16

Harm and Self-Interest

J. FEINBERG*

THE study of *kakapoeics*, or the general theory and classification of harms, should be a central enterprise of legal philosophy. Most writers agree, after all, that the prevention of harms is a legitimate aim of both the criminal law and the coercive parts of the civil law, though of course there is much disagreement over whether it is the *sole* proper concern of coercive law, over *whose* harms are properly considered, and over which types of harm have priority in cases of conflict. There are also conceptual riddles concerning the scope of the term 'harm', three of which provide the excuse for this essay, namely, whether there can be such things as purely *moral harms* (harm to character), *vicarious harms* (as I shall call them), and *posthumous harms*. My discussion of these questions will assume without argument the orthodox jurisprudential analysis of harm as invaded interest, not because I think that account is self-evidently correct or luminously perspicuous, but rather because I wish to explore its implications for the borderline cases of harm, the better to test its adequacy, and to determine the respects in which the concept of self-interest still needs clarification.

The theory of the nature of harms assumed here can be sketched quickly. A person is harmed when someone invades (blocks or thwarts) one of his interests.[1] A person has an interest in Y when he has a *stake* in Y, that is, when he stands to gain or lose depending on the condition or outcome of Y.[2] A person's interest in the singular

* Professor of Philosophy, The Rockefeller University.

[1] Interests can be blocked or defeated by events in impersonal nature or by plain bad luck. But they can only be 'invaded' by human beings, either oneself, acting negligently or perversely, or by others, singly or in groups and organizations. It is only when an interest is invaded by self or others that its possessor is harmed in the usual legal sense, though obviously an earthquake or a plague can cause enormous harm in the ordinary sense.

[2] Strictly speaking, this definition is circular since a person would probably have to know what it is to have an interest in something before he could know what it is to

(his personal interest or self-interest) consists in the harmonious advancement of all his interests in the plural. We speak not only of the things a person 'has an interest in' but also of the various things that are 'in his interest', that is, the things that promote his interests as a group. 'Welfare interests' are interests in the indispensable means to one's ulterior goals, whatever the latter may be. These include health, financial sufficiency, and the like. 'Ulterior interests' are based on stable, long-range objectives, achievements of goals valued at least partly as ends in themselves—for example producing a book, raising a family, building a dream house, advancing a cause. Characteristically human well-being consists in the advancing of such interests.

Welfare and ulterior interests bear somewhat different relations to wants or desires. Anything we believe we have a stake in, whether it be mere minimal health or ultimate achievement, we will desire to some degree, in so far at least as we are rational. But we have some welfare interests in conditions that are good for us even if we should not want them (for example, health), whereas in respect to our more ultimate goals, we have a stake in them because we desire their achievement, not the other way round. In these instances, if our wants were to change, our interests would too. It is not true, however, that wants, even strong wants, are *sufficient* to create interests. Few non-betting football fans, for example, have ulterior interests in their favourite team's victory, though many may have very intense desires for that outcome. As a psychological generalization, it is probably true that few persons can 'invest' enough in a wanted outcome to create a stake in it unless promoting that outcome becomes a personal goal or objective. Surely, no mere 'desire of the moment', like a desire to go to the cinema,[3] can generate an ulterior interest, but only a relatively deep-rooted and stable want whose fulfilment can be both reasonably hoped for (mere idle wishes won't do) and influenced by one's own efforts.[4]

'gain or lose' as well as vice versa. But even a circular definition can have some practical utility in providing an equivalent expression for the *definiendum* that is more easily manipulated to good purpose, or which is more suggestive, or productive of insight. The word 'stake', e.g. brings out with intuitive vividness the connection between interests and risks. The word 'stake' has its primary or literal use to refer to 'the amount risked by a party to a wager, or match, or gambler, a thing whose existence, or safety, or ownership depends on some issue'.

[3] Cf. Brian Barry, *Political Argument* (Routledge & Kegan Paul, London, 1965), p. 183.

[4] There is, I suppose, a respect in which anyone who has a strong desire for anything

1. *Moral harm.* Is interest, then, a wholly 'want-regarding' concept, or does the analysis sketched in the preceding paragraphs leave out something important? The label 'want-regarding' comes from Brian Barry[5] who contrasted it with what he called 'ideal-regarding' concepts and principles. A concept is want-regarding if it can be analysed entirely in terms of the 'wants which people happen to have', whereas it is ideal-regarding if reference must also be made to what would be ideal, or best for people, their wants notwithstanding, or to the wants they ought to have whether they have them in fact or not. The ideal-regarding theory of interest holds that it is in a person's interest ultimately not only to have his wants and goals fulfilled, but also (and often this is held to be more important) to have his tastes elevated, his sensibilities refined, his judgment sharpened, his integrity strengthened: in short to become a better person. On this view, a person can be harmed not only in his health, his purse, his worldly ambition, and the like, but also in his character. One's ultimate good is not only to *have* the things one wants, but (perhaps more importantly) to *be* an excellent person, whatever one may want. We not only degrade and corrupt a man by making him a worse man than he would otherwise be; on this view, we inflict serious *harm* on him, even though all his other interests flourish. Socrates and the Stoics even went so far as to hold that this 'moral harm' is the *only* genuine harm. Epictetus was so impressed with the harm that consists simply in having a poor character that he thought it redundant to punish a morally depraved person for his crimes. Such a person is punished enough, he thought, just by being the sort of person he is.

To a certain extent, the conflict between the two accounts of interest is entirely academic. That is because most forms of excellence, most of the time, tend to promote want-based interests. If there is an antecedent desire for excellence, as there often is, then the achievement of excellence is want-fulfilling, and even in the absence of such a

---

at all stands to 'gain' or 'lose' depending on whether it is satisfied. The pleasant state of mind we call satisfaction is itself a kind of reward or form of 'gain' (although it does not come automatically when we get what we desire) and intense disappointment is a kind of 'loss'. But one cannot do without the inverted commas. There is a distinction, crucial for our present purposes, between being disappointed *because* one has suffered a personal loss, and the 'loss' that consists entirely in disappointment, and between the 'gain' that consists entirely in satisfaction at some outcome, and the satisfaction that occurs *because* there has been some personal gain. The 'losses' and 'gains' in inverted commas have no direct connection with interests or with harms. We are commonly enough disappointed, dissatisfied, even frustrated without suffering harm.

[5] Barry, op. cit., pp. 38 ff.

desire, personal excellence is likely to contribute to the joint satisfaction of other wants. But contrary to Plato and many other ancient sages, there is no necessity that excellence and happiness always coincide, no impossiblity that morally inferior persons can be happy, and excellent persons miserable. There is still room for controversy then over what is truly good for persons in the latter two cases. In particular, philosophers have disagreed over whether it is *in the interest* of the contented moral defective to become a better person. This disagreement can persist even when it is agreed on all sides that it is desirable that the defective's character improve. Desirable, yes; a good thing, to be sure; but in *his* interest? That is another thing.

The source of the appeal of the ideal-regarding theory, I think, is evident: Few of us would wish to exchange places with people we regard as morally flawed, no matter how content they seem to be. It is easy to understand and sympathize with Epictetus' attitude toward the morally depraved criminal. We would not want to be *him* even if he escaped punishment, indeed even if he profited richly from his crime and suffered no remorse for it. Neither would we wish to be contented and vulgar, contented and dull, contented and stupid. We would in fact be prepared to sacrifice a good deal of our (other) want-fulfilments to avoid becoming flawed in these ways. But that is surely because we already have *desires* for excellences of character construed in accordance with our own standards. It is because we have such wants that we think it in our *interest* to be excellent, or at least not defective. Without those antecedent wants, it would not be in our interest to be excellent at all, except of course indirectly through the happy effects (not always to be relied upon) of excellent character on popularity and material success. By the same token, it is not in the interest of the contented moral defective to have *our* idea of virtue, which he doesn't share, imposed on him, unless, of course, we speak of thrift, prudence, diligence, etc., all of which could improve his chances of fulfilling his *other* ulterior wants. But if he is clever enough to make a 'good thing' in material terms out of dishonesty and unscrupulousness, even while he is cold-hearted, mean, vulgar, greedy, and vain, then it can hardly be in his interest to become warm, sensitive, cultivated, and generous; much less witty, perceptive, tactful, disinterested, and wise. We would not trade places with him to be sure, for it would not be in *our* interests to do so in so far as we have a stake, through the investment of our wants, in excellent character. We think, and rightly so in most cases, that we could only

lose by becoming worse persons, and that the change itself would constitute a loss, whatever further losses or gains it caused to our other interests.

Partisans of the ideal-regarding theory often rest their case on the example of child raising. Surely, it is said, we do not educate our children simply to become good want-fulfillers; rather we wish them to have the right wants in the first place, and to acquire the traits of character from which right wants emerge. Thus Stanley Benn claims that we are promoting the interests of the child when, at a time before he has achieved a good character, we commence with 'educating him to be a person of a certain sort'.

> His desires are beside the point [Benn writes], for it is often a question of whether he is to be encouraged to have desires of some approved sort instead of undesirable ones. It might be in the child's interests to deny him satisfaction of some of his desires to save him from becoming the sort of person who habitually desires the wrong thing.[6]

Benn's example supports an important point, but not the one he claims to be making. The point of moral education at the time it is undertaken is not simply to serve the child's interests either as they are or as they might one day become; not simply to promote his gain, profit, or advantage, his happiness or well-being. The aim is rather to lead the child, through creating new wants in him, to seek his happiness *by* pursuing personal excellence: to give him a *stake* in having a good character. The parent who values good character will want to give the child his own interest in it, so that the child's pursuit of his own interests will necessarily involve seeking and preserving virtues of character. The effect of making goodness one of a person's ulterior interests is to make the achievement of happiness impossible without attention to it. So, far from showing that a good character is in a person's interest even if it does not promote want-satisfaction, Benn's example shows instead that good character can be something that is directly in a person's interest only when the person has a want-based interest in *it*.

One of the advantages of the want-regarding theory is that it enables us all the more forcibly to praise personal excellence. Good character would be a good thing to have even if it didn't advance a person's self-interest. Self-interest, after all, isn't everything. It is no aid to clarity to insist that everything that is good *in* a person must be

[6] S. I. Benn, ' "Interests" in Politics', *Proceedings of the Aristotelian Society* 60 (1960), 130–1.

good *for* the person. Nor does it help to say that the evil in a person must be harmful to him. The contented moral defective is an ineligible model for emulation even though his faults cause no harm to himself. He is both evil and well off, and his evil character does not detract from his well-offness. Epictetus's 'pity' for him then is ill-placed. Vice is its own punishment, just as virtue is its own reward, only to the person who has a stake in being good.

It is not merely useful but morally important to preserve in this way the distinction between being good and being well off, for it saves us from speaking as if, and perhaps really believing, that well-offness is the sole good. It is important to be a good person and not merely a happy or fulfilled one. That is why we train children to seek their happiness in part through seeking their goodness. In that way we ensure that they will not be completely happy unless they are good.

Morally corrupting a person, that is, causing him to be a worse person than he would otherwise be, can *harm* him, therefore, only if he has an antecedent interest in being good. (It may in fact harm no one to corrupt him if he is corrupted in a way that does not make him dangerous to others.) The moral corruption or neglect of an unformed child, then, is no direct harm to him, provided that he has the resources to pursue his own interests effectively anyway, but it can be a very real harm to his parents if *they* have a powerful stake in the child's moral development.

2. *Other-regarding wants and vicarious harms.* There are two ways in which one person can have an interest in the well-being of another. In the one case, *A* may be dependent upon the help of *B* for the advancement of his own (*A*'s) interests, so that if *B*'s fortunes should decline, *B* would be less likely to help *A*. What promotes *B*'s interest, in this case, indirectly promotes that of his dependent *A* as well. It is therefore in *A*'s interest that *B*'s interest be advanced. In the extreme version of this case, where *A* is *wholly* dependent on *B*'s help, and so long as *B*'s personal interest flourishes the help is sure to continue, *B*'s good is, in effect, one of *A*'s welfare interests, the advancement of which (like his own health) promotes the whole economy of his ulterior interests and is absolutely essential to his well-being, whatever his ulterior interests happen to be.

In the second kind of case, *C* has 'invested' a desire so strong, durable, and stable in *D*'s well-being, that he comes to have a personal stake in it himself. It becomes, therefore, one of his ulterior

interests or 'focal aims'.[7] This should be contrasted with the more common phenomenon of spontaneous sympathy, pity, or compassion which can be directed at total strangers. It may make *A* very unhappy to see *B* (a stranger) suffer, and *A* may do what he can to help *B*, from genuinely disinterested, compassionate motives. But the harm that has been done *B*, say, by a hit-and-run motorist who knocked him down, is not *also* harm done *A*. The interests of *A* have not been invaded by the harm done *B*; he has only suffered some vicarious unhappiness on *B*'s behalf which will leave his own personal interests largely unaffected. In the case of genuinely other-regarding interest that I have in mind, *C* has an abiding interest of his own in *D*'s well-being which is not merely an episodic 'passing desire'. Further, he desires *D*'s good not simply as a means to the promotion of the other ulterior aims that are components of his own good, but quite sincerely as an end in itself. Such cases are, of course, rare, but no rarer than disinterested love. Indeed, there is one sense of 'love' (that which the New Testament writers called *agape*) which is well defined by the presence of purely other-regarding interest. Ralph Barton Perry once defined 'love' in this sense as an interest in the advancement of someone else's interests.[8] When *C* has a loving interest in *D*'s personal interest, then anything that harms *D* directly *ipso facto* harms *C* indirectly. Can anyone doubt that one harms a loving parent by maiming his child (or as in the previous example, by corrupting his child) or that one harms a loving husband or wife by causing a disappointment that plunges his or her spouse into despair?

The separation of the two kinds of cases distinguished in the preceding paragraphs is somewhat artificial. The distinction is clearly enough conceived, but in real life psychological elements rarely separate so neatly. Most of the things we desire for their own sakes we *also* desire as means to other things. Harm to a child may itself be harm to its loving parent in that it directly violates the parent's 'purely' other-regarding interest, but it may also be instrumentally damaging to various self-regarding interests of the parent, in that it creates a drain on his funds, a burden on his time and energy, and a strain on his emotional stability. Similarly, when one spouse sinks into despair, this not only harms the other person's wholly other-regarding interest in the ailing mate's well-being; it also

[7] This phrase is C. L. Stevenson's. See his account in *Ethics and Language* (Yale U.P., New Haven, 1944), p. 203.

[8] Ralph Barton Perry, *General Theory of Value* (Longmans, Green, & Co., New York, 1926), p. 672. His exact words: '. . . a favourable interest in the satisfaction of the interest of a second person'.

deprives him or her of the myriad services and pleasures that a cheerful partner would contribute.

Loving interests are so commonly intertwined with, and reinforced by, instrumental, essentially self-regarding interests, that many observers are led to discount the former, or even deny altogether their existence in given cases. Others have embraced the apparently cynical view that there are no purely other-regarding interests at all, that human nature being what it is, no one 'really cares' about the well-being of other persons, except in so far as it affects his own self-regarding interests. All interests in the well-being of others, on this view, are of the first type distinguished above. This extreme form of psychological egoism rules out not only disinterested love, but episodic sympathy and compassion as well. Egoism of this sort can never be persuasive to those who are deeply impressed by the genuine purity of their own love for others, so its advocates must posit a good deal of self-deception in their opponents. Since the purity of people's motives is not readily subject to careful scrutiny, the egoistic theory, as a matter of empirical psychology, is not easily refuted, though the stronger philosophical arguments *for* the view are invariably muddled.

Some types of apparently other-regarding interests are so familiar, however, that the burden of explaining them away should be placed on the egoist. One common example is the case of pooled interest, where, either through design or accident, separate persons are so related that they share a common lot. Such common interests, 'all for one and one for all', are found wherever parties are led (or forced) by circumstances to act in concert and share the risk of common failure or the fruits of an indivisible success.[9] Whatever the ultimate truth of the matter, common sense reports that persons with pooled or interdependent interests are sometimes drawn even closer by bonds of sentiment directed toward common objects or reciprocal affection (of an *apparently* disinterested kind) between the parties. And when this happens, as it sometimes seems to in marriages and family groups, each has a genuine stake of a not merely instrumental kind in the well-being of the others, a stable ulterior goal, or focal aim, that the others flourish, partly as an end in itself, partly as a means to a great diversity of other ends.

Despite the familiarity of these observations, some very able philosophers have chosen to exclude purely other-regarding wants

[9] See my 'Collective Responsibility' in *Doing and Deserving* (Princeton U.P., Princeton, N.J., 1970), pp. 233–41.

altogether from their otherwise want-regarding analyses of interest. The writers in question do not necessarily deny that there are purely other-regarding wants. Professor Barry, for example, admits that some of us, some of the time, genuinely want other persons as well as ourselves to enjoy increased opportunities to satisfy ulterior wants. Indeed, he concedes that some persons, some of the time, even voluntarily suffer a diminution of their opportunities for want-satisfaction in order to increase the opportunities of other persons to satisfy *their* wants. But the latter cases, Barry insists, are best described as cases where our *principles* are allowed to override out interests.[10] Barry is right about the cases he seems to be considering, where persons voluntarily sacrifice their own interests for others out of a sense of justice, or for ideal-regarding reasons, or for charity. But he doesn't even consider cases of the kind discussed above where help to others is not thought to be a *sacrifice* at all, but a direct promotion of one's own other-regarding interest in the advancement of the interests of another party.

I think the theoretical motives of writers who exclude other-regarding wants from their analyses of self-interest are clear enough, and worthy of respect. They are simply taking the easiest way out of a kind of linguistic muddle. They are afraid that inclusion of purely other-regarding aims as eligible constituents of a person's own self-interest would commit them to saying various odd-sounding things. They fear that we would have to say when Jones gives his last cent to promote the cause of his favourite political party, or to finance his child's education, or to secure the very best doctor for his sick wife, that he is advancing his own interest *merely* (treacherous word, 'merely'). Hence, we must think of his act as 'selfish', since it was done in his own self-interest, after all. The less paradoxical alternative, they think, is to deny that the act is in the actor's own interest at all, and to say instead that Jones was acting from conscience, or out of principle, or for charity, and against his own interest. After all, how could his act be at once disinterested and self-interested, unselfish yet self-advancing?

There is, however, a more satisfactory, if less direct, way out of the muddle. That is simply to consider very carefully what it means to call an act 'unselfish' and 'disinterested', and to come by this route to appreciate how unselfish and disinterested conduct, without affecting any of the actor's interests other than those he has in the well-being

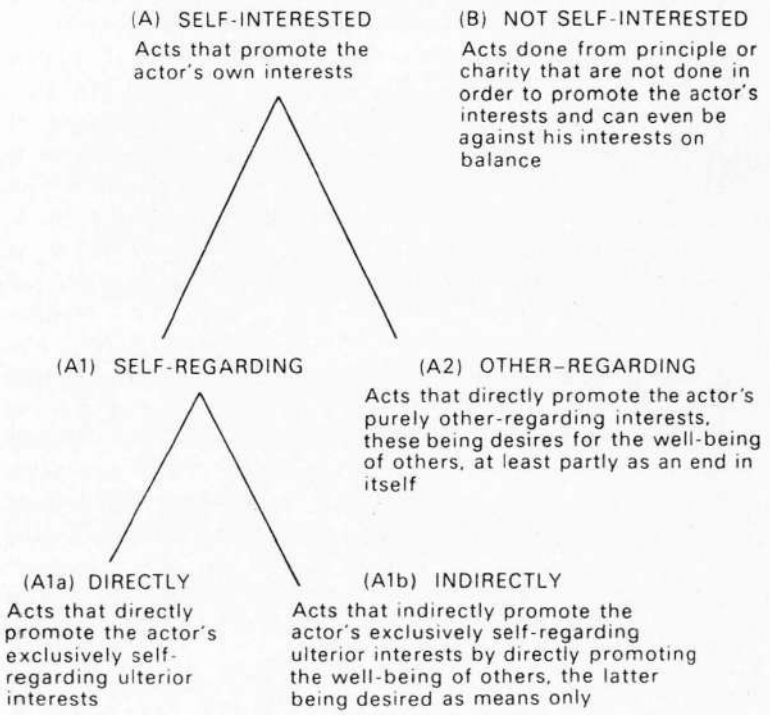
[10] Barry, op. cit., p. 77

of others, can nevertheless be in his own personal interest. A person who has such a stake in the happiness of other people that his own well-being depends on the advancement of their interests is not the proper model of a selfish person. A selfish person is one who pays insufficient attention to the interests of other people, and thus comes to pursue his own self-regarding interests at the expense of, or in disregard of, the interests of others. That is quite another thing than pursuing one's own interest in *promoting* the interests of others. The loving parent or spouse and the public-spirited zealot can make no distinction between their own interests and that of their children, or spouse, or party. Far from indicating their selfishness, that identity of interests shows how unselfish they probably are. They might yet be blamably selfish, however, if they pursue those of their own interests which include the interests of *some* other people (for example, a daughter and a son) at the expense of the interests of still *other* people (for example, their neighbours' children). It is in fact an advantage of our analysis (as opposed to Barry's) that it enables us to explain why conduct of the latter kind is selfish. On Barry's analysis, neither want—that for the well-being of my children nor that for the well-being of my neighbours' children—is one of my own interests. Hence, when I promote the interests of some of these parties at the expense of those of the others, I am acting neither for nor against my own interests. I can be acting oddly or wrongly in that case, but not specifically selfishly. That judgment, however, seems plainly false. It surely *is* selfish wrongly to benefit one's own loved ones at the expense of others.

The best way, it seems to me, to conceive of the relation between self-interested, selfish, unselfish, and disinterested acts is that indicated in a chart with two genera, one of which is further sub-divided (see diagram 1 on p.295).

The generic distinction in the chart is that on the top line between self-interested acts and acts that are not self-interested, particularly those that concerned Barry, namely, conscientious or charitable acts that are not predominantly in the actor's interest. Self-interested acts are then divided into self-regarding and other-regarding species. Depending on our purposes, of course, we would classify the acts in this motley category in various alternative ways, but it is especially useful for our present purposes to divide them into these mutually exclusive and jointly exhaustive categories. The self-regarding class is then further divided into directly and indirectly self-regarding

subclasses. For an example of directly self-regarding activity (A1a) consider an unmarried home-owner's labour at improving his property so that he can take more enjoyment and pride in it, impress those in a



position to help, and disproportionately increase its resale value in a rising market. Such a person is promoting his own purely self-regarding ulterior interests in material possession, career advancement, and capital accumulation. An example of indirectly self-regarding activity (A1b) is found in the story of the gambler, *A*, who bets *B* $50,000 that *C* will recover from a serious illness. Thus *C*'s health is in *A*'s interest, and *A* has a stake (in a literal sense) in *C*'s recovery. To protect that stake he works hard to promote *C*'s recovery, providing at his own expense, the best medical and nursing care that he can find. He thus promotes the well-being of another as ardently as a lover or a saint would, though the other's well-being, his immediate goal, is desired only as a means to the advancement of his own self-regarding interest.

In contrast, acts in the genuinely other-regarding species of the self-interested genus (A2) aim at the promotion of another's good at least partly as an end in itself. An example would be that of a parent whose stake in the well-being of his child is derived from his love for the child simply, and not from any incidental service to his other (self-regarding) interests that the child might contribute. If such a parent depletes his own life savings to advance or protect his child, his act would fall in the other-regarding species of the self-interested genus. This is the species which is thought to be empty, for quite different reasons, by psychological egoists and Brian Barry. The egoists deny that any acts are genuinely other-regarding (that is, motivated by a desire to promote or retard the good of another as an end in itself), while Barry denies that any other-regarding acts are self-interested. But if any person ever does 'really care' whether another person is harmed or benefited, and not simply as a means to his own gain but at least in part for the other's own sake, then the egoists are wrong. And if any person ever does have a genuine stake in the happiness of another person—an independent ulterior interest not wholly derived from its service to other ulterior interests—such that he himself gains or loses directly depending on the condition of the other person, then the view suggested by Barry is wrong.

The chart enables us to distinguish several senses of 'disinterested action' and also two kinds of selfish action. A disinterested act can be defined in a first approximation, as one not done simply to advance the actor's interests.[11] One class of disinterested actions, then, consists of those in the chart's second genus: those not done to advance *any* of the actor's interests, self-regarding or other-regarding. These are actions done from conscience, or out of a sense of justice, or from charity, or from a spontaneous benevolent impulse, often with the conscious expectation that they will be against the actor's own interest. A second kind of disinterested action is one which meets a stricter test; it is neither done to promote the actor's own interest or to favour the interests of any second parties unfairly at the expense of third parties when the actor's own interests simply aren't involved one way or the other. Thus, a judge's decision is disinterested when it is unbiased and impartial. These related senses of the word 'disinterested' are well established in usage. A third sense

[11] Cf. *Webster's New International Dictionary*, 2nd edn. (1954): 'not influenced by regard to personal advantage . . .' and *The Oxford English Dictionary*: 'not influenced by self-interest . . .'

(one which is suggested by our chart) is not so clearly established and may in fact be somewhat extended beyond what is recognized in ordinary usage. I am not sure. Still, it stands for an important category that deserves to be distinguished from the others, whatever name it bears. I refer to actions in the other-regarding species of the self-interested genus (A2), acts done out of the perfectly genuine desire to help another whose well-being is actually a constituent of the actor's own good. When a person promotes the well-being of a loved one in a self-sacrificing or otherwise 'selfless' way, it may be misleading to call his act disinterested since he does have a personal stake, even a predominant ulterior interest, in the outcome. But it can be equally misleading to deny that his act is disinterested since apart from the well-being of the loved one that is his goal, there may be no 'personal advantage' in his action, and no trace of self-interest in his motivation. In an extreme case, he might even sacrifice all his other interests for the good of another person or cause in which he has 'invested' everything. The least misleading thing to say about such conduct is that it is not disinterested in one very familiar sense of the term, but that it is disinterested in another, less familiar, sense. In any event, extreme psychological egoists are likely to deny that there are disinterested acts of either kind, and sometimes put that view by saying that all voluntary actions are 'selfish'.

Now a selfish act, whatever else it may be, is one that is morally defective. A person acts selfishly when he pursues his own interests (or the satisfaction of transitory desires and appetites) *wrongly* at the expense of others. Sometimes, of course, there is nothing blamable in the pursuit of self-interest at the expense of others, as for example, in legitimately or unavoidably competitive contexts. An act is selfish only when its pursuit of self-interest is somehow in excess of what is right or reasonable in the situation.

The more familiar kind of selfish act is a defective specimen of those in the self-regarding species of self-interested actions (A1a and b). The father who refuses to spend money on his children for anything beyond their minimal needs, and uses his surplus instead to buy fine clothes and wines for himself is selfish in this way. But as we have seen, defective specimens of acts in the other-regarding species of the self-interested genus (A2) can also be selfish, as when a parent with a genuinely independent stake in his own children's advancement (an 'other-regarding interest') pursues that interest wrongly at the expense of his neighbour's children. We would be reluctant, I

think, to call the latter actions 'disinterested' in *any* sense, since it would be intolerably odd to think of an act as both disinterested and selfish. Hence I am forced to qualify the account given above of the self-interested acts that can also be, in an 'unfamiliar sense', disinterested, as follows: an act is disinterested in that third sense provided that (i) it is done in order to advance the good of another party, but (ii) not merely as a means to the advancement of the actor's own self-regarding interests, *and* (iii) it is not done to promote the actor's other-regarding interest in the well-being of one party wrongly at the expense of still another party. (This third condition amends the definition, in effect, by requiring that a disinterested act not be a selfish act of the second kind.)

Selfish actions, then, can be defined as those which pursue the actor's self-interest *wrongly* at the expense of, or in disregard of, other people, and the two main types of selfish actions are those which are appropriately defective instances of category A1 on the chart, and those which are appropriately defective instances of A2. (Morally defective instances of B, as we shall see, are not called 'selfish'.) It is best, I think, to define 'selfish' and 'unselfish' as logical contraries rather than contradictories, in recognition of a large and motley class of actions that are neither selfish nor unselfish. An unselfish act then can be defined as one which pursues the interests of others (or the fulfilment of their transitory wants or appetites) *rightly* at the expense of, or in praiseworthy disregard of, the actor's own interests (or wants and appetites).[12] Voluntary actions in the middle group that qualify neither as selfish nor as unselfish include those which pursue the actor's own self-regarding wants or interests (A1) in a non-defective way (not wrong or blamable, not deficient in concern for others) as well as those whose motivation does not include concern for self-interest one way or the other, as in the case of the judge in a controversy between two persons who are strangers to him.

There are blamably defective specimens of acts even in the non-self-interested genus (B on the chart), but these characteristically bear names other than 'selfish'. Acting entirely out of principle, for example, a person might be rigid, cruel, or intolerant. A person might, in another case, act honestly in accord with a dictate of his

[12] Two kinds of unselfish actions then can be distinguished in terms of the categories in the chart: those in category B and those actions in category A2 that are not done wrongly at the expense of, or in blamable disregard of, the interests (or passing wants) of third parties.

own mistaken or confused conscience. Another person might act unjustly or imprudently out of spontaneous compassion. All of these morally defective acts can be against the actor's interest and known to be such, yet deliberately chosen anyway. They may be blameable, but they are not selfish.

According to our provisional definition of 'harm', a violation of an interest in any of the categories in the chart is a harm to its possessor. Any action, omission, or rule that interferes with a person's self-interested action, thus thwarting his interest, causes him harm. But does it follow from the definition that interferences with voluntary acts in the non-self-interested genus (B) are *not* harms? That would seem at first sight to be the case. Since acting out of conscience or benevolence is not acting to advance one's own interest, interference with such action does not violate one's interest, and therefore is not, by definition, a harm to one. Moreover, such interference, when it prevents a person from acting contrary to his own interest, actually *serves* his interest, and would seem therefore to be beneficial to him. Any interference, however, with a voluntary action, even with a non-self-interested one, is an invasion of a person's interest in *liberty*, and is thus harmful to him to that extent. If that seems too trivial a harm in the case at hand to be the basis of a powerful claim to non-interference, the liberal will have to retreat from the harm principle and seek a stronger defensive position, perhaps in the principle that infringements of an actor's *autonomy* are seriously wrongful even when they do him, at most, only trivial harm.

3. *Death and posthumous harms.* If a murderer is asked whether he has harmed his victim, he might well reply: 'Harmed him? Hell no; I killed him outright!' The victim's mourners too might feel that it is something of an understatement to describe the death of their loved one as a harm (to him). The death of the victim, it would seem, is not merely a 'harmed condition' he is put in; it is no 'condition' of him at all, but rather his total extinction. Consider the purest possible hypothetical case of the infliction of death, where all extraneous and distracting harms have been excluded from the example. A man in the prime of his life, with many on-going projects and enterprises, but with no dependents or friends close enough to mourn him, is shot by an unseen assailant in the back of the head. Without ever being aware even that he was in danger, much less that he has been fatally wounded, he dies instantly. Right up to the very instant he was shot, he was unharmed; then at that very moment, perhaps one second

after the killer squeezed the trigger, he was dead. At the very most, he was in a 'harmed condition' for the one half-second, or so, before he died. As for death itself, one might agree with the ancient Epicureans: 'Where he was, death was not, and where death was, he was not'.

Yet for all of that, it seems clear that the murderer did violate his victim's interest in remaining alive. One second before the trigger was pulled, it was true of the victim (as it is now true of both the author and reader of these words) that continued life was something *in his interest*. Indeed, there is nothing a normal person (in reasonable health and tolerable circumstances) dreads more than his own death, and that dread in the vast majority of cases, is as rational as it is unavoidable, for unless we continue alive, we have no chance whatever of achieving the goals that are the ground of our ultimate interests. Some of these goals perhaps might be achieved for us by others after our deaths, publicly oriented and other-regarding goals in particular. But most of our interests require not simply that some result be brought about, but rather that it be brought about *by* us, or if not by us, then *for* us. My interest in producing an excellent book, or a beautiful art object, is not fully satisfied by another person's creation of such objects. My interest was not simply that such objects exist, but that *I* bring them into existence. Similarly my aim to build a dream house, or to achieve leisure in security, is not satisfied when such a house or such leisure comes into existence, but only when I am present to enjoy and use it. Our interest in avoiding death is a supreme welfare interest, an indispensable condition for the advancement of most, if not all, of the ulterior interests that constitute our good. There is something bare minimal about it on the one hand, yet something supremely important on the other. Apart from the interests it serves, it has no value in itself; yet unless it is protected, hardly any of a person's ulterior interests will be advanced. To extinguish a person's life is, at one stroke, to defeat almost all of his self-regarding interests: to ensure that his on-going projects and enterprises, his long-range goals, and his most earnest hopes for his own achievement and personal enjoyment, must all be dashed.

There is a case then both for saying that death is not a harm and that it *is* a violation of an antecedent interest in staying alive. That makes death a very hard case indeed for the analysis of harm as invaded interest. There may be no way out of this for the writer who has strong theoretical incentives for saving the invaded interest

theory other than to stipulate an admittedly extended sense of 'harm' broad enough to include death as a harm.[13] This would be a minor and quite excusable departure from the conventions of ordinary language for the sake of theoretical economy; still, it would make things tidier all around if we could *show* that, ordinary language to the contrary—and not as a matter of mere arbitrary stipulation—death *is* a harm.[14] It would be unreasonable to expect that this conclusion could be demonstrated, and indeed, there are various common sense considerations other than the oddness of its sound to the ear that militate against it. But there is also a way of conceiving death (even without the assumption of survival or immortality) that mitigates its paradox and lends it some plausibility. That is all that can be claimed, at best, for the view that death can be a harm to the one who dies.

To be sure, death is not always and necessarily a harm to the one who dies. To the person in hopeless, painful illness, who has already 'withdrawn his investments' in all ulterior interests, there may be nothing to lose, and cessation of agony or boredom to be 'gained', in which case death is a blessing. For the retired nonogenarian, death may not exactly be ardently desired, but still it will be a non-tragedy. Those who mourn his death will not think of themselves as mourning *for him*, but rather for his dependants and loved ones, if any, or simply in virtue of the capacity of any *memento mori* to evoke sadness. In contrast, when a young vigorous person dies, we think of *him* as chief among those who suffered loss.

One way of saving the 'invaded interest' theory of harm, at minimal cost to common sense, is to think of all harm as done to interests themselves, and interpret talk of harm done to men and women as convenient elliptical references to, and identification of, the interest that was thwarted or set back. Thus, when Cain harms Abel by punching him in the nose, it is Abel's interest in the physical integrity of his nose that is the immediate object of the harm, and

[13] For the writer who is interested in formulating a more precise and defensible version of Mill's 'harm principle' there is another alternative. He can simply amend his statement of that principle so that it restricts interferences with liberty to those necessary to prevent *harm or death* (implying that they are not the same thing). The cost of this amendment, however, would be the abandonment of the analysis of harm as 'any invasion of interest', for there *is* an interest in avoiding death, yet the amendment implies that death is not a harm.

[14] Thomas Nagel has argued ingeniously but inconclusively that death is an 'evil' or a 'misfortune' to the one who dies. This is not quite the same perhaps as saying that a person is *harmed* when he is killed, but it is close. See his article 'Death' in its expanded form in *Moral Problems*, edited by James Rachels (Harper & Row, New York, 1971), pp. 361–70.

Abel himself is harmed in the derivative sense of being the owner of a harmed interest. This is perhaps a step beyond (but only a small step beyond) saying what is obviously true: that it is only in virtue of having interests that people can be harmed, and that the only way to harm any person is to invade his interests. If Abel had no interest of the usual welfare kind in the integrity and normal functioning of his body, then Cain could not have harmed him by punching him in the nose, but at most only hurt, annoyed, or disappointed him. The next step is to point out that most of a person's self-regarding interests, at least, are thwarted permanently, and thus harmed, by his death. Although *he* no longer exists, we can refer to his earlier goals (as a matter of identification) as *his* interests, and *they* were the interests directly harmed by his death.

What then does it mean to say that an *interest* has been harmed? Our answer to this question will depend on which of two conceptions of interest enhancement and impairment we adopt. As we have seen, interests are 'stakes' that are derived from and linked to wants, in the case of ulterior interests to more ulterior goals or focal aims. Now we can apply to these wants W. D. Ross's distinction between want-fulfilment and want-satisfaction.[15] The *fulfilment* of a want is simply the coming into existence of that which is desired. The *satisfaction* of a want is the pleasant experience of contentment or gratification that normally occurs in the mind of the desirer when he believes that his desire has been fulfilled. When the object of a want does not come into existence, we can say that the want has been *unfulfilled* or *thwarted*; the experience in the mind of the desirer when he believes that his desire has been thwarted is called *frustration* or *disappointment*. Notoriously, fulfilment of desire can fail to give satisfaction. There is no more melancholy state than the disillusionment that comes from getting what we wanted and finding it disappointing. Such disillusionment can usually be explained as the consequence of a rash or ill-considered desire and unrealistic expectations. On other occasions, the original desire will bear up under retrospective scrutiny, and yet its fulfilment gives no pleasure. Indeed, the occurrences of subjective satisfaction is a highly contingent and unreliable phenomenon. Sometimes when our goals are achieved, we don't experience much joy, but only fatigue and sadness, or an affective blankness. Some persons, perhaps, are disposed by temperament normally to receive their achievements in this unthrilled fashion. Still, even in these cases,

[15] W. D. Ross, *Foundations of Ethics* (Clarendon Press, Oxford, 1939), p. 300.

re-examination of the goal whose fulfilment failed to satisfy may disclose no hidden defects, no reasons for regret, in a word, no disillusionment. Not only can one have fulfilment without satisfaction; one can also have satisfaction of a want in the absence of its actual fulfilment, provided only that one is led to believe, falsely, that one's want has been fulfilled. Similarly, pleasant states of mind resembling 'satisfaction' can be induced by drugs, hypnosis, and other forms of manipulation that have no relation whatever to prior wants.

Similarly, one's wants can be thwarted without causing frustration, or disappointment, and one can be quite discontented even when one's wants have in fact been fulfilled. These negative cases are perfectly parallel with the positive ones. Non-fulfilment of a want yields no disappointment when the want was ill advised in the first place. In such a case, the want can happily be renounced after rational reassessment. Disillusionment, however, is often not involved. A perfectly genuine and well-considered goal may be thwarted without causing mental pain when the desirer has a placid temperament or a stoic philosophy. And discontent does not presuppose thwarting of desire any more than satisfaction presupposes fulfilment. One can have feelings of frustration and disappointment caused by false beliefs that one's wants have been thwarted, or by drugs and other manipulative techniques.

For these reasons, harm to an interest is better defined in terms of the objective blocking of goals and thwarting of desires than in subjective terms; and the enhancement or benefiting of an interest is likewise best defined in terms of the objective fulfilment of well-considered wants than in terms of subjective states of pleasure. Most persons will agree, I think, that the important thing is to get what they want, even if that causes no joy.[16] The pleasure that normally attends want-fulfilment is a welcome dividend, but the object of our efforts is to fulfil our wants in the external world, not to bring about states

[16] This judgment is probably too confident if understood to extend to cases where what is wanted is expected to cause actual *disappointment*. Derek Parfit has reminded me of the distinction between cases where fulfilment *can't possibly* produce satisfaction because the person will never be in a position to know that his want has been fulfilled, and cases where fulfilment can produce satisfaction but in fact won't. In the former case, all would agree that the important thing is that what we want to happen will happen (our desire will be fulfilled). But in the latter case, if people know or confidently expect that fulfilment will not only 'not cause joy' but will actually produce disappointment, it is not so clear, as Parfit points out, that the important thing is 'to get what one wants'. There is some question, however, whether the existence of the want could even survive such conditions.

of our own minds. Indeed, if this were not the case, there would be no way to account for the pleasure of satisfaction when it does come; we are satisfied only because we think that our desires are fulfilled. If the object of our desires were valuable to us only as a means to our pleasant inner states, those inner glows could never come.

The object of a focal aim that is the basis of an interest, then, like the object of any want, is not simply satisfaction or contentment, and the defeat of an interest is not to be identified with disappointment or frustration. Hence, death can be a thwarting of the interests of the person who dies, and must be the total defeat of most of his self-regarding interests, even though, as a dead man, he can feel no pain.

This account helps explain, I think, why we grieve for a young, vigorous 'victim of death' himself, and not *only* for those who loved him and depended on him. We grieve for him in virtue of his unfulfilled interests. We think of him as one who has invested all his energies and hopes in the world, and then has lost everything. We think of his life as a whole as not as good a thing as it might have been had he lived on. In some special circumstances, death not only does its harm in this wholly 'negative' way, preventing the flowering of the interests in which a person's lifetime good consists, it also does direct and 'positive' harm to a person by undoing or setting back important interests that were already prospering. Death, in these cases, leads to the harming of surviving interests that might otherwise have been prevented.[17]

Because the objects of a person's interests are usually wanted or aimed-at events that occur outside of his immediate experience and at some future time, the area of a person's good or harm is necessarily wider than his subjective experience and longer than his biological life. The moment of death is the terminating boundary of one's biological life, but it is itself an important event within the life of one's future-oriented interests. When death thwarts an interest, the interest is harmed, and the harm can be ascribed to the man who is no more, just as his debts can be charged to his estate.

The interests that die with a person are those that can no longer be helped or harmed by posthumous events. These include most of his self-regarding interests, those based, for example, on desires for personal achievement and personal enjoyment, and those based on 'self-confined' wants that a person could have 'if he were the only

[17] The most vivid example I know in literature of a 'positively harmful' death is that foreseen by Pip at the hands of the villainous Orlick in Dickens's *Great Expectations*.

person that had ever existed',[18] for example, the desire to be a self of a certain kind, or the desire for self-respect. Other self-regarding wants, in contrast, seem more like other-regarding and publicly oriented wants, in that they can be fulfilled or thwarted after the death of the person whose wants they are. I refer to some of a person's desires to stand in certain relations to other people where 'the concern is primarily with the self . . . and with others only as objects or as other terms in a relation to me'.[19] These desires can be called 'self-centred', and include as a class such wants as the desire to assert or display oneself before others, to be the object of the affection or esteem of others, and so on. In particular, the desire to maintain a good reputation, like the desire that some social or political cause triumph, or the desire that one's loved ones flourish, can be the basis of interests that survive their owner's death, in a manner of speaking, and can be promoted or harmed by events subsequent to that death. Fulfilment and thwarting of interest, after all, may still be possible, even when it is too late for satisfaction or disappointment.

The above account might still contain elements of paradox, but it can be defended against one objection that is sure to be made. How can a man be harmed, it might be asked, by what he can't know? Dead men are permanently unconscious; hence they cannot be aware of events as they occur; hence (it will be said) they can have no stake one way or the other, in such events. That this argument employs a false premiss can be shown by a consideration of various interests of *living* persons that can be violated without them ever becoming aware of it. Most of these are 'possessary interests' whose rationality can be doubted, for example, a landowner's interest in the *exclusive* possession and enjoyment of his land—an interest that can be invaded by an otherwise harmless trespasser who takes one unobserved step inside the entrance gates; or the legally recognized 'interest in domestic relations' which is invaded when one's spouse engages in secret adulterous activity with a lover. The latter is an interest in being the exclusive object of one's spouse's love, and has been criticized by some as implying property in another's affections. But there is no criticizing on such grounds the interest every person has in his own reputation, which is perhaps the best example for our present purposes. If someone spreads a libellous description of me,

[18] C. D. Broad, 'Egoism as a Theory of Human Motives' in *Ethics and the History of Philosophy* (Routledge and Kegan Paul, London, 1952), p. 220.
[19] Op. cit., p. 221.

without my knowledge, among hundreds of persons in a remote part of the country, so that I am, still without my knowledge, an object of general scorn and mockery in that group, I have been injured in virtue of the harm done my interest in a good reputation, even though I *never* learn what has happened. That is because I have an interest, so I believe, in having a good reputation *as such*, in addition to my interest in avoiding hurt feelings, embarrassment, and economic injury. And *that* interest can be seriously harmed without my ever learning of it.

How is the situation changed in any relevant way by the death of the person defamed? If knowledge is not a necessary condition of harm before one's death why should it be necessary afterward? Suppose that after my death, an enemy cleverly forges documents to 'prove' very convincingly that I was a philanderer, an adulterer, and a plagiarist, and communicates this 'information' to the general public that includes my widow, children, and former colleagues and friends. Can there be any doubt that I have been harmed by such libels? The 'self-centred' interest I had at my death in the continued high regard of my fellows, in this example, was not thwarted by my death itself, but by events that occurred afterward. Similarly, my other-regarding interest in the well-being of my children could be defeated or harmed after my death by other parties overturning my will, or by thieves and swindlers who cheat my heirs of their inheritance. None of these events will embarrass or distress me, since dead men can have no feelings; but all of them can harm my interests by forcing non-fulfilment of goals in which I had placed a great stake.

This liability, to which we are all subject, to drastic changes in our fortune both before and after death was well understood by the Greeks. Aristotle devotes a chapter of his *Nicomachean Ethics* to a saying already ancient in his time, and attributed by some to Solon, that we can 'call no man fortunate before his death'.[20] On one interpretation, this dark saying means that 'only when he is dead is it safe to call a man . . . beyond the arrows of outrageous fortune'. On the day before he dies, his interests can be totally smashed and his life thus ruined. But as Aristotle shrewdly observes (attributing the point to the general popular wisdom), some of a person's interests are not made safe even by his death, and we cannot call him fortunate with perfect confidence until several more decades have passed; 'For a dead man is popularly believed to be capable of experiencing both

[20] Aristotle, *Nicomachean Ethics*, I. 10.

good and ill fortune—honour and dishonour, and prosperity and the loss of it among his children and descendants generally—in exactly the same way as if he were alive but unaware or unobservant of what was happening'.[21]

Three hypothetical cases can illustrate the 'popular belief' mentioned by Aristotle, and the case for posthumous harm must rest with them.

Case A. A man devotes thirty years of his life to the furtherance of certain ideals and ambitions in the form of one vast undertaking. He founds an institution dedicated to these ends and works single-mindedly for its advancement, both for the sake of the social good he believes it to promote, and for the sake of his own glory. One month before he dies, the 'empire of his hopes' collapses utterly as the establishment into which he has poured his life's energies crumbles into ruin, and he is personally disgraced. He never learns the unhappy truth, however, as his friends, eager to save him from disappointment, conceal or misrepresent the facts. He dies contented.

Case B. The facts are the same as in Case A, except that the institution in which the man had so great an interest remains healthy, growing and flourishing, until the man's death. But it begins to founder a month later, and within a year, it collapses utterly, while at the same time, the man and his life's work are totally discredited.

Case C. The facts are the same as in Case B, except for an additional surmise about the cause of the decline and collapse of the man's fortune after his death. In the present case, a group of malevolent conspirators, having made solemn promises to the man before his death, deliberately violate them after he has died. From motives of vengeance, malice, and envy, they spread damaging lies about the man and his institution, reveal secret plans, and otherwise betray his trust in order to bring about the ruin of his interests.

It would not be very controversial to say that the man in Case A had suffered grievous harm to his interests although he never learned the bad news. Those very same interests are harmed in Case B to exactly the same extent, and again the man does not learn the bad news, in this case because he is dead, and dead men hear no news at all. There seems no relevant difference between Case A and Case B

[21] Ibid., first paragraph. Aristotle's primary concern in this chapter, however, was not to show that a person's interests can be affected after his death, but rather that well-being, whether before or after death, cannot be destroyed by the caprice of events, but at worst, only somewhat tarnished. The point about interests surviving death he simply assumed as beyond need of argument.

except that in Case B there might seem to be no subject of the harm, the man being dead. But if we consider that the true subjects of harms are interests, and that interests are harmed by thwarting or non-fulfilment rather than by subjective disappointment, we can think of posthumous harms as having subjects after all. But if that point is not convincing, the argument must depend on its reinforcement by Case C. In that example, the man is not *merely* harmed (if he is harmed at all); rather he is exploited, betrayed, and wronged. When a promise is broken, someone is wronged, and who if not the promisee? When a confidence is revealed, someone is betrayed, and who, if not the person whose confidence it was? When a reputation is falsely blackened, someone is defamed, and who, if not the person lied about? If there is no 'problem of the subject' when we speak of wronging the dead, why should there be, when we speak of harming them, especially when the harm is an essential ingredient of the wrong?

To summarize then: Death can thwart a person's ulterior, self-regarding interests in personal achievement and enjoyment, by totally defeating the welfare interest that is necessary for fulfilment of the goals and focal aims that are their bases. It is for this reason alone that death is a harm to the one who dies suddenly, in the prime of life, never knowing what hit him, and unmourned by loved ones or dependents. We grieve for such a person (as opposed to grieving for our own loss) because of his unfulfilled interests. Events after death can thwart or promote those interests of a person which may have 'survived' his death. These include his publicly oriented and other-regarding interests, and also his 'self-centred' interests in being thought of in certain ways by others. Posthumous harm occurs when the deceased's interest is thwarted at a time subsequent to his death. The awareness of the subject is no more necessary than it is for harm to occur to certain of his interests at or before death.