Harvester: London 1988, p 113.

1. Hume: Psychological Egoism is False

According to Hume, since there is neither a God nor any mind-independent moral facts, morals has to be grounded in human nature. Moral philosophy needs a philosophical anthropology. Human beings are very egoistic, but the theory of psychological egoism is false. To human nature belong also feelings of sympathy, benevolence, and humanity. These feelings are the roots of morals. If there are no conscious beings beside human beings, Hume's point can be put in the following way: If human nature did not contain a nose there would be no smells, if human nature did not contain eyes there would be no colors, and if our human nature did not contain dispositions and sentiments of sympathy, benevolence, and humanity, there would be neither norms, values, or virtues. For blind people there are no colors. For purely egoistic people, if there are any, there are no morals. A purely egoistic man is morally blind.

Three "Humean truths" ought to be distinguished. (a) If psychological egoism is true we cannot ground ethics philosophically, (b) psychological egoism is *not* true, and (c) *all* human beings have a capacity to feel with and to feel for *every* other human being.[4] The first claim is philosophical, whereas the other two are primarily empirical propositions, although with problems of interpretation which make them partly philosophical. I will argue that Hume is quite right in his claims (a) and (b), but somewhat wrong in (c). Some psychopaths and autistic persons simply lack sentiments of sympathy, i.e. *all* human beings do *not* have feelings of sympathy. However, the overwhelming majority do. Proposition (c) can be false both with respect to the expression 'all' and with respect to the expression 'every'. We should therefore ask, next, if within the assumed majority there is a sentiment of humanity which extends sympathetic approval to *every* man. Hume himself says the following:

> The notion of morals implies some sentiment common to all mankind, which recommends the same object to general approbation, and makes every man, or most men, agree in the same opinion or decision concerning it. It also implies some sentiment, so universal and comprehensive as to extend to all mankind, and render the actions and conduct, even of the persons the most remote, an object of applause or censure, according as they agree or disagree with that rule of sight which is established. These two requisite circumstances belong alone to the sentiment of humanity here insisted on.[5]

Hume is very assured on this point. He says, in a beautiful passage, that we cannot "without the greatest absurdity" dispute "that there is some benevolence, however small, infused into our bosom; some spark of friendship for human kind; some particle of the dove kneaded into our frame, along with the elements of the wolf and serpent."[6] He stresses that our feelings of sympathy in relation to our closest friends and relatives are much stronger than in relation to anonymous people, but even if it is often the case that our "generous sentiments" are "insufficient to move even a hand or finger of our body"[7] they make us able to make moral distinctions. We can judge morally in a particular case even if we have no propensity to act in favor of what is in the case at hand good and right.

(Hume, it can be noted, never keeps pure feelings distinct from propensities to act. His favorite term, sentiment, embraces both feelings and desires. As the dictionary says: "The sentiment of pity is made up of the *feeling* of sympathy and of a *desire* to help and protect."[8])

---

[4] I have taken the phrase "feel with and feel for" from J.H. Sobel, *Walls and Vaults. David Hume's Natural Science of Morals*, (mimeographed) University of Toronto 1995, p 148.

[5] *Enquiries concerning Human Understanding and concerning the Principles of Morals* (ed. Selby-Bigge), Clarendon Press: Oxford 1975, p 282.

[6] *Enquiries*, p 271.

[7] *Enquiries*, p 271.

[8] *The Advanced Learner's Dictionary of Current English*, Oxford UP: London 1963.

A close reading of Hume reveals that he makes both a distinction between *private* and *public* benevolence, and a distinction between *interested* and *disinterested* benevolence.[9] The distinctions are, I think, self-explanatory; public benevolence is directed at the whole of mankind or some relevant large part of it. Obviously, some psychological law makes disinterested benevolence more connected with public benevolence than with private. It is easier to be disinterested with regard to suffering anonymous people than with regard to a suffering friend. But there is no universal correlation. When two beloved children intensely want something which only one of them can get, then disinterested benevolence is needed from their parents.

Whether or not it is true that most of us have a sentiment of public and disinterested benevolence, I think today's psychology and anthropology have shown that we have a sentiment of private benevolence. And that is enough to refute psychological egoism. *Private* benevolence is a desire to regard, *now and then*, some other being's well-being *as an end in itself*. Mutual private benevolence must be carefully distinguished from mutual *instrumental* benevolence. If I promise to scratch your back if you promise to scratch mine, I am not benevolent. I am merely in search of a contract where we use each other as instruments for our own satisfaction. 'If I help him, he will probably help me', that is the pattern of reasoning behind instrumental benevolence. Pure benevolence, benevolence as an end in itself, is something else. When we are benevolent we just want to help, neither more nor less.

Private and interested benevolence consists in a desire to make at least one other conscious being (human or animal[10]) more happy for the moment. This desire is most apparent in the case of parent-child relationships. It then takes the form of a disposition for caring.[11] Some philosophers and a lot of economists of course think that this is mere illusory appearance. The essence even of caring is egoistic.[12] Much could be said about this position, but this is not the place to do it. However, I take a brief quotation from Hume:

> Whatever contradiction may vulgarly be supposed between the *selfish* and *social* sentiments or dispositions, they are really no more opposite than selfish and ambitious, selfish and revengeful, selfish and vain.[13]

From an egoistic point of view one may loose (or win) a lot by being ambitious, revengeful and vain. The dispositions and the corresponding passions or impulses may arise independently of the outcome. Egoism should not be confused with impulse satisfaction and desire satisfaction. We have impulses to be benevolent just as we have impulses which are directed only at our own satisfaction. The satisfaction may in both kinds of cases cause us feelings of pleasure, but that does not turn this pleasure into the intentional object of the impulse. When we are hungry we want food for ourselves to eat, but when we meet a hungry friend and feel for him we *want him* to have food. In both cases we primarily desire food, not

---

[9] Hume distinguishes between private and public benevolence in *A Treatise of Human Nature* (ed. Selby-Bigge), Clarendon Press: Oxford 1975, p 482; the concept of disinterested benevolence is used in *Enquiries*, p 301.

[10] Although Hume has no discussion of animals, he is of course aware that we can have feelings of sympathy for animals. See *Treatise* p 481. It gives rise to a problem which Hume does not notice, and which I shall consciously bypass. Private and interested benevolence towards pet animals is often stronger than public and disinterested benevolence towards even not too-anonymous people. We simply do not care for our whole species *before* we start to care for other animals.

[11] For a heavy stress on caring as a basic category, see Nel Nodding, *Caring. A feminine approach to ethics and moral education*, University of California Press: Berkeley 1984. Even if Hume was a bachelor his whole life, he of course noticed "certain instincts ... as ... kindness to children". *Treatise* p 417.

[12] I think the most elaborated defense of psychological egoism of the latter decades are to be found in the works of the Nobel prize winner in Economics in 1992, Gary Becker.

[13] *Enquiries,* p 281.

pleasure. Survival comes before pleasure. As C.D. Broad pointed out long ago, we have to keep four things conceptually distinct:

(i) the intentional object of a desire;
(ii) its exciting causes;
(iii) what will in fact satisfy the desire;
(iv) the collateral effects of satisfying the desire.[14]

In my view, whatever *the exciting causes* are, most of us *sometimes* simply want to make someone else happier (= *the intentional object*), and, furthermore, we will be satisfied if we manage to make this person happier. When this happens, we will mostly but *not* always, as a collateral effect, get feelings of pleasure. Those who think, like Thomas Hobbes, that they always give alms only because they find pleasure in satisfying the poor, ought to consider themselves in the light of Broad's distinctions.

My comments mean that I agree with Hume's claims that (b) psychological egoism is false, and that (c) our human nature contains a sentiment of humanity. Hume is right in his rejection of psychological egoism, but he overstates the case for the sentiment of humanity. What seems to be true, is merely that almost all human beings have a sentiment for *private benevolence*. Next we shall look at Hume's reasons for his claim that (a) if psychological egoism is true we cannot ground ethics philosophically. In so doing I shall focus attention on a concept which is indispensable in a modern ethics, the concept of *self-subordination*. Hume, however, finds this concept unintelligible.

2. Hume: Self-Subordination is Impossible

If psychological egoism is true, then, at bottom, all our sentiments are only concerned with ourselves. Even in such a case there could of course arise common imperatives. But they would have to be grounded in mutual *instrumental* benevolence, and they would not, therefore, constitute *moral* imperatives. An imperative which people adhere to merely because they profit from it, is not a moral imperative. If, as Hume thinks, ethics can be grounded in nothing but sentiments, then ethics requires something like the sentiment of humanity. But are there really nothing beside sentiments in human nature in which morals might be grounded? Hume is famous for making it clear that morals cannot be grounded in reason, but what about the will? What does he say about that possibility?

There is a section in *Treatise* which I think supplies us with Hume's reasons for his view that ethics cannot be grounded in acts of the will. When Hume analyzes justice, he devotes one section to the question whether obligations can be grounded in promises, or more specifically, in *natural* promises.[15] As *non-natural*, i.e. as human *conventions*, promises can give rise to sympathetic approvals and in this way get a moral value. That is in conformity with his general views, and constitutes no consistency problem for him. What is at stake now is whether a promise in the state of nature, and independently of all sentiments, can create an obligation. What, then, is the essence of a natural promise?

> If promises be natural and intelligible, there must be some act of the mind attending these words, *I promise*; and on this act of the mind must the obligation depend. Let us, therefore, run over all the faculties of the soul, and see which of them is exerted in our promises.

---

[14] Broad, *Five Types of Ethical Theory*, Routledge & Kegan Paul: London 1930, p 67-71.
[15] *Treatise*, book III, part II, section V.

> The act of the mind, exprest by a promise, is not a *resolution* to perform any thing: For that alone never imposes any obligation. Nor is it a *desire* of such a performance: For we may bind ourselves without such a desire, or even with an aversion, declar'd and avow'd. Neither is it the *willing* of that action, which we promise to perform: For a promise always regards some future time, and the will has an influence only on present actions. It follows, therefore, that since the act of the mind, which enters into a promise, and produces its obligation, is neither the resolving, desiring, nor willing any particular performance, it must necessarily be the *willing* of that *obligation*, which arises from the promise.[16]

According to Hume, to make a natural promise is to will an obligation to arise from the promise. This is the feature which all natural promises have in common. If the content of a promise is disregarded, then a promise is a promise for promise's sake. This sounds Kantian and is Kantian. It is duty for duty's sake. In a promise we exert a will to be bound in a very specific way. We *exert a will to be subordinated to an obligation created by our own will*, i.e. we want self-subordination. Kant finds such a will and obligation not only intelligible, he claims that it is necessary for rational beings to have it. Hume, however, finds the idea of such an obligation wholly unintelligible. This is the gulf between them. I shall discuss Kant's view later on, now we stay with Hume for a while. First, we shall distinguish between direct and indirect binding of oneself. In the case of *indirect* self-subordination, our will has to create something in ourselves that is distinct from our will but which then commands our will. For Hume, of course, the will, in this case, has to create a sentiment.

> a creation of a new obligation supposes some new sentiment to arise . But 'tis certain we can naturally no more change our own sentiments, than the motions of the heavens; ... It would be absurd, therefore, to will any new obligation, ... A promise, therefore, is *naturally* something altogether unintelligible, nor is there any act of the mind belonging to it.[17]

Indirect self-subordination is impossible since the will cannot create any sentiments. The possibility of *direct* self-subordination is discussed only in a footnote.[18] There, Hume says that "Morality is suppos'd to consist in relation", by which, I think, he means that if a promise creates a moral obligation it must create a relation between the act of promising and that future action which is promised. The act of promising, however, is a self-contained act and cannot create this relation. A new will, therefore, is needed to create this relation between the promise and the promised action. But this will, in turn, suffers the same fate. In Hume's own words:

> To will a new obligation, is to will a new relation of objects; and therefore, if this new relation of objects were form'd by the volition itself, we shou'd in effect will the volition; which is plainly absurd and impossible. The will has here no object to which it cou'd tend; but return upon itself *in infinitum*.[19]

Obscurities are sometimes suppressed into footnotes where the reader is told that something is obvious. This sad truth applies to great thinkers, too. Is it really, as Hume said, *plainly absurd* to will a volition, i.e. are second-order willings logically impossible? With regard to *desires*, for instance, no absurdities arise when we distinguish between a first-order level and a second-order level. I can at one and the same time have a (first-order) desire to smoke and a (second-order) desire not to have this desire. This second-order desire *opposes* the first-order desire but does not necessarily cancel it. In a similar way, if a second-order desire *supports* a first-order desire, this does not mean that the second-order desire is submerged in the first-

---

[16] *Treatise*, p 516.
[17] *Treatise*, p 517.
[18] *Treatise*, p 517, footnote.
[19] *Treatise*, p 517-18.

order desire in such a way that it loses its identity. A man can at one and the same time both have a desire to smoke and have a desire to have this desire. That these desires can be kept distinct from each other is shown by the fact that most persons with such desires can discuss both whether or not they ought to smoke and whether or not they *ought to* have a desire to smoke.

What is true of desires need not be true of willings. In fact, I think that Hume had the right intuition that there is something mysterious with second-order willings. However, I do not think that he managed to find the right arguments. But here, I hope, they come. If I freely will smoke a cigarette, I *cannot* at the same time freely will not to will this. My opposing second-order will would, if real, immediately cancel my first-order will. If, on the other hand, my second-order will goes in the same direction as my first-order will, then the second-order willing is wholly submerged in the first-order willing. A free will to have a free will is to have this will. Perhaps an analogy with how a 'truth-operator' works may make the idea clearer. If I have said 'It is true that p is the case', then I add no new descriptive content by saying 'It is true that (it is true that p is the case)'.[20] On the other hand, if I say 'It is false that (it is true that p is the case)' then the 'falsity-operator' takes away the effect of the 'truth-operator'. When I have freely decided to smoke, nothing is added by a new free decision to decide to smoke. And if, instead, I add a new free decision to decide *not* to smoke, then my first decision is automatically canceled.

My argument above, just like Hume's own arguments, is about the *general* concept of willings of volitions, but the oddity Hume is trying to explicate, comes out even better if one restricts the argument to *one kind* of willings, namely self-subordinative willings. There is a close analogy between self-subordination and self-deception which I want to highlight.

The classical paradox of self-deception looks as follows. When one person, A, tries to deceive another, B, into believing that p is the case, then A himself is believing that not-p is the case. If A succeeds, then B believes that p and A himself believes that not-p. In self-deception, however, the deceiver and the deceived, A and B, is one and the same person. Therefore, a victim of self-deception has to believe both that p and that not-p. Self-deception, construed in this way, amounts to an explicit *belief in a contradiction*. For mentally sane people, such a strong kind of self-deception is impossible. Like A.R. Mele, I think that *ordinary* self-deception is something else, and that we should regard *direct* self-deception as something absurd.[21]

With regard to self-subordination, a paradox similar to the paradox of direct self-deception can be construed. When one person, A, tries to command another, B, into doing D (i.e. A tries to bind B's will so that B does D) then A regards his own will as free. If A succeeds, B's will is bound but A's will is free. Now, in self-subordination of course the subordinator and the subordinated, A and B, is one and the same person. This means that a self-subordinated person has to will both unfreedom and freedom (at the same time, in the same person, and in the same respect). Such self-subordination, i.e. direct self-subordination, is self-canceling; it tries to bring a logical contradiction into the world. Later on (section 7), I shall discuss *indirect* self-subordination and claim, against Hume, that some feelings can be subject to our will. Now, I merely want to repeat that I find, with Hume, *direct* self-subordination to be an absurdity.

Hume's analysis of promises comes, it ought to be added, very close to an Existentialist view of promises. In *Being and Nothingness*, Sartre describes a gambler who has promised to

---

[20] It may, *pace* Austin, add illocutionary force.

[21] A.R. Mele, *Irrationality. An Essay on Akrasia, Self-Deception, and Self-Control*, Oxford UP: Oxford 1987, chapter 9.

stop gambling.[22] When the gambler at the sight of a gaming table suddenly perceives it in anguish, this is not, according to Sartre, due to the fact that the gambler is suddenly caught by a violent desire, a desire to gamble, which is in conflict with his promise. His anguish is wholly due to the fact that the gambler realizes that he is *free* to gamble, that his promise has no ontological hold on him. He can choose either to play or not to play. In Sartre's own prose:

> I should have liked so much not to gamble anymore; yesterday I even had a synthetic apprehension of the situation (threatening ruin, disappointment of my relatives) as *forbidding me* to play. It seemed to me that I had established a *real barrier* between gambling and myself, and now I suddenly perceive that my former understanding of the situation is no more than a memory of an idea, a memory of a feeling. In order for it to come to my aid once more, I must remake it *ex nihilo* and freely.[23]

Sartre calls this case ”anguish in the face of the past”, and it is part and parcel of Sartre's argument that we are necessarily ontologically free. Even if we can fool ourselves into ”bad faith”, we are essentially free. Those who really think that promises are binding are living in such bad faith. A will which has put itself under an obligation is an ontological impossibility; it is freedom and unfreedom at one and the same time. The difference between Hume and Sartre is that Sartre believed in ontologically free wills, but that Hume did not. (According to Hume, beneath the appearances of substances, causal relations, and an ego with a free will, the world in fact contains only passions, impressions (of sensations), and ideas.[24]) However, when Hume toys with free willings, then he finds the idea of a will which obliges itself unintelligible.

3. Hume: Virtues are like Colors

Morals, according to Hume, are created by our sentiments of sympathy, benevolence, and humanity.[25] These sentiments function the same ontological way our sensory organs function. Just like our faculty of vision creates colors, our sentiment of sympathy creates sympathetic approval and disapproval. This is the usual short-hand presentation of Hume's view on the foundations of morals. However, following J.H. Sobel,[26] I think it is important to stress that for Hume the sentiment of sympathy is no more or no less hallucinating than our visual faculty is. Even though, as being scientifically educated people, we believe that mind-independent things around us have no colors, we do think that there is some structure in the things themselves which *together* with our visual faculty (and the reflected light) create colors. There is something in the things themselves which make us see and project colors out onto these things. Therefore, ordinary color perception is quite different from a colorful hallucination. Virtues, according to Hume, should be analyzed in exactly the same way. When

---

[22] J-P. Sartre, *Being and Nothingness*, Washington Square Press: New York 1966, p 69-70. In fact, Sartre talks of a *decision* not to play, not a *promise*. But I am sure that he could just as well have talked about a promise.

[23] *Being and Nothingness*, p 70.

[24] I have written ‘impressions (of sensation)’ since Hume regards passions as one particular kind of impressions. *Treatise*, p 1.

[25] There is no reason for me to discuss why Hume in his *Enquiry* stresses humanity instead of sympathy as in his *Treatise*. But I am in agreement with the following quotation: ”In the Treatise ”benevolence” is, indeed, regarded as an original instinct; but it is a confined benevolence, a desire for the happiness of our own friends; there is no natural and original love of man for man as such. The concern for the general happiness is due to the mechanism of sympathy, the natural attraction of ideas and impressions. In the Enquiry all this is dropped, and a natural and universal benevolence or ”sentiment of humanity” is substituted.” From D.G.C. MacNabb, *David Hume. His Theory of Knowledge and Morality*, Hutchinson: London 1951, p 190.

[26] J.H. Sobel, *Walls and Vaults. David Hume's Natural Science of Morals*, (mimeographed) University of Toronto 1995, in particular chapters I and II.

we find someone virtuous, it is not *only* that we have pleasurable feelings of a special sort (sympathetic approval). The virtuous person causes, together with our sentiment of sympathy, these approvals. As colors in one sense are both powers in things and sensations in us, virtues are both powers of other persons and pleasurable sensations in us. These latter kind of powers are powers of other minds to produce in our own minds sympathetic approvals. Before starting section VIII of his *Enquiry*, which is entitled ”Of qualities immediately agreeable to others”, Hume added a footnote, which reads as follows.

It is the nature and, indeed, the definition of virtue, that it is *a quality of the mind agreeable to or approved of by every one who considers or contemplates it*. But some qualities produce pleasure, because they are useful to society, or useful or agreeable to the person himself; others produce it more immediately, which is the case with the class of virtues here considered.[27]

The qualities which Hume find *immediately agreeable to others,* make up a list which is a mix of traditional Greek, pagan and Christian virtues.[28] I find it a nice list, although, from a modern point of view one virtue is missing, the virtue of tolerance. The remarkable thing, however, is that Hume never discusses virtue conflicts. What about the fact that different cultures have partly different sets of virtues? Obviously, Hume regards his list as universal. He writes as though every human being is caused into sympathetic approval by the same character traits. Such a conclusion is in no way entailed by his principle of humanity. In a world where there are no psychological egoists, there may nonetheless exist groups of people whose sympathetic approvals are different. In a world where there are no blind people, there may nonetheless exist people who are color blind.

4. Hume: It is Impossible to Will to be a Free Rider

One of the main problems that confront a modern ethics is that of the free rider. Even if there are a lot of good reasons to let some good imperatives rule one’s society, it seems to be hard to find reasons which totally blocks the opportunity that, now and then, one makes oneself an exception to the rules. Hume, mostly, take a spectator’s view both of others and of himself. As an onlooker he explains to the reader, as another onlooker, what virtues and vices are, what justice is, and how they arise. As long as people have the sentiments of sympathy, benevolence, and humanity, there will be moral orders in the world. From the spectator’s view, there simply is no free rider problem. If I am not going to act, I do not have to ask myself whether I shall allow myself to be vicious or unjust despite the fact that I realize that virtues and justice are necessary for the moral order that I want to exist. The pure philosophical anthropologist can rest content with noticing (1) that norms, values, and virtues are needed in any society, (2) that free riders are possible, and (3) that if there are too many free riders the moral order will disappear. Hume is quite convinced that the free riders are not too many. In at least one place, however, he is considering the case we nowadays call the problem of the free rider.

And he**,** it may perhaps be thought, conducts himself with most wisdom who observes the general rule and takes advantage of all the exceptions.

---

[27] *Enquiries*, p 261.

[28] I refer to J. Casey, *Pagan Virtue. An Essay in Ethics*, Clarendon Press: London 1990, for details.

I must confess that if a man think that this reasoning much requires an answer, it will be a little difficult to find any which will to him appear satisfactory and convincing. ... But in all ingenious natures the antipathy to treachery and rougery is too strong to be counterbalanced by any views of profit or pecuniary advantage.[29]

This passage with its evading answer is part of a larger argument, where, really, Hume tries to answer the free rider problem. He makes a thought experiment and asks whether a man can choose, i.e. *can will*, to be vicious. He writes:

let a man suppose that he has full power of modelling his own disposition, and let him deliberate what appetite or desire he would choose for the foundation of his happiness and enjoyment.[30]

And he answers:

But were they /the knaves/ ever so secret and succesful, the honest man, if he has any tincture of philosophy, or even common observation and reflection, will discover that they themselves are, in the end, the greatest dupes, and have sacrificed the invaluable enjoyment of a character, with themselves at least, for the acquisition of worthless toys and gewgaws.[31]

Hume is not particularly clear, but as I understand him, his argument differs from the common argument that one will in the long run, *from an egoistic point of view*, be better off if one chooses to adhere to a set of given norms. This is norm-following based on pure egoism; it is *instrumental* benevolence. But this argument in no way blocks oneself from being an occasional free rider. For Hume, however, the free rider becomes impossible because Hume is not thinking of how to choose *actions*, but of how to choose a *character*. Even in this thought experiment he takes it for granted that it is only sentiments which can make us act. What is to be chosen are sentiments, i.e. a character. I think Hume should be interpreted as saying that if you can choose your own character, you should from an egoistic point of view choose not to be egoistic. You should choose a virtuous character, and with such a character you cannot become a free rider even if you now and then would benefit from it. On Hume's premises, a will informed by reason cannot will otherwise than be installed in a virtuous character. For him, a reflective second-order egoism cancels our ordinary first-order egoism.

5. Kant: Oughts Originate Outside Nature

'*Ought*' expresses a kind of necessity and of connection with grounds which is found nowhere else in the whole nature. … We cannot say that anything in nature *ought to be* other than what in all these time-relations it actually is. When we have the course of nature alone in view, '*ought*' has no meaning whatsoever.[32]

As persons, in contradistinctions to mere animals, we do, according to Kant, belong to three realms. Two of the realms contain objects and deserve to be called worlds. It is on the one hand our immanent spatiotemporal world of phenomena, i.e. nature with its sensible objects, and it is on the other hand the transcendent non-spatiotemporal and non-sensible world of noumena, i.e. intelligible objects.[33] In between them, so to speak, as a borderland there is a

---

[29] *Enquiries*, p 282-83.
[30] *Enquiries*, p 281.
[31] *Enquiries*, p 283.
[32] Kant, *Critique of Pure Reason* (transl. Kemp-Smith), MacMillan: London 1968, p 472-473 (B 575).
[33] In a footnote, Kant commented upon the term 'intelligible': "We must not, in place of the expression *mundus intelligibilis*, use the expression 'an *intellectual* world', … For only modes of knowledge are either intellectual or sensuous. What can only be an *object* of the one or the other kind of intuition must be entitled … intelligible or sensible." *Critique of Pure Reason*, p 273 (B 312).

transcenden*tal* realm consisting of non-personal *faculties*; in particular, the faculties of understanding and reason. Faculties are powers; they are neither thing-like nor property-like entities. All our ordinary feelings, desires, and sentiments belong to our sensible world and are directed at objects and states of affairs in this world. We can know our feelings and desires by means of empirical self-consciousness, but we can do nothing about them. They are created by nature and ruled by the causality of nature. Morality, however, requires agency or causality of freedom, and cannot be found in nature.

We can have knowledge, in the proper sense of that word, of the immanent world and the transcendental realm but not of the noumenal world. This epistemological view of Kant creates a problem for him with agency, since, if freedom exists, it exists in the transcendent world of noumena; and only if we belong to this world are we really free. The categorical imperative (the moral law), however, is created by the transcendental faculty of *reason*; not by the noumenon of freedom, nor by the transcendental faculty of *understanding* which creates the categories (substance, causality, etc.) which are necessary in order for the spatiotemporal world to be structured and known. Nonetheless, the noumenon of freedom is a necessary condition for the moral law to be created. The moral law, in turn, is a condition for us to be able to assume that there is freedom. Our freedom is not known in the sense in which we know nature, but it is there, it shows itself, and is in some sense known.[34] We had better listen to Kant himself:

> To avoid having anyone imagine that there is an inconsistency when I say that freedom is the condition of the moral law and later assert that the moral law is the only condition under which freedom can be known, I will only remind the reader that, though freedom is certainly the *ratio essendi* of the moral law, the latter is the *ratio cognoscendi* of freedom. For had not the moral law already been distinctly thought in our reason, we would never have been justified in assuming anything like freedom, even though it is not self-contradictory. But if there were no freedom, the moral law would never have been encountererd in us.[35]

Hume uses the terms ‘understanding’ and ‘reason’ interchangeably, but Kant does not. When Hume says that reason cannot move us to act, he could equally well have said that understanding cannot so move us. In Kant’s view and terminology, understanding cannot move us to action, but reason can. Reason, seemingly, has two kinds of employments, one which is speculative and one which is practical.[36] Its speculative employment, however, cannot, contrary to the claims of all rationalist philosophers, give us knowledge. It is a sham employment. Reason *is* practical reason, and the faculty of practical reason *is* a will, and this will *is* free; it is causality of freedom.

> Only a rational being has the capacity of acting according to the conception of laws, i.e., according to principles. *This capacity is will*. Since reason is required for the derivation of actions from laws, *will is nothing else than practical reason*.[37] (Emphasis added.)

Nature cannot contain a free will, but transcendental practical reason can. In each empirical ego there is an *animal will* which is wholly determined by sensuous impulses and desires.

---

[34] It is, though, it should be noted, the one and only idea of reason which can be known to have a noumenal (or intelligible) object. He says that ”of all intelligible objects nothing [is known] except freedom (through the moral law)”, *Critique of Practical Reason*, Prentice Hall: New Jersey 1993, p 73. About the existence of God and immortal souls we know nothing in any way.

[35] *Critique of Practical Reason*, p 4 (note).

[36] See, for instance, Kant, *Critique of Pure Reason*, MacMillan: London 1968, p 24-25 and 617.

[37] I. Kant, *Foundations of the Metaphysics of Morals*, Liberal Arts Press: Indianapolis 1959 (transl. L.W. Beck), p 29.

Transcendental practical reason, on the other hand, is a *free will*[38], i.e. a will which is determined only by motives created by itself. As I have remarked above, however, this *transcendental* free *will* is for its existence dependent upon the existence of freedom as a *transcendent* intelligible *object* in the noumenon-world. Reason creates, all by itself, an *idea* of freedom. But for freedom to be real it must exist in itself as an intelligible object. In other words, if the idea of freedom has a referent, then the will which created the idea is also in fact free. This curious dependence which Kant finds between our *transcendental* free will and our *transcendent* freedom will be left out of account in this paper.

According to Kant, our free will does not exist in space and time, and it cannot obstruct any causal laws of nature. Nonetheless it can in some peculiar way make us act spontaneously. Agency appears in nature but is created outside nature. When Kant says that (practical) reason can and ought to affect our actions, he is *not* saying that knowledge or beliefs about matters of fact can move us to act. He is saying that our *free will* can. This transcendental will is also, to repeat, the creator of the categorical imperative.

The categorical imperative is often described as a second-order principle by which ordinary rules of action (*maxims*) are to be tested.[39] But that is its function only in relation to the sensible world and all the maxims which arise because of our sensuous desires. Kant's categorical imperative functions in one way in the transcendental realm where it is created, and in another in relation to our world of desires and sentiments. I shall present the purely transcendental use in this section, and postpone the discussion of its application in our spatiotemporal world to the next section.

Sensuous desires can be directed at particular spatiotemporal objects, but transcendental reason can only be concerned with universals; the faculty of reason can only work on universal things like concepts, judgments, logical relations, and principles. Therefore, if practical reason wills an imperative, this imperative can directly only be concerned with things or states of affairs which are universal. A transcendental will cannot possibly pick out and favor one particular person. It can only will universal maxims. In the way now sketched, Kant finds the first formulation of the categorical imperative, a formulation which supplies us with the *form* of this imperative: *Act only according to that maxim by which you can at the same time will that it should become a universal law*.[40] To repeat: This imperative is primarily *willed* by reason, although, secondarily, of course, it must be cognized at the same time.

As long as the transcendental will is willing only with regard to the transcendental realm, it can only be directed at other entities in this realm. From a moral point of view, the concepts of understanding and the corresponding judgments are uninteresting. Therefore, the only morally interesting thing in the transcendental realm is the transcendental free will (i.e. practical reason) itself and its willings. And this kind of will, Kant tells us, must be regarded as an end in itself. There might be a premise to the effect that entities which are able to create ends for themselves are ends in themselves. It is the fact that we human beings in some way have a part of us in the transcendental free will, which turn us from mere animals into human beings and persons and ends in ourselves. In Kant's view, animals are not persons even if they happen to have both perceptions, desires, and feelings of pleasure and pain. In some kind of reasoning like this, Kant reaches the second formulation of the categorical imperative, a formulation which supplies us with the *matter* of the transcendental will: *Act so that you treat*

---

[38] These concepts are explicitly used in *Critique of Pure Reason*, p 633.

[39] Broad, *Five Types of Ethical Theory*, p 120.

[40] Kant, *Foundations of the Metaphysics of Morals*, Liberal Arts Press: Indianapolis 1959 (transl. L.W. Beck), p 39 and 54.

humanity, whether in your own person or in that of another, always as an end and never as a means only.[41]

(Out of these two imperatives Kant then deduces his third formulation: *viz., the idea of the will of every rational being as making universal law*.[42] The third formulation means that one should always act as if one were a legislating member in 'the universal kingdom of ends'. This formulation will be left wholly out of account here.)

In itself, in the transcendental realm, the categorical imperative is merely a part of the activity of practical reason. It is a spontaneously created willing, and it cannot exist apart from such a willing. Since the transcendental will is pure will, it automatically and immediately conforms to the law it itself creates in its willing. Maxims, in contradistinction to the moral law, are concerned with the sensible world.[43] Without feelings and desires in this world directed at other things in the world, there would be no maxims like 'Help others!' and 'Don't make false promises!'. Without human beings with feelings and desires there would be nothing for the categorical imperative to test.

When practical reason has no sensibly rooted maxims to evaluate, it is in the same kind of predicament as the faculty which creates the categories, the faculty of understanding, is in when it has no sensations to structure. There is then only *empty* thinking. Kant's famous saying that ”Thoughts without contents are empty, intuitions without concepts are blind”[44] has a counterpart in his practical philosophy. He could also have written that 'the categorical imperative without maxims is empty, maxims without the categorical imperative are morally blind'. These two 'emptinesses', however, are not absolutely empty. The empty *categorical imperative* (created by practical reason) is even in its emptiness wholly distinct from every empty *category* (created by the understanding). In itself, the categorical imperative delivers us the idea of Oughts, i.e. the completely general idea of a moral imperative. According to Kant, such an idea can never arise in nature. Of course, nature contains desires which give rise to imperatives, but nature cannot give rise to *moral* imperatives.

The faculty of reason can only reach sensible nature indirectly. It can only work with material supplied to it by the faculties of understanding and judgment, i.e. it can only will with regard to concepts, judgments, and principles. Therefore, a necessary prerequisite for a this-worldly-ethics is that there is a language in which maxims can be formulated. If a sentiment tells me 'I want this!', reason cannot interfere directly since the term 'this' refers to a spatiotemporal particular, not to a concept. Maxims must have the form 'Do D!' or 'Don't do D!', where 'D' describes a generic action or deed.

If we fuse the first two formulations of the categorical imperative, we get both the form and the matter of the categorical imperative in one single formulation: *Act only according to that maxim by which you can at the same time will that it should become a universal law for all persons, i.e. a law for all beings who are ends in themselves*.

Two remarks are important. First, according to the categorical imperative one should not merely be able to *think* a maxim as a law, one should also be able to *will* it. Second, Kant interchangeably talks of *universal* law and *natural* law, but the point is all the time that, *in thought*, one treats persons counterfactually *as if* they have no free will. In my opinion, his

---

[41] Ibid. p 47 and 54.

[42] Ibid. p 49 and 55.

[43] This may not be wholly true. There may be maxims which try to connect our empirical egos in the sensible world with the noumenal world. An example would be 'Believe in God!'. The idea of God refers to the world of noumena, and the maxim is in this sense not concerned with the sensible world. As far as I know, however, Kant did not discuss such maxims, and I will write as if the imperative 'Believe in God!' is not a real maxim.

[44] *Critique of Pure Reason*, p 93.

idea comes out more clearly in the following formulation: *Act only according to maxims which you can will are collectively willable by all persons*.[45]

So much for Kant himself. Since the aim of this paper is to investigate the possibilities of a modern ethics, i.e. a wholly this-worldly ethics, Kant has to be de-transcendentalized if he shall be of any real interest. Happily, what such a transformation would amount to is easy to tell. It consists of two steps. In the transcendental realm we find understanding, reason and human freedom. In terms of today's philosophy, Kant places conceptual language (which we today can contrast with body language and language by gestures) and agency outside the spatiotemporal world. We, on the other hand, shall as our first step away from the transcendental realm, place conceptual language and agency *within* this world.

Kant identifies nature with the spatiotemporal world of ours, but we should not. Like Kant we should say that conceptual language and agency do not belong to nature, but we should *not*, like Kant, identify nature and the spatiotemporal world. We need to make a distinction between nature and non-nature within the reality of the spatiotemporal world. For the purposes at hand, it is enough to say that nature is that part of reality which is studied by the natural sciences, and that conceptual language and agency belong to the non-nature part of spatiotemporal reality. In particular, they belong to social reality.

The second step away from transcendentalism consists in the claim that something like Kant's reason can exist within the spatiotemporal world. Personally, I would like to bring in Mead's theory about the emergence of the human self, and the human self as something which have both biological and intersubjective conditions of existence,[46] but I will rest content with merely making this statement.

In today's contrast between human beings and other primates, we think of ourselves as having a kind of self-consciousness which the other primates lack. We seem to have a kind of reflexive consciousness which is possible only for linguistic animals. This kind of consciousness, furthermore, is a necessary condition for the existence of moral agency. We do not regard the other known primates as animals capable of morality. An animal which cannot possibly explicitly ask itself 'Ought I do this?' cannot possibly be a moral agent susceptible to moral gratification and reproach. In order to understand the question 'Ought I do this?', one needs to understand not only what 'ought' means. It is equally necessary to understand 'I'. The point I have now tried to make, has been concisely made by Charles Taylor.

> To be a moral agent is to be sensitive to certain standards. But 'sensitive' here must have a strong sense: not just that one's behaviour follow a certain standard, but also that one in some sense recognize or acknowledge the standard.
>
> Animals can follow standards in the weaker sense. My cat will not eat fishmeal below a certain quality. With knowledge of the standard I can predict his behaviour. But there need be no recognition here that he is folowing a standard. This kind of thing, however, would not be sufficient to attribute moral action to an agent. We could imagine some animal who was systematically beneficient in his behaviour; what he did always redounded to the good of man and beast. We still would not think of him as a moral agent, unless there were some recognition on his part that in acting this way he was following a higher standard. ...
>
> Moral agency, in other words, requires some kind of reflexive awareness of the standards one is living by (or failing to live by). ... I think we can say that being a linguistic animal is essential to one's having these concerns;[47]

---

[45] The formulation 'collectively willable' is taken from H. Sobel's paper "Kant's Compass", *Erkenntnis 46* (1997), p 368, although this paper in other ways contradicts some of the views I put forward.

[46] G.H. Mead, *Mind, Self, and Society*, Chicago University Press: Chicago 1967.

[47] C. Taylor, *Human Agency and Language. Philosophical Papers 1*, Cambridge UP: Cambridge 1985, p 102-03. Cf. also pp 260-63 where he distinguishes between standards *an sich* and standards *für sich*; moral standards have to be of the latter kind.

There is something to Kant's analysis of the origin of the moral Ought even if we drop all the transcendental and transcendent aspects of Kant's analysis. If obligations are man-created, but not created wholly by our feelings, desires, or sentiments, then, in a sense, they must be created by something beyond nature. They belong to language, language belongs to social reality, and social reality presupposes but is not identical with nature. Moral Oughts originate outside nature.

6. Kant: It is Impossible to Will to be a Free Rider

Both in Kant's own formulations, and in my condensed reformulation, the categorical imperative puts forward only a *necessary* condition (it may be called 'person-universalization') for a maxim being a duty. A maxim which passes the test is merely *permitted*. Such a maxim of course allows for free riders, but Kantian ethics does not. How, then, did Kant get a *sufficient* condition for a maxim being a real deontological duty, a duty which rules out all possibilities to make oneself a free rider? Like many other commentators, I think that Kant did not succeed in making all his intuitions clear, so what now follows is partly an attempt at a rational reconstruction of the steps which took Kant from the transcendental realm and the categorical imperative down this world and its duties.

That a certain maxim 'Do A!' passes Kant's universalization test and is permitted to act on does not mean that one is prohibited to act on the counter maxim 'Don't do A!'. Both may be permitted! For some maxims it is even the case that if 'Do A!' is prohibited by the categorical imperative, one *cannot* infer that its opposite 'Don't do A!' must be permitted. There are cases where one is prohibited to act on both 'Don't do A!' and 'Do A!'; see case (ii) below.[48] Kant should have said that in order to find out if a proposed maxim constitutes a duty or not, one has to make the test *both* on the maxim *and* on its negation. If a maxim is a duty it must (a) be *possible to will it as being collectively willed*, and it must (b) be *impossible to will its negation as being collectively willed*.[49] The maxim must be permitted and its negation must be prohibited by the categorical imperative.

There are four logically possible results of a simultaneous test of the maxims 'Do A!' and 'Don't do A!': (i) both the maxims are permitted, (ii) both of them are prohibited, (iii) the positive maxim is permitted and the negative one prohibited, and conversely, (iv) the positive maxim is prohibited and the negative maxim is permitted. Only the last two cases give rise to duties, but let us take a look at all four cases, anyhow.

When, (i), both the maxims are permitted, then both of them can be willed as being willed by all persons. '*Can* be willed' does not mean '*need necessarily* be willed'. For instance, if the maxims are 'Color your hair red!' and 'Do not color your hair red!', respectively, then both of them *can* be willed as being willed by everybody. I can will that everybody wants to color their hair red, but I can also will that everybody wants definitely not to color their hair red. Both the maxim and the counter-maxim pass the test, and both ways of acting are permitted. There is no duty to be found in relation to hair-coloring.

The simplest examples of case (ii) - both the maxims are prohibited - are maxims where the description of the action is in itself either a tautology or a contradiction. Let 'Do A!' be the *whole disjunction* 'Color your hair red *or* do not color your hair red!' or the *whole*

[48] Such odd cases must be expelled from formal deontic logic where $\neg P \neg A \equiv OA$; 'P' means permitted and 'O' means obligatory.

[49] R. Norman has made the point that the application of the categorical imperative "is primarily a *negative* test" (*The Moral Philosophers,* Clarendon Press: Oxford 1983, p 104), but he has not realized the need for 'the double test' I am proposing.

conjunction 'Color your hair red *and* do not color your hair red!'. Orders to implement or to refrain from implementing tautologies and contradictions are equally *impossible to will for a logically enlightened will*.[50] Since it is logically impossible to do anything at all in order to get a tautology or a contradiction as a specific state of affairs in the world, it is impossible for a rational being to will that anybody (including himself) should have such a will.

When, now, we come to those cases - (iii) and (iv) - where duties arise, I shall take as examples the four this-worldly duties which Kant himself offers his readers in *Foundations of the Metaphysics of Morals*.[51] Let us start with the maxim-pair 'Develop your talents!' and 'Don't develop your talents!'. Nothing peculiar seems to happen if I will that everybody wants to develop their talents. As persons they can will this just as they can will to color their hair red, and I can will that they all will it if they like to. Of course, this maxim brings with it all the ordinary paradoxes of freedom. The development of my talents may interfere with the development of yours, and vice versa. But these paradoxes are beside the point. They can be taken care of in specific formulations; they constitute demands for specification, and such specifications do not make the *unspecified* maxim non-universalizable. Therefore, the positive maxim is permitted.

What, then, happens if I try to will that everybody (including myself) chooses to neglect their talents? As Kant himself admits, a world where nobody *in fact* cares for their talents is quite possible. This means that person-universalization can be *thought*, but can it be *willed*? Or, rather, can it be willed by an enlightened will? Such a will knows that, as a matter of fact, all human beings inevitably seek happiness. It also knows that in order to will an end one has to will some means for the end. Now, one's talents are means for one's happiness. Therefore, an enlightened will cannot will that it itself, or other persons, will not develop their talents. Consequently, one is prohibited to act on the maxim 'Don't develop your talents!', and since the positive maxim 'Develop your talents!' is permitted, this latter maxim is a real duty. A duty which admits of no exceptions and free riders.

In order to understand this argumentation, one should bear in mind that Kant distinguishes between *three* kinds of imperatives: (i) categorical, necessary, and moral; (ii) hypothetical, assertorical, and pragmatic; (iii) hypothetical, problematical, and technical.[52] The necessary imperatives are to be found in the transcendental realm, but, in fact, there is one and only one, *the* categorical imperative. Both assertorical and technical imperatives prescribe means for a given end; they are both of them *hypothetical* in the sense that they have the form 'If G is the goal, then do D!'. The difference is that in assertorical imperatives the goal, as a matter of fact, is inevitable, whereas in problematical imperatives the goals are wholly contingent. The imperative '*If* you want more happiness *then* develop your talents!' states, I think, in Kant's view, an assertoric imperative since we inevitably seek happiness, whereas '*If* you want to become a good runner *then* train running a lot'. In the assertoric imperative, the if-clause can be taken away and we are left with a formal categorical imperative 'Develop your talents!'. Out of problematical imperatives no duties can be generated.

In passing, I want to say that Kant, in contradistinction to Hume, seems to be a psychological egoist. Happiness he defines as the satisfaction of self-regard ("Selbstsucht").[53]

A test of the maxim-pair 'Help others!' and 'Don't help others!' gives the same kind of result as the last example. There is no problem with willing that everybody should will to help each other. It promotes the happiness everybody wants to have. But there is a problem with

---

[50] In deontic logic, different logical systems handle tautological and contradictory actions in different ways. See for instance, D. Føllesdal & R. Hilpinen, 'Deontic Logic: An Introduction', in Hilpinen (ed.), *Introductory and Systematic Readings*, Reidel: Dordrecht 1971, p 1-35.
[51] *Critique of Practical Reason*, p 39-42.
[52] Ibid. p 31-32, note the footnote 4.
[53] Ibid. p 76.

willing the opposite. If I will that no-one wants to help any other, I negate my own search for happiness, and I negate other's similar searching. And that is impossible to will for a rational person in this world. I can will the positive maxim as being collectively willed, but I cannot will the negative one as being collectively willed. Therefore, 'Help others!' represents a duty.

I now turn to the kind of cases where negative maxims become duties. The first pair of maxims to be tested is 'Make false promises if you need to!' and 'Don't make false promises even if you need to!'. Nothing peculiar happens if I try to will that nobody ever in any situation wants to make a false promises. However, if I try to will that everybody wants to make false promises when needed, then something does happen. In order to *will* the person-universalization, I have to be able to *think* it as a state of affairs existing in this world. But as a logically enlightened person I cannot do that. In a world where everybody knows that everybody will make false promises as soon as they need to, no one will trust anyone. Promises, however, presuppose trust, which means that in such a world it is not possible to make a promise. If we person-universalize the maxim 'Make false promises if you need to!' it destroys itself. Kant says that "Some of our actions are of such a nature that their maxim cannot even be *thought* as a universal law of nature without contradiction."[54] False promises cannot even be thought as universal; nor can they, therefore, be willed. This being so, the maxim 'Don't make false promises even if you need to!' is a duty.

Kant was well aware of the distinction between the distributive and the collective sense of 'all'.[55] Some properties and actions have the feature that though it may be possible for all (i.e. *each)* members of a set *distributively* to have it, all the members cannot *collectively* have it. Winning a race is one example. All who start can win, but all of them cannot win. Likewise, it is possible for some persons to act on the maxim 'Give gifts, but refuse to receive any', but it is logically impossible for all persons to act on it. If everybody refuses to receive gifts, then nobody can give a gift. A logically enlightened person cannot *think* this maxim universalized. The one who believes he can think it commits a logical fallacy, the so-called fallacy of composition.[56] When one tries to universalize the maxim 'Make false promises if you need to!', one ends up with a description of an impossible state of affairs.

It is astonishing that a philosopher of Kant's stature does not realize that he cannot rest content here. He should of course also have discussed whether or not it is possible to think and to will the maxim 'Create a world without promises!' as universalized. Obviously, it can be thought. (This remarkable simple failure was recognized already by Hegel.[57]) There is nothing *logically* wrong with thinking that everybody wants a world where nobody makes promises whether they need it or not. One may, though, argue that it is impossible to make the corresponding *willing*. Such a willing would negate our inevitable search for happiness. The existence of promises is a very important means for increased happiness.

Kant exemplifies negative duties with yet another example, 'Don't commit suicide!'. It is easy both to think and to will that no-one wants to commit suicide. It must be *permitted* to act on this maxim. But if we person-universalize the counter-maxim 'Commit suicide!', it will no longer be possible to act on this maxim. Why? Answer: For the simple reason that then there is no-one around any more. If everybody wants to take their own life and does it, the maxim is in a (peculiar) sense canceled. Therefore, the counter maxim is not *thinkable* as universalized. Again, it is astonishing that Kant did not realize that, next, he should have asked whether it is possible to universalize the maxim 'Create a world without human

---

[54] Ibid. p 41-42.
[55] See for instance *Critique of Pure Reason*, p 494.
[56] For a modern and interesting discussion of this fallacy, see J. Elster, *Logic and Society*, John Wiley & Sons: Chichester 1978, p 97-106. I do not however think that Elster gets his distinction between logical and conceptual impossibilities quite right.
[57] I am relying on Ch. Taylor, *Hegel*, Cambridge UP: Cambridge 1975, p 371.

beings!'.[58] Such a world can be thought, but perhaps not willed, since it contradicts the search for happiness.

When Kant tries to prove his examples of positive duties, 'Develop your talents!' and 'Help others!', he (more or less) explicitly relies on the premise that in this world all persons seek happiness, and that we should take this into account. These duties, by the way, are called *imperfect duties*. Their corresponding negative maxims are impossible to *will* as being universally willed. However, when he tries to prove his examples of negative duties, 'Don't make false promises even if you need to!' and 'Don't commit suicide!', he argues that the corresponding positive maxims are impossible to *think* as being universally willed. These duties are called *perfect duties*.[59] But here, as just remarked, he fails completely. Even here he has to rely on the impossibility of willing the counter-maxim; and this impossibility, in turn, relies on the premise that all persons seek happiness. Really, the two negative duties rely as much on a fact about human nature (we seek happiness) as the two positive duties do.

I will also take the opportunity to comment on two further maxims, and show the fruitfulness of my way of handling the universalization test. Let us look at 'Love your enemies!' and 'Help the poor!'. The first one was put forward F.H. Bradley,[60] and the second by Charles Taylor,[61] to show the incredibility of Kant's way of testing maxims. These maxims are as impossible to *think* as universalized as the maxims 'Make false promises!' and 'Commit suicide!' were. If everybody loves their enemies there will be no enemies, and the maxim is made inapplicable. Similarly, if everybody helps the poor there will be no poor, and the maxim is inapplicable. This fact, however, does not turn the maxims 'Don't love your enemies!' and 'Don't help the poor!' into duties. The last two maxims are impossible to *will* as universalized. That goes against the pursuit of happiness since you may imagine a situation in which you are hated as an enemy but want to be loved, and a situation in which you are poor but want to be rich. We are back in the curious kind of case, (ii), where both a maxim and its counter maxim is prohibited. This must mean that when we love an enemy or help someone who is poor we cannot (if rational) be acting on a maxim. If Kant had realized the point I have made, he might have said that love and charity are necessarily related only to *particular persons* in space and time. With regard to love and charity there can be no *duties*. But that does not make love and charity uninteresting from a happiness point of view.

I think this analysis shows that even those who accept Kant's categorical imperative have to rely on a philosophical anthropological hypothesis in order to get any duties. The relevant impossible willings become impossible only in a world where everybody is searching for happiness. Kant's presumed solution to the free rider problem rests on two presuppositions, one is about human nature and the other about the transcendental will. Both presuppositions can, and ought to be, questioned.

---

[58] It is astonishing because in the *Critique of Pure Reason* Kant very explicitly states that "To posit a triangle, and yet to reject its three angles, is self-contradictory; but there is no contradiction in rejecting the triangle together with its three angles. The same holds true of an absolutely necessary being. If its existence is rejected, we reject the thing itself with all its predicates; and no question of contradiction can then arise." (p 502). To will universal promise-breaking as posited is self-contradictory, but there is no contradiction in rejecting promises altogether.

[59] The distinction between perfect and imperfect duties are to be found in *Critique of Practical Reason*, p 39.

[60] Bradley, *Ethical Studies*, Clarendon Press: Oxford 1988, p 155.

[61] Taylor, *Hegel*, Cambridge UP: Cambridge 1975, p 371.

7. Kant: Self-Subordination is Possible

Kant attempted to ground morals in a will cut loose both from pleasures and pains and from desires and aversions. A transcendental will wills only the categorical imperative: *Act only according to that maxim by which you can at the same time will that it should become a universal law for all persons, i.e. a law for all beings who are ends in themselves*. This formulation, however, contains concepts which refer to the spatiotemporal world. In the transcendental realm we find neither maxims nor acting persons. In its purity and emptiness, the categorical imperative is merely the moral Ought in itself. Transcendental willing and the moral Ought is for Kant one and the same thing. Now, one may ask a question similar to the old theological one whether God's will is subordinated to the concept of Good or whether his willing constitutes the concept of Good. Is the transcendental will subordinated to the categorical imperative or does the transcendental will constitute the categorical imperative? That is a question I would like to have been able to ask Kant, but now I have to try to handle it on my own.

The concept of self-subordination which I regard as necessary in a modern moral philosophy, might be useful in the theological and in the Kantian case, too.[62] Perhaps God wants to subordinate himself to his own good will, and perhaps the transcendental will wants to subordinate itself to its own imperative. However, be that as it may, in a secularized moral philosophy the moral person has in some way to subordinate himself to a freely chosen norm system.

If my interpretation of Kant is at least broadly correct, then Kant is looking for the kind of will which Hume found when he was analyzing promises. Hume found *a will which wanted to be subordinated to an obligation which it itself creates*, and he found such a will unintelligible. I did agree. This is *direct* self-subordination, and since Kant does not discuss it I now leave it, but I remind the reader of the comments I made in relation to Hume in section 2. Instead of Kant's weary transcendental way, we will look at his views on self-subordination in the spatiotemporal world.

When Kant, in *Foundations of the Metaphysics of Morals*, tried to explain his concept of duty, he said that ”Duty is the necessity of an action executed from respect for law”.[63] Before he made this claim, he had made it perfectly clear that he thinks that there is one and only one thing which, without qualification, is good, and that is a good will.[64] A will which wants to conform to the categorical imperative is a good will, and nothing else can make a will a good will. Sentiments and feelings can neither add nor subtract to the goodness we ascribe to a person. That was important for Kant. Is he then not contradicting himself when, as in the sentence just quoted, he grounds duty in a *feeling* of respect? Kant noted this possible objection, and tried to counter it, but this he did in the *Foundations*, remarkably enough, only in a footnote. He said:

> It might be objected that I seek to take refuge in an obscure feeling behind the word ”respect,” instead of clearly resolving the question with a concept of reason. But though respect is a feeling, it is not one *received* through any [outer] influence but is self-wrought by a rational concept; thus it differs specifically from all feelings of the former kind which may be referred to inclination or fear. What I recognize directly as a law for

[62] Whether or not my concept of ”self-subordination” is identical with Kant's concept of ”self-legislation” I will not try to decide. The two concepts are of course closely related, but ”self-legislation” takes away some of the connotations that I want to retain. Even if self-legislation implies a *will* to self-subordination, self-legislation may perhaps exist without self-subordination. And vice versa: even if self-subordination always means that it is oneself who chooses to obey a law, one may not oneself originally have created the law.

[63] Kant, *Foundations of the Metaphysics of Morals*, op. cit., p 16. I have consciously chosen a translation where *Achtung* is translated with 'respect' instead of the other alternatives, 'reverence' and 'awe'.

[64] Kant, *Foundations of the Metaphysics of Morals*, op. cit., p 9.

myself I recognize with respect, which means merely the consciousness of the submission of my will to a law without the intervention of other influences on my mind. The direct determination of the will by the law and the consciousness of this determination is respect; thus respect can be regarded as the effect of the law on the subject and not as the cause of the law. Respect is properly the conception of a worth which thwarts my self-love. Thus it is regarded as an object neither of inclination nor of fear, though it has something analogous to both. The only object of respect is the law, and indeed only the law which we impose on ourselves and yet recognize as necessary in itself. As a law, we are subject to it without consulting self-love; as self-imposed on us by ourselves, it is a consequence of our will. In the former respect it is analogous to fear and in the latter to inclination. All respect for a person is only respect for the law (of righteousness, etc.) of which the person provides an example.[65]

In his *Critique of Practical Reason*, published three years later, Kant discusses respect at more length.[66] He is now no longer reluctant to call respect a feeling, but he calls it a *moral* feeling, and he claims that it is the one and only self-produced (self-wrought) feeling. Also, he claims that respect is caused by the causality of freedom, and that such a causation is necessary for morality to enter the sensible world. Our free will gives rise to respect for the moral law, and this respect may sometimes counter our sensuous desires. When this is the case, we have (in my terminology) self-subordination. One part of ourselves (sensuous inclination) becomes subordinated to another part (our will) by means of a feeling (respect). It is *indirect* self-subordination.[67]

At the end of section 5, I took the liberty to de-transcendentalize Kant's view that the moral Ought originates outside the sensuous world. This transformation means that the spatiotemporal sensuous world comes to contain not only nature, but conceptual language, agency, and reflective reason as well. In what follows, I will write *as if* Kant talked about a will and a reason which wholly belong to our spatiotemporal world. Such a de-transcendentalization turns some of Kant's analyses into phenomenological analyses. Thereby, the philosophical problem of how to combine agency with physical determinism (or indeterminism) is put within that parenthesis where it always exists in courts and everyday moral contexts.

The problem of self-subordination is, I would like to add, not dissolved by a claim that self-reflective agency is intersubjectively constituted, e.g. in the way G.H. Mead assumed. Even if agency is so constituted, which I happen to think, the problem of the free rider remains. Even intersubjectively constituted selves are able to ask themselves: Why should *I* subordinate myself to this morally perfect norm? However, back to Kant.

I am quite convinced that Kant has captured something of phenomenological importance in his analysis of respect, something which, as far as I know, has not been noted by any other philosopher. Surely, not by Hume when he analyzed respect.[68] We may remind ourselves here of Hume's view that willings cannot give rise to any sentiments and feelings; ”we can naturally no more change our own sentiments, than the motions of the heavens”. For Kant this is possible, but only in *one* kind of case, that of respect. There are, according to Kant, two kinds of feelings, self-produced feelings and naturally caused feelings, i.e. sensuous feelings; respect is of the former kind.

In contradistinction to natural feelings, *respect is in some sense created by our own will*. This is the first peculiarity of respect which Kant notes. The second peculiarity is that, as

---

[65] Op. cit. p 17-18.

[66] Kant, *Critique of Practical Reason*, first part, first book, chapter 3.

[67] My interpretation of Kant is in line with Richard McCarty's criticism of the so-called *intellectualist* view of Kantian moral motivation. McCarty calls his own interpretation an *affectivist* view. He explicitly (p 434) leaves out of account, however, what will interest me in what follows, namely the phenomenology of respect. See his paper ”Kantian Moral Motivation and the Feeling of Respect”, *Journal of the History of Philosophy* vol. XXXI (1993), pp 421-35.

[68] *Treatise*, book II, part II, chapter X.

Kant says, respect has *affinities* with both inclination and fear without *being* either. When one respects an imperative it is *as if* there is an inclination or desire to act on it, at the same time as one feels *as if* it were dangerous not to act on it. Normally, one has a positive feeling towards that which one desires and a negative feeling towards that which one fears. Therefore: *Respect for a norm appears as if this feeling were a contradictory unity of a positive and a negative feeling in relation to the same norm*. Since, according to Kant, respect is a feeling which cannot be divided into parts it is not a matter of *mixture* but of *unity*.

The third and last trait of respect which Kant stresses is that *respect is always respect for a law* (or some kind of principle, I would say). This means, among other things, that it is impossible to respect another person *as person*. At bottom, respect for a person is respect for a principle which the person incarnates.

My de-transcendentalization of Kant's analysis of the feeling of respect supplies us with three phenomenological observations:

i) respect appears as a self-produced feeling;
ii) respect appears as a contradictory unity of a positive and a negative feeling for exactly the same thing;
iii) respect appears as respect for a principle.

I shall now try to say something more, and of my own, about these three aspects of respect; I will deal with them bottom up.

Respect is intentional, i.e. it is directed towards something. Respect is always respect *for* something. In that *complex unity* which is the feeling of respect, both a *conative* and an *affective* aspect supervenes on a *cognition* directed at a principle of some kind. Just like a perceived color necessarily has a hue, an intensity, and a degree of saturation, a feeling of respect necessarily has a cognitive object, a certain affection (pleasure/displeasure), and tendencies towards certain ways of acting (=conations). According to Kant, respect can only be directed at the categorical imperative, but this view seems to be wrong even given his own transcendental-philosophical presuppositions. Let us take a look at logic.

As philosophers we can ask ourselves why we should be logical and follow the laws of formal logic. Why, for instance, do we never consciously contradict ourselves? Why do we never claim that something both is and is not the case? There seems to be no sense-given inclination to be logical and, mostly, we have not much to fear if we are a little illogical. In the way Kantians have spoken about duty for duty's sake, we may speak of logic for logic's sake. We *can* be illogical but we do not want to be. Why? The simple answer is, I think, that philosophers are logical, partly out of habit, but partly out of *respect* for logic. At bottom, respect for logic seems to be grounded in respect for *truth-telling and truth-seeking*. A person who contradicts himself does not speak the truth, and logical rules of derivation shall guarantee that truth is transferred from true premises to conclusions. However, this fact does not cancel the point about the feeling of respect which I am trying to bring home. It only means that respect for logic is logically secondary to respect for truth. This also means that in contexts where truth is of no interest, there is no reason to respect the laws of logic either.

Formal logic puts constraints on our speech acts, constraints which we are, no doubt, able to neglect. But out of respect for logic we constrain ourselves. Note now the following distinctions. First, a person may be logical in the sense that he always makes speech acts in accordance with the law of non-contradiction without ever having thought of the law, he then talks *in accordance with* logical principles. Second, a student may try to avoid contradictions only because his teacher in logic has told him to do so; he then makes speech acts *on* logical principles. Such persons, however, are not really logical. Logical in a praiseworthy sense are

only those who have grasped logic and then speaks logically because of this insight. Those who are to be appraised for being logical shall have no other motive than respect for logic. They, alone, make speech acts *for* logical principles.[69] Respect for a logician, in turn, is only respect for 'the laws of which that person gives us an example' (cf. the last Kant-quotation). Kant is of the opinion that neither fear for the moral law nor a sensuous inclination to follow the moral law have any moral worth; only respect for the moral law has such a worth. In my view, we can similarly say that neither fear of logic nor sensuous inclination to follow the laws of logic, but only respect for logic, can give a *speech act* logical *worth*.

It is, I want to stress, possible for a person to study formal logic and to cognize its necessary relations, but simply neglect these relations when he is making speech acts. A logical person makes two things simultaneously. He both cognizes necessary relations and respects them. Respect is a future-directed and action-constraining feeling, but it is not in itself a moral feeling. Respect becomes moral only when its cognitive aspect has a moral content. However, in all cases where respect appears, be it logic or morals, *indirect* self-subordination is at hand. In contradistinction to the concept of direct self-subordination, the concept of *indirect* self-subordination is *not* a contradiction in terms.

People now and then speak of the necessity of having respect for the (legal) law, but equally often they speak of respect for persons. Is it then phenomenologically true that, as Kant said: ”all respect for a person is only respect for the law (of righteousness, etc.) of which the person provides an example”? Those familiar with political organizations and democratic political life can, I think, easily understand what I am going to say. But, probably, my remarks are just as applicable to a lot of other kinds of organizational activities. There are two seemingly different situations I want to highlight.

First, you are member of a political party in which a leader of your own taste has been elected. He does a god job, he has a keen eye for how to apply your ideology in very different circumstances. You begin to respect him. But what is it then exactly that you respect? In my view, it is mainly *the ideological principles which he embodies*. All members of the party (hopefully) accept a list of verbally formulated ideological principles, but few of them can concretize them in the way the leader does. One may *admire* the leader for his concretizing *ability*, but when one *respects* him it is, in fact, the concretized *principles* one respects.

Second, in politics it is not uncommon that one respects a person with whom one disagrees radically in ideological matters; one would never, for instance, vote the way the respected person does. In cases like this, one often stresses that one respects the other *as a person*. Mostly, this means that one respects that the other person has the courage to stand up for his views also when this may be inconvenient and even harmful to him. Oneself, then, thinks that people, including oneself, ought to behave in this way. Beneath the ideological difference between oneself and the respected person there is, in spite of everything, a meta-ideological principle which is shared. When we explicitly respect a person *as a person*, we respect him because he embodies a certain principle. Just like Kant said.

Next, what are we to say about Kant's claim that the feeling of respect is similar to both fear and inclination, but that it is a specific kind of unity all of its own? Respect is respect and nothing else. Is it really possible, at one and the same time, both to fear a norm and to wish to follow it? Note that this problem is not a problem with *contradictory desires*. Buridan's ass had contradictory desires; she wanted two stacks of hay equally strongly. However, a fear for being a norm-breaker and a desire to conform to the norm contain *the same goal and corresponding tendency to act*, namely to act in accordance with the norm. Our problem is situated in the *affective* aspect of the feeling of respect, *not* in its conative aspect. Can there be

[69] This distinction between 'act in accordance with', 'act on', and 'act for' a law is made by Broad in his discussion of Kant in *Five Types of Ethical Theory*, p 119.

feelings and emotions which are unities of a pro-affect and a con-affect directed at exactly the same thing? Must not the concept for such a feeling be a contradiction in terms and, therefore, the corresponding real feeling an impossibility? In order to understand what Kant had a glimpse of, one must realize that there is a whole group of feelings and emotions which deserve the label *contradictory feeling and emotions*.

When we experience a situation or an event as tragi-comical, we have a contradictory emotion. We do not oscillate between being sad because we find the event tragic and being amused because we find the event comic; nor de we have a contradictory desire both to weep and to laugh. We smile, but we smile in a very specific way. The tragi-comical is a phenomena of its own. Next, think of that kind of relationship which is called hate-love relationship. Here, the persons involved do sometimes hate each other and do sometimes love each other. But I think there is more to it. There is also a third emotional state, a contradictory emotional state which can be called exactly 'hate-love'. This is a peculiar state which is related to hate and love as the tragi-comical is related to the tragical and the comical. It is, just like the tragi-comical, a phenomena *sui generis*, even though we in our conceptualization of it have borrowed concepts whose extensions are made up of other kind of phenomena. The Swedish language, by the way, actually has a word, 'hatkärlek', which seems to refer directly to a specific emotion; a literal translation into English would be 'hatelove'. This emotion is captured already in Catullus' famous ”Odi et amo”: ”I hate and I love. Perhaps you ask why I do so. I do not know, but I feel it, and I am in torment.” (from Ode lxxxv).

The one who hates is not necessarily in torment, nor is one who loves, but the one who feels hate-love is necessarily tormented. As I am interpreting Catullus, his view confirms mine which says that ”odi et amo” cannot be reduced to an oscillation between ”odi” and ”amo”.

The aura of peculiarity which seems to surround contradictory feelings and emotions disappears if one notes and understands the real import of two things: (a) feelings and emotions are partly perceptual phenomena, (b) perceptual phenomena may contain the contradictory as a really existing phenomenon.

When we are glad the world looks nice; when we are sad the world looks gray; when we fear X in front of us, X looks fearful. Or something like that. Both feelings and emotions may contain (i) thoughts and cognitions, (ii) a way of being aware of one's body, and (iii) dispositions to act in more or less specific ways. However, what is of interest at the moment is that they also (iv) affect our perceptions. To have a certain feeling or emotion is to perceive the world in a certain way.

Think now of the so-called Müller-Lyer's illusion. It consists of two equally long lines which look as if they are of different length. The remark I want to make I have found in M. Merleau-Ponty's writings:

> The two straight lines in Müller-Lyer's optical illusion are neither of equal nor of unequal length; it is only in the objective world that this question arises. The visual field is that strange zone in which contradictory notions jostle each other because the objects - the straight lines of Müller-Lyer - are not, in that field, in the realm of being, in which a comparison would be possible, but each is taken in its private context as if it did not belong to the same universe as the other. Psychologists have for a long time taken great care to overlook these phenomena. --- We must recognize the indeterminate as a positive phenomenon.
>
> It /science/ requires that two perceived lines, like two real lines, should be equal or unequal, that a perceived crystal should have a definite number of sides, without realizing that the perceived, by its nature, admits of the ambiguous, the shifting, and is shaped by its context. In Müller-Lyer's illusion, one of the lines ceases to be equal to other without becoming 'unequal': it becomes different.[70]

[70] M. Merleau-Ponty, *Phenomenology of Perception*, Routledge & Kegan Paul: London 1962, pp 6 & 11.

In the perceptual world ambiguity is a positive phenomenon. It is *sui generis*; it cannot be reduced away. Sometimes, this kind of ambiguity in perception deserves to be called *contradictory perception*. The lower line in Müller-Lyer's illusion is at one and the same time both longer than and equally long as the upper line.

There are more illusions of the same kind. One is called the waterfall illusion. If you look for a while on an a lively waterfall and then fasten your look on a non-moving object, for instance a stone, it looks as if the stone is both moving and not moving:

> … although the after-effect gives a very clear illusion of movement, the apparently moving features nevertheless seem to stay still! That is, we are still aware of features remaining in their 'proper' locations even though they are seen as moving. What we see is logically impossible![71]

In the waterfall illusion there is *one* distinct but contradictory perception of movement and rest, just as in Müller-Lyer's illusion there is *one* distinct perception of two things being both of equal and of unequal length. In my view, this is the kind of phenomenon which appears both in tragi-comical situations, in 'hate-love', and in the feeling of respect. A tragi-comical situation can be described by means of a *seemingly contradictory* proposition, i.e. as a situation which is both tragic and comical in the same respect. The emotion of 'hate-love' can be described by means of a *seemingly contradictory* proposition, i.e. as an emotion which consists of a feeling of love and a feeling of hate directed at the same person in the same respect. Also, when we respect a principle we have a contradictory feeling; a feeling which is a real state of affairs but which may be described by a seemingly contradictory description. It is a feeling whose object (the principle) is both feared and loved at the same time. In analogy with the construction 'hate-love', respect may be denoted by 'fear-love'.

It is hard to discuss emotions by means of concrete examples since we perceive similar situations differently. Nonetheless, I shall try to show the difference between fear, inclination, and respect by means of an example with three variations. (a) You are driving in your car, far above the speed limit; at the same time as you look upon the speedometer you are seeing that ahead of you there is a speed control with some policemen. You then press the brake and feel fear for having driven too fast. (b) You are driving in your car, far above the speed limit; you look upon the speedometer and when you have done that you take a quick look in the driving mirror at your little child. You then get an inclination to drive at the speed limit and slow down. (c) You are driving in your car, far above the speed limit, and you are thinking about arguments, pro and con, for having speed limits; when the pro-arguments take the lead you look at the speedometer. You then slow down, but not out of fear for anything and nor out of any inclination, but out of respect for the law; you have a feeling of 'fear-love'.

At last, to the last of the three phenomenological features of respect which Kant observed, namely that respect appears as subject to our will in another way than other feelings and emotions are. Here I would like to amend Kant a little. A feeling can be subject to our will in two different ways. One way, of course, is to *produce* the feeling by sheer will power; another way is to be able to *cancel* the feeling by an act of willing even if the feeling cannot be produced at one's will. When it comes to the feeling of respect, my experience is that it can be canceled but not produced by willings. When we feel respect for a principle, we can be described as asking ourselves ”Shall I act in accordance with this principle?” and answering ”Yes, I shall!”. Respect is not produced by merely willing and telling ourselves ”Now I will start to respect A”. I think we can freely deny or consent to feelings of respect, but that this is not possible in relation to feelings and emotions like joy, sorrow, love, and hate. In the short run, the latter seem to be wholly beyond our conscious control. But not so with respect.

---

[71] J. Frisby, *Seeing*, Oxford UP: Oxford 1979, s 101.

Quite independently of the rest of his philosophy, Kant made phenomenological observations of the feeling of respect which, both before and after him, have gone unnoticed by others. The question for me now is whether it is possible to move from the phenomenologist's "appears as" to the ontologist's "is". Can Kant's three observations be re-written in the following way?

i) respect *is* a self-produced feeling;
ii) respect *is* a contradictory unity of a positive and a negative feeling for exactly the same thing;
iii) respect *is* respect for a principle.

The interesting and problematic case is the first statement. In the other two, the phenomenological being of the phenomena spoken of *is* their being. A feeling of respect *is* a contradictory feeling for a principle. In relation to the first statement, however, one may with good reasons ask whether the perceived causality really is an operative causality. Here, Nietzsche and Freud comes to the modern mind. Does not "the hermeneutics of suspicion" tell us that, probably, a feeling like respect, with its curious character, is not at all under the spell of our conscious ego? Is not, we may continue, indirect self-subordination by means of the feeling of respect in fact only a repressed fear for breaking norms that once were established in an authoritarian way? I have no intention whatsoever to try to answer that question here, but there are three remarks I want to make in relation to it. These remarks end this section:

a) The contradictory character of the feeling of respect in no way proves that this feeling must be due to repression, and that it cannot be primordial. Respect is merely one of several contradictory feelings and emotions.
b) If indirect self-subordination by means of respect for norms is not possible (and Kant is wrong), then *reflective* moral agency seems not to be possible.
c) Even if reflective moral *agency* is not possible, no absolute moral chaos will emerge. From a spectator's perspective one can say (as Hume did; see section 4) that men's benevolent impulses will always bring some order into society. In that sense there will always be non-reflective "morals".

8. Summary

This paper is not intended to improve on Hume and Kant. It is merely meant as an evaluation of their attempts to ground a secular ethics. Therefore, I shall end by listing the main findings in the seven sections above:

(1) Hume is right in his claim that psychological egoism is false. But he overstates the case for a general principle of humanity.
(2) Hume is right in his claim that direct self-subordination is impossible, but he is wrong when he takes it for granted that no feeling can be subject to our will.
(3) Hume's causal explanation of virtues seems to be true as such, but it in no way solves the problem of the relativity of virtues.
(4) Hume is on the right track when he takes character formation as a possible way of solving the free rider problem.
(5) Kant is right when he notes that the moral Ought originates outside sensuous nature, but wrong when he thinks that it originates outside the spatiotemporal world. The transcendental derivation of the categorical imperative must be wholly rejected.

(6) Kant’s logical derivation of this-worldly deontological duties fails, and, consequently, he does not solve the free rider problem.

(7) Kant’s phenomenological analysis of respect is correct, and he is right when he makes respect central to moral philosophy, too. Indirect self-subordination by means of respect is perhaps a real possibility.