1   Persons, character and morality

Much of the most interesting recent work in moral philosophy has been of basically Kantian inspiration; Rawls' own work[1] and those to varying degrees influenced by him such as Richards[2] and Nagel[3] are very evidently in the debt of Kant, while it is interesting that a writer such as Fried[4] who gives evident signs of being pulled away from some characteristic features of this way of looking at morality nevertheless, I shall suggest later, tends to get pulled back into it. This is not of course a very pure Kantianism, and still less is it an expository or subservient one. It differs from Kant among other things in making no demands on a theory of noumenal freedom, and also, importantly, in admitting considerations of a general empirical character in determining funda-mental moral demands, which Kant at least supposed himself not to be doing. But allowing for those and many other important differences, the inspiration is there and the similarities both significant and acknowledged. They extend far beyond the evident point that both the extent and the nature of opposition to Utilitarianism resembles Kant's: though it is interesting that in this respect they are more Kantian than a philosophy which bears an obvious but superficial formal resemblance to Kantianism, namely Hare's. Indeed, Hare now supposes that when a substantial moral theory is elicited from his philosophical premises, it turns out to be a version of Utilitarianism. This is not merely because the universal and prescriptive character of moral judgements lays on the agent, according to Hare, a requirement of hypothetical identification with each person affected by a given

---

[1] John Rawls, *A Theory of Justice* (Oxford, 1972).
[2] D. A. J. Richards, *A Theory of Reasons for Action* (Oxford, 1971).
[3] Thomas Nagel, *The Possibility of Altruism* (Oxford, 1970).
[4] Charles Fried, *An Anatomy of Values* (Cambridge, Mass., 1970).

decision – so much is a purely Kantian element. It is rather that each identification is treated just as yielding 'acceptance' or 'rejection' of a certain prescription, and they in turn are construed solely in terms of satisfactions, so that the outputs of the various identifications can, under the usual Utilitarian assumptions, be regarded additively.

Among Kantian elements in these outlooks are, in particular, these: that the moral point of view is basically different from a non-moral, and in particular self-interested, point of view, and by a difference of kind; that the moral point of view is specially characterized by its impartiality and its indifference to any particular relations to particular persons, and that moral thought requires abstraction from particular circumstances and particular characteristics of the parties, including the agent, except in so far as these can be treated as universal features of any morally similar situation; and that the motivations of a moral agent, correspondingly, involve a rational application of impartial principle and are thus different in kind from the sorts of motivations that he might have for treating some particular persons (for instance, though not exclusively, himself) differently because he happened to have some particular interest towards them. Of course, it is not intended that these demands should exclude other and more intimate relations nor prevent someone from acting in ways demanded by and appropriate to them: that is a matter of the relations of the moral point of view to other points of view. But I think it is fair to say that included among the similarities of these views to Kant's is the point that like his they do not make the question of the relations between those points of view at all easy to answer. The deeply disparate character of moral and of non-moral motivation, together with the special dignity or supremacy attached to the moral, make it very difficult to assign to those other relations and motivations the significance or structural importance in life which some of them are capable of possessing.

It is worth remarking that this detachment of moral motivations and the moral point of view from the level of particular relations to particular persons, and more generally from the level of all motivations and perceptions other than those of an impartial character, obtains even when the moral point of view is itself explained in terms of the self-interest under conditions of ignorance of some abstractly conceived contracting parties, as it is by Rawls, and by Richards, who is particularly concerned with applying directly to the characterization

of the moral interest, the structure used by Rawls chiefly to characterize social justice. For while the contracting parties are pictured as making some kind of self-interested or prudential choice of a set of rules, they are entirely abstract persons making this choice in ignorance of their own particular properties, tastes, and so forth; and the self-interested choice of an abstract agent is intended to model precisely the moral choice of a concrete agent, by representing what he would choose granted that he made just the kinds of abstraction from his actual personality, situation and relations which the Kantian picture of moral experience requires.

Some elements in this very general picture serve already to distinguish the outlook in question from Utilitarianism. Choices made in deliberate abstraction from empirical information which actually exists are necessarily from a Utilitarian point of view irrational, and to that extent the formal structure of the outlook, even allowing the admission of *general* empirical information, is counter-Utilitarian. There is a further point of difference with Utilitarianism, which comes out if one starts from the fact that there is one respect at least in which Utilitarianism itself requires a notable abstraction in moral thought, an abstraction which in this respect goes even further than the Kantians': if Kantianism abstracts in moral thought from the identity of persons, Utilitarianism strikingly abstracts from their separateness. This is true in more than one way. First, as the Kantian theorists have themselves emphasized, persons lose their separateness as beneficiaries of the Utilitarian provisions, since in the form which maximizes total utility, and even in that which maximizes average utility, there is an agglomeration of satisfactions which is basically indifferent to the separateness of those who have the satisfactions; this is evidently so in the total maximization system, and it is only superficially not so in the average maximization system, where the agglomeration occurs before the division. Richards,[5] following Rawls, has suggested that the device of the ideal observer serves to model the agglomeration of these satisfactions: equivalent to the world could be one person, with an indefinite capacity for happiness and pain. The Kantian view stands opposed to this; the idea of the contractual element, even between these shadowy and abstract participants, is in part to make the point that

[5] Richards, op. cit., p. 87 al; cf. Rawls, op. cit., p. 27; also Nagel, op. cit., p. 134. This is not the only, nor perhaps historically the soundest, interpretation of the device: cf. Derek Parfit, 'Later Selves and Moral Principles', in A. Montefiore, ed., *Philosophy and Personal Relations* (London, 1973), pp. 149–50 and nn. 30–4.

there are limitations built in at the bottom to permissible trade-offs between the satisfactions of individuals.

A second aspect of the Utilitarian abstraction from separateness involves agency.[6] It turns on the point that the basic bearer of value for Utilitarianism is the *state of affairs*, and hence, when the relevant causal differences have been allowed for, it cannot make any further difference who produces a given state of affairs: if S1 consists of my doing something, together with consequences, and S2 consists of someone else doing something, with consequences, and S2 comes about just in case S1 does not, and S1 is better than S2, then I should bring about S1, however *prima facie* nasty S1 is. Thus, unsurprisingly, the doctrine of negative responsibility has its roots at the foundation of Utilitarianism; and whatever projects, desires, ideals, or whatever I may have as a particular individual, as a Utilitarian agent my action has to be the output of *all* relevant causal items bearing on the situation, including all projects and desires within causal reach, my own and others. As a Utilitarian agent, I am just the representative of the satisfaction system who happens to be near certain causal levers at a certain time. At this level, there is abstraction not merely from the identity of agents, but, once more, from their separateness, since a conceivable extension or restriction of the causal powers of a given agent could always replace the activities of some other agent, so far as Utilitarian outcomes are concerned, and an outcome allocated to two agents as things are could equivalently be the product of one agent, or three, under a conceivable redistribution of causal powers.

In this latter respect also the Kantian outlook can be expected to disagree. For since we are concerned not just with outcomes, but at a basic level with actions and policies, *who* acts in a given situation makes a difference, and in particular I have a particular responsibility for *my* actions. Thus in more than one way the Kantian outlook emphasizes something like the separateness of agents, and in that sense makes less of an abstraction than Utilitarianism does (though, as we have seen, there are other respects, with regard to causally relevant empirical facts, in which its abstraction is greater). But now the question arises, of whether the honourable instincts of Kantianism to defend the individuality of individuals against the agglomerative indifference of Utilitarianism can in fact be effective granted the impoverished and abstract character of persons as moral agents which

[6] For a more detailed account, see 'A Critique of Utilitarianism', in J. J. C. Smart and B. Williams, *Utilitarianism: For and Against* (Cambridge, 1973).

the Kantian view seems to impose. Findlay has said 'the separateness
of persons...is...the basic fact for morals',[7] and Richards hopes to
have respected that fact.[8] Similarly Rawls claims that impartiality does
not mean impersonality.[9] But it is a real question, whether the
conception of the individual provided by the Kantian theories is in fact
enough to yield what is wanted, even by the Kantians; let alone enough
for others who, while equally rejecting Utilitarianism, want to allow
more room than Kantianism can allow for the importance of individual
character and personal relations in moral experience.

II

I am going to take up two aspects of this large subject. They both
involve the idea that an individual person has a set of desires, concerns
or, as I shall often call them, projects, which help to constitute a
character. The first issue concerns the connection between that fact and
the man's having a reason for living at all. I approach this through
a discussion of some work by Derek Parfit; though I touch on a variety
of points in this, my overriding aim is to emphasize the basic
importance for our thought of the ordinary idea of a self or person
which undergoes changes of character, as opposed to an approach
which, even if only metaphorically, would dissolve the person, under
changes of character, into a series of 'selves'.

   In this section I am concerned just with the point that each person
has a character, not with the point that different people have different
characters. That latter point comes more to the fore on the second issue,
which I take up in part III, and which concerns personal relations. Both
issues suggest that the Kantian view contains an important
misrepresentation.

   First, then, I should like to comment on some arguments of Parfit
which explore connections between moral issues and a certain view
of personal identity: a view which, he thinks, might offer, among other
things, '*some* defence'[10] of the Utilitarian neglect of the separateness
of persons. This view Parfit calls the 'Complex View'. This view takes
seriously the idea that relations of psychological connectedness (such
as memory and persistence of character and motivation) are what really

[7] Findlay, *Values and Intentions* (London, 1961), pp. 235–6.
[8] Richards, op. cit., p. 87.   [9] Rawls, op. cit., p. 190.
[10] Parfit, op. cit., p. 160, his emphasis. In what follows and elsewhere in this chapter
    I am grateful to Parfit for valuable criticisms of an earlier draft.

matter with regard to most questions which have been discussed in relation to personal identity. The suggestion is that morality should take this seriously as well, and that there is more than one way of its doing so. Psychological connectedness (unlike the surface logic of personal identity) admits of degrees. Let us call the relevant properties and relations which admit of degrees, *scalar* items. One of Parfit's aims is to make moral thought reflect more directly the scalar character of phenomena which underlie personal identity. In particular, in those cases in which the scalar relations hold in reduced degree, this fact should receive recognition in moral thought.

Another, and more general, consequence of taking the Complex View is that the matter of personal identity may appear altogether less deep, as Parfit puts it, than if one takes the Simple View, as he calls that alternative view which sees as basically significant the all-or-nothing logic of personal identity. If the matter of personal identity appears less deep, the *separateness* of persons, also, may come to seem less an ultimate and specially significant consideration for morality. The connection between those two thoughts is not direct, but there is more than one indirect connection between them.[11]

So far as the problems of *agency* are concerned, Parfit's treatment is not going to help Utilitarianism. His loosening of identity is diachronic, by reference to the weakening of psychological connectedness over time: where there is such weakening to a sufficient degree, he is prepared to speak of 'successive selves', though this is intended only as a *façon de parler*.[12] But the problems that face Utilitarianism about agency can arise with any agent whose projects stretch over enough time, and are sufficiently grounded in character, to be in any substantial sense *his* projects, and that condition will be satisfied by something that is, for Parfit, even *one* self. Thus there is nothing in this degree of dissolution of the traditional self which can help over agency.

In discussing the issues involved in making moral thought reflect more directly the scalar nature of what underlies personal identity, it is important to keep in mind that the talk of 'past selves', 'future selves' and generally 'several selves' is only a convenient fiction. Neglect of this may make the transpositions in moral thought required by the

---

[11] Parfit develops one such connection in the matter of distributive justice: pp. 148ff. In general it can be said that one very natural correlate of being impressed by the separateness of several persons' lives is being impressed by the peculiar unity of one person's life.   [12] Ibid., n. 14, pp. 161–2.

Complex View seem simpler and perhaps more inviting than they are, since they may glide along on what seems to be a mere multiplication, in the case of these new 'selves', of familiar interpersonal relations. We must concentrate on the scalar facts. But many moral notions show a notable resistance to reflecting the scalar: or, rather, to reflecting it in the right way. We may take the case of promising, which Parfit has discussed.[13] Suppose that I promise to A that I will help him in certain ways in three years time. In three years time a person appears, let us say A^*, whose memories, character etc., bear some, but a rather low, degree of connectedness to A's. How am I to mirror these scalar facts in my thought about whether, or how, I am to carry out my promise?

Something, first, should be said about the promise itself. '*You*' was the expression it used: 'I will help *you*', and it used that expression in such a way that it covered both the recipient of these words and the potential recipient of the help. This was not a promise that could be carried out (or, more generally, honoured) by helping anyone else, or indeed by doing anything except helping that person I addressed when I said 'you' – thus the situation is not like that with some promises to the dead (those where there is still something one can do about it).[14] If there is to be any action of mine which is to count as honouring that promise, it will have to be action which consists in now helping A^*. How am I to mirror, in my action and my thought about it, A^*'s scalar relations to A?

There seem to be only three ways in which they could be so mirrored, and none seems satisfactory. First, the action promised might itself have some significant scalar dimension, and it might be suggested that this should vary with my sense of the proximity or remoteness of A^* from A. But this will not do: it is clearly a lunatic idea that if I promised to pay A a sum of money, then my obligation is to pay A^* some money, but a smaller sum. A more serious suggestion would be that what varies with the degree of connectedness of A^* to A is the degree of stringency of the obligation to do what was promised. While less evidently dotty, it is still, on reflection, dotty; thus, to take a perhaps unfair example, it seems hard to believe that if someone had promised to marry A, they would have an obligation to marry A^*, only an obligation which came lower down the queue.

What, in contrast, is an entirely familiar sort of thought is, last of all, one that embodies degrees of doubt or obscurity whether a given

[13] Ibid., pp. 144ff.   [14] Ibid., p. 144 fin.

obligation (of fixed stringency) applies or not. Thus a secret agent might think that he was obliged to kill the man in front of him if and only if that man was Martin Bormann; and be in doubt whether he should kill this man, because he was in doubt whether it was Bormann. (Contrast the two analogously dotty types of solution to this case: that, at any rate, he is obliged to wound him; or, that he is obliged to kill him, but it has a lower priority than it would have otherwise.) But this type of thought is familiar at the cost of not really embodying the scalar facts; it is a style of thought appropriate to uncertainty about a matter of all-or-nothing and so embodies in effect what Parfit calls the Simple View, that which does not take seriously the scalar facts to which the Complex View addresses itself.

These considerations do not, of course show that there are no ways of mirroring the Complex View in these areas of moral thought, but they do suggest that the displacements required are fairly radical. It is significant that by far the easiest place in which to find the influence of the scalar considerations is in certain *sentiments*, which themselves have a scalar dimension – here we can see a place where the Complex View and Utilitarianism easily fit together. But the structure of such sentiments is not adequate to produce the structure of all moral thought. The rest of it will have to be more radically adapted, or abandoned, if the Complex View is really to have its effect.

One vitally important item which is in part (though only in part) scalar is a man's concern for (what commonsense would call) his own future. That a man should have some interest now in what he will do or undergo later, requires that he have some desires or projects or concerns now which relate to those doings or happenings later; or, as a special case of that, that some very general desire or project or concern of his now relate to desires or projects which he will have then. The limiting case, at the basic physical level, is that in which he is merely concerned with future pain, and it may be that that concern can properly reach through any degree of psychological discontinuity.[15] But even if so, it is not our present concern, since the mere desire to avoid physical pain is not adequate to constitute a character. We are here concerned with more distinctive and structured patterns of desire and project, and there are possible psychological changes in these which could be predicted for a person and which would put his future after such changes beyond his present interest. Such a future would be, so to speak, over the horizon of his interest, though of course if the future

[15] Cf. 'The Self and the Future', in *Problems of the Self* (Cambridge, 1973).

picture could be filled in as a *series* of changes leading from here to there, he might recapture an interest in the outcome.

In this connection, to take the language of 'future selves' at all literally would be deeply misleading: it would be to take the same facts twice over. My concern for my descendants or other relatives may be, as Parfit says, to some degree proportional to their remoteness from me; equally, my concern for other persons in general can vary with the degree to which their character is congenial to my own, their projects sympathetic to my outlook. The two considerations, of proximity and congeniality, evidently interact – ways in which they can reinforce or cancel one another are, for instance, among the common-places of dynastic fiction. But the proximity of Parfitian 'later selves' to me, their ancestor, just consists of the relations of their character and interests to my present ones. I cannot first identify a later self 'descendant', and then consider the relations of his character to mine, since it is just the presence or lack of these relations which in good part determines his proximity and even his existence as a separate self.

Thus if I take steps now to hinder what will or may predictably be my future projects, as in Parfit's Russian nobleman case,[16] it would be a case of double vision to see this as my treating my future self as another person, since, spelled out, that would have to mean, treating my future self as another person *of whose projects I disapprove*; and therein lies the double vision. To insist here that what I would be doing is to hinder *my own* future projects (where it is understood that that is not necessarily a foolish thing to do) is to keep hold on a number of deeply important facts. One is that to contemplate, or expect, or regard as probable, such changes in my own character is different from my relation to them in someone else (still more, of course, from my attitude to the mere *arrival* of someone else with a different character). The question must arise, how prediction is, in my own case, related to acquiescence, and special and obscure issues arise about the range of methods that it could be appropriate or rational for a man to use to prevent or deflect predicted changes in his own character. Thought about those issues must take as basic the *he* for whom these changes would be changes in *his* character.

Relatedly, there is the question of why I should regard my present projects and outlook as having more authority than my future ones. I do *not* mean by that the question, why I should not distribute consideration equally over my whole life: I shall later touch on the

[16] Parfit, op. cit., pp. 145ff.

point that it is a mistake of Kantians (and perhaps of some kinds of Utilitarians too) to think it *a priori* evident that one rationally should do that. I mean rather the question of how, in the supposed type of example, I evaluate the two successive outlooks. Why should I hinder my future projects from the perspective of my present values rather than inhibit my present projects from the perspective of my future values? It is not enough in answer to that to say that evidently present action must flow from present values. If the future prospect were of something now identified as a growth in enlightenment, present action would try to hinder present projects in its interest. For that to be so, there indeed would have to be now some dissatisfaction with one's present values, but that consideration just turns attention, in the Russian nobleman case, to the corresponding question, of why the young man is so unquestioningly satisfied with his present values. He may have, for instance, a theory of degeneration of the middle-aged, but then he should reflect that, when middle-aged, he will have a theory of the naiveté of the young.

I am not saying that there are no answers to any of these questions, or that there is no way out of this kind of diachronic relativism. The point is that if it is true that this man will change in these ways, it is only by understanding his present projects *as the projects of one who will so change* that he can understand them even as his present projects; and if he knows that he will so change, then it is only through such an understanding that he could justifiably give his present values enough authority to defeat his future values, as he clear-headedly conceives them to be. If he clear-headedly knows that his present projects are solely the projects of his youth, how does he know that they are not *merely* that, unless he has some view which makes sense of, among other things, his own future? One cannot even start on the important questions of how this man, so totally identified with his present values, will be related to his future without them, if one does not take as basic the fact that it is his own future that he will be living through without them.

This leads to the question of why we go on at all.

It might be wondered why, unless we believe in a possibly hostile after-life, or else are in a muddle which the Epicureans claimed to expose, we should regard death as an evil.[17] One answer to that is that we desire certain things; if one desires something, then to that extent one has reason to resist the happening of anything which

[17] The argument is developed in more detail in *Problems of the Self*, pp. 82ff.

prevents one getting it, and death certainly does that, for a large range of desires. Some desires are admittedly contingent on the prospect of one's being alive, but not all desires can be in that sense conditional, since it is possible to imagine a person rationally contemplating suicide, in the face of some predicted evil, and if he decides to go on in life, then he is propelled forward into it by some desire (however general or inchoate) which cannot operate conditionally on his being alive, since it settles the question of whether he is going to be alive. Such a desire we may call a categorical desire. Most people have many categorical desires, which do not depend on the assumption of the person's existence, since they serve to prevent that assumption's being questioned, or to answer the question if it is raised. Thus one's pattern of interests, desires and projects not only provide the reason for an interest in what happens within the horizon of one's future, but also constitute the conditions of there being such a future at all.

Here, once more, to deal in terms of later selves who were like descendants would be to misplace the heart of the problem. Whether to commit suicide, and whether to leave descendants, are two separate decisions: one can produce children before committing suicide. A person might even choose deliberately to do that, for comprehensible sorts of reasons; or again one could be deterred, as by the thought that one would not be there to look after them. Later selves, however, evade all these thoughts by having the strange property that while they come into existence only with the death of their ancestor, the physical death of their ancestor will abort them entirely. The analogy seems unhelpfully strained, when we are forced to the conclusion that the failure of all my projects, and my consequent suicide, would take with me all my 'descendants', although they are in any case a kind of descendants who arise only with my ceasing to exist. More than unhelpfully, it runs together what are two quite different questions: whether, my projects having failed, I should cease to exist, and whether I shall have descendants whose projects may be quite different from mine and are in any case largely unknown. The analogy makes every question of the first kind involve a question of the second kind, and thus obscures the peculiar significance of the first question to the theory of the self. If, on the other hand, a man's future self is not another self, but the future of his self, then it is unproblematic why it should be eliminated with the failure of that which might propel him into it. The primacy of one's ordinary self is given, once more, by the

thought that it is precisely what will not be in the world if one commits suicide.

The language of 'later selves', too literally taken, could exaggerate in one direction the degree to which my relation to some of my own projects resembles my relation to the projects of others. The Kantian emphasis on moral impartiality exaggerates it in quite another, by providing ultimately too slim a sense in which any projects are mine at all. This point once more involves the idea that my present projects are the condition of my existence,[18] in the sense that unless I am propelled forward by the conatus of desire, project and interest, it is unclear why I should go on at all: the world, certainly, as a kingdom of moral agents, has no particular claim on my presence or, indeed, interest in it. (That kingdom, like others, has to respect the natural right to emigration.) Now the categorical desires which propel one on do not have to be even very evident to consciousness, let alone grand or large; one good testimony to one's existence having a point is that the question of its point does not arise, and the propelling concerns may be of a relatively everyday kind such as certainly provide the ground of many sorts of happiness. Equally, while these projects may present *some* conflicts with the demands of morality, as Kantianly conceived, these conflicts may be fairly minor; after all – and I do not want to deny or forget it – these projects, in a normally socialized individual, have in good part been formed within, and formed by, dispositions which constitute a commitment to morality. But, on the other hand, the possibility of radical conflict is also there. A man may have, for a lot of his life or even just for some part of it, a *ground* project or set of projects which are closely related to his existence and which to a significant degree give a meaning to his life.

I do not mean by that they provide him with a life-plan, in Rawls' sense. On the contrary, Rawls' conception, and the conception of practical rationality, shared by Nagel, which goes with it, seems to me rather to imply an external view of one's own life, as something like a given rectangle that has to be optimally filled in.[19] This perspective

[18] We can note the consequence that present projects are the condition of future ones. This view stands in opposition to Nagel's: as do the formulations used above, p. 10. But while, as Nagel says, taking a rational interest in preparing for the realization of my later projects does not require that they be my present projects, it seems nevertheless true that it presupposes my having some present projects which directly or indirectly reach out to a time when those later projects will be my projects.

[19] It is of course a separate question what the criteria of optimality are, but it is not surprising that a view which presupposes that no risks are taken with the useful area

omits the vital consideration already mentioned, that the continuation and size of this rectangle is up to me; so, slightly less drastically, is the question of how much of it I care to cultivate. The correct perspective on one's life is *from now.* The consequences of that for practical reasoning (particularly with regard to the relevance of proximity or remoteness in time of one's objective), is a large question which cannot be pursued here; here we need only the idea of a man's ground projects providing the motive force which propels him into the future, and gives him a reason for living.

For a project to play this ground role, it does not have to be true that if it were frustrated or in any of various ways he lost it, he would have to commit suicide, nor does he have to think that. Other things, or the mere hope of other things, may keep him going. But he may feel in those circumstances that he might as well have died. Of course, in general a man does not have one separable project which plays this ground role: rather, there is a nexus of projects, related to his conditions of life, and it would be the loss of all or most of them that would remove meaning.

Ground projects do not have to be selfish, in the sense that they are just concerned with things for the agent. Nor do they have to be self-centred, in the sense that the creative projects of a Romantic artist could be considered self-centred (where it has to be *him,* but not *for* him). They may certainly be altruistic, and in a very evident sense moral, projects; thus he may be working for reform, or justice, or general improvement. There is no contradiction in the idea of a man's dying for a ground project – quite the reverse, since if death really is necessary for the project, then to live would be to live with it unsatisfied, something which, if it really is his ground project, he has no reason to do.

That a man's projects were altruistic or moral would not make them immune to conflict with impartial morality, any more than the artist's projects are immune. Admittedly *some* conflicts are ruled out by the projects sincerely being *those* projects; thus a man devoted to the cause of curing injustice in a certain place, cannot just insist on his plan for doing that over others', if convinced that theirs will be as effective as

of the rectangle should also favour a very low risk strategy in filling it: cf. Rawls (on prudential rationality in general), op. cit., p. 422: 'we have the guiding principle that a rational individual is always to act so that he need never blame himself no matter how things finally transpire.' Cf. also the passages cited in Rawls' footnote. For more on this and the relations of ground projects to rationality, see chapter 2, below.

his (something it may be hard to convince him of). For if he does insist on that, then we learn that his concern is not merely that injustice be removed, but that *he* remove it – not necessarily a dishonourable concern, but a different one. Thus some conflicts are ruled out by the project being not self-centred. But not all conflicts: thus his selfless concern for justice may do havoc to quite other commitments.

A man who has such a ground project will be required by Utilitarianism to give up what it requires in a given case just if that conflicts with what he is required to do as an impersonal utility-maximizer when all the causally relevant considerations are in. That is a quite absurd requirement.[20] But the Kantian, who can do rather better than that, still cannot do well enough. For impartial morality, if the conflict really does arise, must be required to win; and that cannot necessarily be a reasonable demand on the agent. There can come a point at which it is quite unreasonable for a man to give up, in the name of the impartial good ordering of the world of moral agents, something which is a condition of his having any interest in being around in that world at all. Once one thinks about what is involved in having a character, one can see that the Kantians' omission of character is a condition of their ultimate insistence on the demands of impartial morality, just as it is a reason for finding inadequate their account of the individual.

III

All this argument depends on the idea of one person's having a character, in the sense of having projects and categorical desires with which that person is identified; nothing has yet been said about different persons having different characters. It is perhaps important, in order to avoid misunderstanding, to make clear a way in which difference of character does *not* come into the previous argument. It does not come in by way of the man's thinking that only if he affirms these projects will they be affirmed, while (by contrast) the aims of Kantian morality can be affirmed by anyone. Though that thought could be present in some cases, it is not the point of the argument. The man is not pictured as thinking that he will have earned his place in the world, if his project is affirmed: that a distinctive contribution to the world will have been made, if his distinctive project is carried forward. The point is that he wants these things, finds his life bound

[20] Cf. 'A Critique of Utilitarianism', sections 3–5.

up with them, and that they propel him forward, and thus they give
him a reason for living his life. But that is compatible with these drives,
and this life, being much like others'. They give him, distinctively, a
reason for living this life, in the sense that he has no desire to give
up and make room for others, but they do not require him to lead
a *distinctive* life. While this is so, and the point has some importance,
nevertheless the interest and substance of most of the discussion
depends on its in fact being the case that people have dissimilar
characters and projects. Our *general* view of these matters, and the
significance given to individuality in our own and others' lives, would
certainly change if there were not between persons indefinitely many
differences which are important to us. The level of description is of
course also vital for determining what is the same or different. A similar
description can be given of two people's dispositions, but the concrete
detail be perceived very differently – and it is a feature of our
experience of persons that we can perceive and be conscious of an
indefinitely fine degree of difference in concrete detail (though it is only
in certain connections and certain cultures that one spends much time
rehearsing it).

One area in which *difference* of character directly plays a role in the
concept of moral individuality is that of personal relations, and I shall
close with some remarks in this connection. Differences of character
give substance to the idea that individuals are not inter-substitutable.
As I have just argued, a particular man so long as he is propelled
forward does not need to assure himself that he is unlike others, in
order not to feel substitutable, but in his personal relations to others
the idea of difference can certainly make a contribution, in more than
one way. To the thought that his friend cannot just be equivalently
replaced by another friend, is added both the thought that he cannot
just be replaced himself, and also the thought that he and his friend
are different from each other. This last thought is important to us as
part of our view of friendship, a view thus set apart from Aristotle's
opinion that a good man's friend was a duplication of himself. This I
suspect to have been an Aristotelian, and not generally a Greek,
opinion. It is connected with another feature of his views which seems
even stranger to us, at least with regard to any deeply committed
friendship, namely that friendship for him has to be minimally
risky – one of his problems is indeed to reconcile the role of friendship
with his unappetizing ideal of self-sufficiency. Once one agrees that
a three-dimensional mirror would not represent the ideal of friendship,

one can begin to see both how some degree of difference can play an essential role, and, also, how a commitment or involvement with a particular other person might be one of the kinds of project which figured basically in a man's life in the ways already sketched – something which would be mysterious or even sinister on an Aristotelian account.

For Kantians, personal relations at least presuppose moral relations, and some are rather disposed to go further and regard them as a *species* of moral relations, as in the richly moralistic account given by Richards[21] of one of the four main principles of supererogation which would be accepted in 'the Original Position' (that is to say, adopted as a moral limitation):

a principle of mutual love requiring that people should not show personal affection and love to others on the basis of arbitrary physical characteristics alone, but rather on the basis of traits of personality and character related to acting on moral principles.

This righteous absurdity is no doubt to be traced to a feeling that love, even love based on 'arbitrary physical characteristics', is something which has enough power and even authority to conflict badly with morality unless it can be brought within it from the beginning, and evidently that is a sound feeling, though it is an optimistic Kantian who thinks that much will be done about that by the adoption of this principle in the Original Position. The weaker view, that love and similar relations presuppose moral relations, in the sense that one could love someone only if one also had to them the moral relations one has to all people, is less absurd, but also wrong. It is of course true that loving someone involves some relations of the kind that morality requires or imports more generally, but it does not follow from that that one cannot have them in a particular case unless one has them generally in the way the moral person does. Someone might be concerned about the interests of someone else, and even about carrying out promises he made to that person, while not very concerned about these things with other persons. To the extent (whatever it may be) that loving someone involves showing some of the same concerns in relation to them that the moral person shows, or at least thinks he ought to show, elsewhere, the lover's relations will be examples of moral relations, or at least resemble them, but this does not have to be because they are *applications to this case* of relations which the lover, *qua* moral person, more generally enters into. (That might not be the best

[21] Richards, op. cit., p. 94.

description of the situation even if he *is* a moral person who enters into such relations more generally.)

However, once morality is there, and also personal relations to be taken seriously, so is the possibility of conflict. This of course does not mean that if there is some friendship with which his life is much involved, then a man must prefer any possible demand of that over other, impartial, moral demands. That would be absurd, and also a pathological kind of friendship, since both parties exist in the world and it is part of the sense of their friendship that it exists in the world. But the possibility of conflict with substantial moral claims of others is there, and it is not only in the outcome. There can also be conflict with moral demands on how the outcome is arrived at: the situation may not have been subjected to an impartial process of resolution, and this fact itself may cause unease to the impartial moral consciousness. There is an example of such unease in a passage by Fried. After an illuminating discussion of the question why, if at all, we should give priority of resources to actual and present sufferers over absent or future ones, he writes:[22]

surely it would be absurd to insist that if a man could, at no risk or cost to himself, save one or two persons in equal peril, and one of those in peril was, say, his wife, he must treat both equally, perhaps by flipping a coin. One answer is that where the potential rescuer occupies no office such as that of captain of a ship, public health official or the like, the occurrence of the accident may itself stand as a sufficient randomizing event to meet the dictates of fairness, so he may prefer his friend, or loved one. Where the rescuer does occupy an official position, the argument that he must overlook personal ties is not unacceptable.

The most striking feature of this passage is the direction in which Fried implicitly places the onus of proof: the fact that coin-flipping would be inappropriate raises some question to which an 'answer' is required, while the resolution of the question by the rescuer's occupying an official position is met with what sounds like relief (though it remains unclear what that rescuer does when he 'overlooks personal ties' – does *he* flip a coin?). The thought here seems to be that it is unfair to the second victim that, the first being the rescuer's wife, they never even get a chance of being rescued; and the answer

---

[22] Fried, op. cit., p. 227. [Note 1981] Fried has perhaps now modified the view criticised here. He has himself used the idea of friendship as creating special moral relations, but in a connexion where, it seems to me, it is out of place: for criticism, see chapter 4, below.

(as I read the reference to the 'sufficient randomizing event') is that at another level it is sufficiently fair – although in this disaster this rescuer has a special reason for saving the other person, it might have been another disaster in which another rescuer had a special reason for saving them. But, apart from anything else, that 'might have been' is far too slim to sustain a reintroduction of the notion of fairness. The 'random' element in such events, as in certain events of tragedy, should be seen not so much as affording a justification, in terms of an appropriate application of a lottery, as being a reminder that some situations lie beyond justifications.

But has anything yet shown that? For even if we leave behind thoughts of higher-order randomization, surely *this* is a justification on behalf of the rescuer, that the person he chose to rescue was his wife? It depends on how much weight is carried by 'justification': the consideration that it was his wife is certainly, for instance, an explanation which should silence comment. But something more ambitious than this is usually intended, essentially involving the idea that moral principle can legitimate his preference, yielding the conclusion that in situations of this kind it is at least all right (morally permissible) to save one's wife. (This could be combined with a variety of higher-order thoughts to give it a rationale; rule-Utilitarians might favour the idea that in matters of this kind it is best for each to look after his own, like house insurance.) But this construction provides the agent with one thought too many: it might have been hoped by some (for instance, by his wife) that his motivating thought, fully spelled out, would be the thought that it was his wife, not that it was his wife and that in situations of this kind it is permissible to save one's wife.

Perhaps others will have other feelings about this case. But the point is that somewhere (and if not in this case, where?) one reaches the necessity that such things as deep attachments to other persons will express themselves in the world in ways which cannot at the same time embody the impartial view, and that they also run the risk of offending against it.

They run that risk if they exist at all; yet unless such things exist, there will not be enough substance or conviction in a man's life to compel his allegiance to life itself. Life has to have substance if anything is to have sense, including adherence to the impartial system; but if it has substance, then it cannot grant supreme importance to the impartial system, and that system's hold on it will be, at the limit, insecure.

It follows that moral philosophy's habit, particularly in its Kantian forms, of treating persons in abstraction from character is not so much a legitimate device for dealing with one aspect of thought, but is rather a misrepresentation, since it leaves out what both limits and helps to define that aspect of thought. Nor can it be judged solely as a theoretical device: this is one of the areas in which one's conception of the self, and of oneself, most importantly meet.