OXFORD

ON HUMAN RIGHTS

James Griffin

individual has.[12] He goes on to distinguish *innate* from *acquired* rights. An 'innate' right belongs 'to everyone by nature independently of any act that would establish a right', while 'acquired' rights are those that require such an act.[13] There is only one innate right: namely, '*freedom* (independence of constraint by another's choice) insofar as it can coexist with the freedom of every other in accordance with a universal law'; it is 'the only original right belonging to every man by virtue of his humanity'.[14] Kant has made here a fateful move: the content of the one overarching right is the same as what Kant calls 'The Universal Principle of Right',[15] which I stated a moment ago, suggesting that the one innate right and the rights that follow from it cover much of morality. It is not that Kant thinks that rights cover all of morality; they omit, for example, duties arising from the Doctrine of Virtue (Part II of *The Metaphysics of Morals*) and those duties arising from a mix of the a priori and a posteriori.[16] 'Natural' rights, in contrast to 'positive' rights, rest only on a priori principles, specifically on The Universal Principle of Right.[17] More fully, a 'natural right', Kant says, 'is one derived from *a priori* principles for a civil constitution'.[18] Of the several natural rights Kant derives from the one innate right, he mentions, among others, rights to procedural and distributive justice,[19] to retributive justice (a right exercised on our behalf by the sovereign power),[20] to grant clemency (also the sovereign's exercise),[21] to have one's reputation defended against unjust charges even after one's death,[22] to marriage if one's partner wishes to use one's 'sexual attributes',[23] to help in dire need,[24] and perhaps also not to be subjected to gratuitous suffering (I can find no explicit mention of this final right in the texts, but it is a negative right; it can be publicly commanded and enforced).

There is much overlap between what Kant says of 'natural rights' and my personhood account of 'human rights', because they are both centred on the idea of respect for persons. But there is also a great difference. What I mean by 'liberty' is freedom to pursue one's conception of a worthwhile life; liberty is one among other rights, the other ones on the same high level of abstraction being autonomy and minimum provision.[25] These rights are protections of something quite specific: our status as normative agents. What Kant means by 'freedom' is much broader than this: it is the area of action left to us after excluding what we are required to do and prohibited from doing by the Doctrine of Right. So what I called Kant's fateful move does indeed result in a list of rights considerably longer than the ones in the Enlightenment tradition. As I pointed out earlier,[26] although that tradition includes procedural justice (fair procedure in law) among human rights, it

strikingly does not include many forms of distributive or retributive justice or a right to grant clemency, highly important moral matters though these be. Nor does it include a right to marriage to one's sexual partner, although there might be a moral case for it, though the feeble one that Kant himself mounts is not it.[27] And if Kant believed there to be a natural right not to be made gratuitously to suffer, that does not count as a right in the Enlightenment tradition either. Recall my earlier example of one spouse's subjecting the other to relentless unpleasantness, which causes the other spouse greatly to suffer but not so much, or in such a way, as to lose personhood.[28] Again, most persons raised in the use of the language of the Enlightenment tradition of natural rights would find it counter-intuitive to say that the first spouse thereby violated the second's human rights. The first spouse does the second a serious moral wrong, but that is different. In Kant's hands, natural rights cover not only more ground than those in the Enlightenment tradition; they also have a different moral weight. For Kant, natural rights are absolute. In the tradition, they are not.

What should we make of these differences? Are they so great, especially in the extensions they yield, that we should doubt that Kant's account is, after all, an account of 'natural rights'? Or is his high-level Doctrine of Right, in association with the Categorical Imperative, so compelling that we should be willing to abandon the tradition, including our linguistic intuitions arising from it. It is clear from Part I of the *Metaphysics of Morals* that Kant's interest is in spelling out what can be derived from the Categorical Imperative, in particular from the Doctrine of Right, and not at all in accommodating how those around him at the height of the Enlightenment were using the term 'natural right'. Kant simply commandeers the term to do service in his grand theory.

To return now to our question: does Kant's account of natural rights establish their co-possibility? I think not. I am not persuaded by Kant's case for the Categorical Imperative and the Doctrine of Right. But suppose that one were so persuaded, and furthermore accepted his notion of natural rights. That would still leave the Enlightenment notion, with its considerably different extension. The only way that the Enlightenment notion would disappear is if it were to be shown to be so flawed that it would be better abandoned. I said earlier that, though arguments to that effect have occasionally been advanced, none has succeeded in establishing anything approaching such a strong conclusion,[29] and the last chapter offered my reasons.[30] In any case, it is the Enlightenment notion that virtually all of

us, Kantians or not, understand by 'human rights' and have in mind when we ask whether human rights are co-possible. So let us, for now, answer the questions so understood.

The example of detention without trial that I gave a short while ago used a government as one of the agents involved, because that is the form in which we actually encounter this sort of case. But the presence of a government is an unnecessary complication—a complication because the government's duty to act here may be thought to arise from more than just the rights to self-defence of the citizens for whom it acts. It might, for instance, be thought to arise from a contract-like relation between the government and its citizens. But the example can be simplified. The members of a certain group, let us say, have a human right to defend themselves against a clear and present danger to their lives, and the only effective way for them to exercise it is to round up those members of another group threatening them, knowing that it is likely that some of those rounded up will be innocent. Their exercising their right to self-defence, derived from their right to life, conflicts with the innocent detainees' exercising their legitimate right to liberty. This, then, is a simpler counter-example: a case of some persons' exercise of a human right conflicting with other persons' exercise of a human right.

If there are absolute rights, they must be co-possible. But it is false that, on the Enlightenment notion, natural rights are co-possible. So, by *modus tollens*, not all of them are absolute. On my personhood account everyone has maximum liberty compatible with equal liberty for all, 'equal liberty' here meaning that all persons possess the same right to liberty, where the content of the right is not so full, nor capable of being made so full, that it guarantees harmony in the exercise of all human rights.

3.4 CONFLICTS BETWEEN A HUMAN RIGHT AND OTHER KINDS OF MORAL CONSIDERATION

Do human rights sometimes conflict with welfare? There are abundant pseudo-conflicts here too. Many writers regard any restriction on one's doing what one wants as an infringement, no doubt often justified, of one's liberty. Liberty, they say, conflicts with efficiency. The one-way traffic restriction infringes my liberty, but this minor liberty is outweighed by the increase in efficiency. But, as we saw earlier, liberty does not protect one in doing

whatever one wants. Liberty, the moral and political value that is the ground of the right, is not even at stake in this case. There is no conflict.

Still, there are genuine conflicts here as well. A country, let us say, decides to hold a referendum on whether to devote a percentage of its GDP to foreign aid and, if so, what percentage it should be. In a democracy, the right to autonomy, at least in modern conditions, requires consulting citizens on certain major decisions affecting what is done to them and for them. A crisis arises in a neighbouring country, causing great suffering (but not deaths, let us suppose, in order to simplify things) and needing quick remedy. The government announces that, as there is not time to await the results of the referendum, it will send aid. The suffering is widespread and severe, it explains, the neighbours have long-standing ties to us of friendship and mutual help, and the autonomy denied our own citizens is not of a particularly high order (after all, it is a one-off action on the part of the government, and if the coming referendum goes against further foreign aid, it will not be repeated). There must be some level of suffering and some level of importance of an exercise of autonomy at which the suffering outweighs the loss of autonomy.

That is an example of the most widely discussed sort of conflict: a right–welfare conflict. But there are kinds of moral consideration besides human rights and welfare. There are, for example, considerations of justice, and though some parts of justice overlap with human rights, not all do. Let me briefly remind you of my earlier argument, which was an appeal to strong and widespread linguistic intuitions.[31] Justice has many departments: retributive justice, distributive justice, procedural justice, fairness (there are reasons, I should say, for thinking that fairness is not exhausted by those departments already mentioned), and still more. Now, if you free-ride on the bus because you know that no harm will come, as the rest of us are paying our fare, you do not infringe my human rights, though you do, clearly, act unfairly. That helps to explain why the tradition regards the whole of procedural justice in courts as a matter of human rights, but not the whole of distributive justice. It regards the requirements of procedural justice in courts as human rights because, to put it briefly, they are important protections of life, liberty, and supporting goods—all necessary conditions for agency. Of course, human rights have their own distributive consequences, say the right to the minimum material resources needed to function as an agent. On distributive matters above that level, however—say, whether resources or welfare should be distributed equally, or whether deviations from equality

that make the worst off better off should be allowed—human rights are silent.

If the domain of rights does not exhaust the domain of justice, then there may be right–justice conflicts. For example, think of a properly convicted criminal who is imprisoned or executed. One is supposed to have human rights simply in virtue of being human. The criminal, in virtue of the crime, does not cease being human in the relevant sense—namely, a normative agent—so the criminal, it seems, retains all human rights, including rights to life and liberty. If one thinks that what justifies punishment is its good consequences, then one has here a potential right–welfare conflict. But many people would insist that what justifies punishment is, or is also, desert. Would we have here, if that were so, a right–justice conflict?

It would, of course, be a grotesque theoretical embarrassment if human rights were to prove an obstacle to fair punishment. The embarrassment was meant to have been avoided by a doctrine of forfeit.[32] A criminal forfeits human rights. If that were so, then this case too would thereby become another pseudo-conflict. The doctrine of forfeit, however, is a factitious measure, never deeply worked out. What exactly, according to the doctrine, does a murderer or a thief forfeit? Rights generally? The right to life? The right to liberty? The right to security of person (e.g. the right not to have one's hand cut off, as it might be under Shariah law)? The appropriate way to answer those questions is by appeal to desert: what punishment would fit the crime? Punishment involves taking away something typically valuable to human beings—for example, life, liberty, or property. What, and how much, good is to be taken is determined by the offender's desert. What the offender might be thought to forfeit is not a right *simpliciter*; the 'forfeit' is whatever punishment turns out to be just. There are cases in which different punishments are all equally just: the guilty party, let us say, might appropriately be given either six months in jail or a £20,000 fine or three years' community service. So, in the language of 'forfeits', the offender would forfeit either the right to liberty to the extent of six months or the right to property to the extent of £20,000 or occasional liberty to the extent of three years. But what are important here are the judgements about desert, from which one can derive, if one should care to, judgements about what is forfeited. But why should one care to? To speak of 'forfeit' suggests that the right in some way disappears from the scene. But it does not disappear. On the contrary, we have just seen that the personhood account leaves no space for forfeits; an offender is still a person. That is why an offender retains,

among others, a right not to be tortured. It is more perspicuous to say that the demands of justice can sometimes, and to some appropriate degree, outweigh the protection of human rights. That is why it is indeed reasonable to talk here in terms of conflict—a conflict between human rights and justice.

Might justice, then, not only outweigh a human right but also upset the calculations that I say go into the resolution of conflicts? Might the resolution of a right–right conflict, for instance, based as it is on degrees of loss of personhood, be altered by considerations of distributive justice—say, maxi–min? In principle, yes; but rarely. We can apply the principle of maxi–min only if we know the welfare levels of those involved, either individually or as a group. Think of cases of temporary arbitrary detention. We typically do not know the welfare levels of those involved, either individually or as a group. Even if we did, it would be unlikely to make a difference, because so many lives are at stake.

To summarize: there are, then, right–right conflicts, right–welfare conflicts, right–justice conflicts, and possibly more. I think that there are indeed more, some of which are neither dissoluble nor resolvable.[33]

3.5 A PROPOSAL AND A QUALIFICATION

There is a general requirement on the resolution of conflicts of values: if it is not to be arbitrary, one must know what values are at stake and how to attach weight to them. No matter how basic in the whole moral structure human rights may be, there is still language available that allows us to articulate, at least in part, why they are so valuable. It need not be the language of ends. It could be the language of duties. Kant, as we saw, speaks of persons' having 'dignity' in virtue of their freedom, and of their dignity's giving them inviolability. Non-Kantians can make a similar point. We all need to understand why persons are regarded as especially valuable. Since the late Middle Ages many writers have explained it by pointing to the special feature of humanity that I have been stressing: our capacity for normative agency. We can easily understand why the term 'dignity' was attached to our status as normative agents, and why other forms of animal life were thought to lack that dignity. It is not that the direction of explanation need run solely from 'dignity' to 'human right'; it might be that adumbrating human rights is an indispensable way of understanding that particularly vague term 'dignity'.

Our agency is far from simple. It has parts: autonomy, liberty, and minimum provision. One can lose one part and not others. And each part itself can be at stake in different degrees. There are, for example, minor liberties and major ones, minor exercises of autonomy and major ones. The ground for our drawing this difference between liberties and exercises of autonomy is their degree of centrality to personhood. For instance, we are at liberty to visit other countries, other cultures, other political systems, and people with other attitudes. The United States government's prohibition of its citizens' visiting Castro's Cuba was an infringement of their liberty, but less drastic than a prohibition on all foreign travel would have been, which in turn would have been a less serious infringement than a prohibition of all foreign contact. So behind the idea of major and minor infringements of rights is the idea of an attack on something nearer to or further from the centre of one's agency. More or less of one's liberty, more or less of one's autonomy, can be at stake at different times. What is more, there is the temporal dimension; the loss can be for a short or a long time.[34]

Is my talk here of degrees of agency inconsistent with what I said earlier? I say now that a person can lose one component of agency but not others, and each component to different degrees. Yet I said earlier that personhood is a threshold concept: once inside the class of persons, there are no degrees of being a person. There are two senses of agency that chiefly concern us. There is the sort of agency that makes us bearers of human rights—namely, our capacities for autonomy and liberty—and there is the sort of agency that human rights are meant to protect—that is, not only the possession of these capacities but also their exercise. It can happen, either through the action of other agents or, more commonly, through illness (polio, say, or motor neurone disease) or accident that one loses one's natural capacity to exercise liberty. One can no longer pursue various parts of one's conception of a worthwhile life because one has insufficient control over one's body. Is this a loss of agency? And of one's human rights? It is clear to me that, however one explains personhood and the possession of rights, such an alert person trapped inside a non-functioning body must be classed as a *person* and as a *bearer of rights*. This is uncontroversial. It can also be accommodated within the personhood account. Because of the value of personhood, we have a duty, correlative to human rights, to restore that person's capacity to act. We think so, for instance, in the case of the crippled; we build special access for them to schools, museums, concert halls, and so on. And we can similarly help

persons trapped in non-functioning bodies; we can, for instance, become their surrogate arms and legs; we can become the executors of their rational conceptions of a worthwhile life—perhaps even including a merciful death. I shall come back to that last issue later.[35]

Look now at the idea that human rights are absolute. We have seen that human rights can conflict—both one with another and one with other kinds of moral considerations. Therefore they cannot be absolute. There are, it is true, intuitions that seem to support absolutism. Some pairs of values are such that, no matter how much one of them increases, it can never reach the level of the other. Elsewhere I have called this sort of relation 'discontinuity'.[36] We constantly meet welfare–welfare conflicts, and we all believe that utilitarian calculation is usually the appropriate way to resolve them. For instance, we weigh up how many airplane passengers are convenienced, and how much, by there being early morning and late night flights, and then weigh up how many people living under the flight path of the airport are disturbed (assuming that it is no worse than disturbance), and how much. But certain other cases seem radically different. Suppose that, if I were to live the only sort of life that I regard as worth living, my neighbours would be upset and distressed. But, unlike in the airport case, we do not think we should now count heads. We do not, because upset and distress, so long as they remain ordinary upset and distress, can never add up to anything as important as one's being able to live out what one regards as a worthwhile life—not if there were a hundred neighbours upset and distressed, or a hundred thousand, or a million. Upset and distress are not the kind of thing that could ever match the centre of a person's liberty.

But the existence of discontinuities does nothing to support the existence of absolute human rights. On the contrary, the admission of discontinuities is compatible with utilitarianism; it does not even introduce an incomparability; it is indeed an especially emphatic comparison. And the example I have just used constitutes a discontinuity only because the two conflicting values come in degrees. It matters that it is my most *central* liberty at stake, and *wholly* at stake. It matters that my neighbours do not experience anything worse than upset and distress. If, on the contrary, they were caused considerable suffering, and I lost only briefly a relatively minor liberty, we should think again.

There are, I have proposed, right–right conflicts. Virtually everyone would agree that an important part of their resolution comes by determining the degrees of the values constitutive of personhood at stake. The innocent

detainees have their liberty violated for the period of their arbitrary detention—not totally violated, because they ought still to retain several liberties while in detention. The innocent victims of terrorism, however, lose their liberty, and all other freedoms and protections, totally, if they lose their lives. Now, comparisons need a bridging notion: that is, some conceptual background that supplies the terms in which the conflicting items are compared.[37] The bridging notion need not itself be a substantive value; it could be the notion of 'value' itself: 'the relief of this suffering is more valuable than the temporary loss of this element of autonomy'. Or it could be, say, the notion of a 'reason': 'this is a stronger reason than that'. A bridging notion in the resolution of right–right conflicts is protection of personhood. Saving the lives of twenty or thirty innocent bystanders is more protective of personhood than detaining a handful of innocent suspects for six months is destructive of it.

Much the same can be said about right–welfare conflicts. Once we understand the value of human rights, we see how their value admits of degrees. One might not be prepared to accept the denial of many people's very status as agents to relieve a certain suffering, but one might, if the suffering were great enough, accept some partial, short-term surrender of autonomy to avoid it. In this case one is making a relative judgement about importance to life. That broader notion, importance to life, is a bridging notion in right–welfare conflicts.

We have now reached a point in the argument at which the bottom-up approach must rise appreciably in abstraction. It may not have to rise as high as some top-down approaches, but it must move in that direction. Both of the proposals for resolution of conflict that I have just sketched would command wide acceptance, as far as they go. Virtually everyone would agree that the considerations they make central are indeed central. But, as they stand now, the proposals look consequentialist. Can we not then resolve all conflicts involving human rights by calculating consequences for the quality of lives? And if so, are not human rights co-possible after all, now on a consequentialist basis? Many writers would deny it. The resolution of conflict, they would insist, is more complicated than that. We must take account of certain deontological elements, some would say; or we have still to introduce certain teleological but non-consequentialist elements; or we have yet to accommodate the foundational role of the virtues.

This, as forecast, brings us to the heart of normative ethics. The heart of normative ethics is a large subject to be introducing into my argument in

what will have to be a fairly summary way, but anyone wishing to understand human rights must at this point, willy-nilly, face several questions. It is important, at the very least, to mark out what these questions are. I have said something in answer to them elsewhere;[38] but more still is needed, although it will have to wait for another occasion. I think that many readers will have accepted the claims I have made so far about human rights—in effect, the personhood account. They are likely to agree that personhood, at least, is a large part of the explanation of human rights. But one cannot expect wide agreement about the heart of normative ethics. Anyone who disagrees with what I shall now go on to say must find something else to put in its place. So it is worth marking out what that place is.

There are limits to our capacity to calculate consequences. We do not, of course, need our calculations to be certain; we generally live by probabilities. But we do need the probabilities to be high enough for us to be prepared to stake our lives on them. Clearly, sometimes our calculations have a high enough probability, and sometimes not. The crucial question, then, is how often they do not, and how central to moral life these failures are. If, as I think, the most plausible form of consequentialism is highly indirect, then a consequentialist must be able to answer some such question as: which set of rules and dispositions, if they were the dominant ones in our society, would have best consequences in the society at large and in the long run? No one has ever come even close to answering that question. The closest we have come is certain cost–benefit analyses, but they have gaping holes. They have trouble finding adequate expression in their formalization for certain key values—famously, for the value of human life and the value of the environment (especially those environmental values that cannot be reduced to good and bad outcomes for human beings). The assumptions on which the calculations in cost–benefit analysis are based are often so oversimplified that we are rightly hesitant to act on them. And an indirect consequentialist would have to calculate on a vastly greater scale than any cost–benefit analyst has yet attempted.

Some will think that this worry can be met by what Bernard Williams called 'the early days reply'. Purely secular ethics, the reply goes, is in its youth—a little over two centuries old, at most—and our modern ethical 'theories' have been developed as yet only in the roughest terms. But this reply, though consisting of two true claims, is reassuring only if the obstacles we face in calculating consequences are the sort that will yield to more time. But that is an extraordinarily strong assumption. It is an assumption that

many philosophers today indeed make, largely because they assume that if an ethical 'theory' needs agents to have certain powers of intellect or will, agents have them. But that assumption has only to be articulated to be seen to be ridiculous.

One of the greatest under-discussed questions of ethics is: what, in fact, are the capacities of the agents whom ethics seeks to regulate? We are able *sometimes* to calculate, fairly reliably, the good and bad consequences of very large-scale, long-term social arrangements. If the arrangements being contemplated are extreme enough, we clearly can. For example, if we were to consider moving to a world without the rule of law or fair means of resolving conflict, we could make a fairly safe guess that we should be worse off in it. But the cases that we think of as live options, worth taking seriously, are not the extreme ones. And when they are not extreme, we flounder. Take a realistic case: suppose that our rules and dispositions concerning respect for innocent human life become less strict; we start deliberately killing non-combatants in time of war (Shock and Awe tactics); we use terrorism widely as a political instrument; we make troublesome political opponents 'disappear'; surgeons begin to kill one patient on the sly to save several others; and so on. They are all cases that, if justified, are justified on grounds of better consequences overall. But are they better? I doubt that anyone can answer that question to a reliable degree of probability. And it is hard to see how 'early days' could matter. We know the great problem we face: identifying and collecting all the relevant information and then expressing it in a form that will allow reliable reduction to a single answer. How will later days make that any easier?

This example of lessening respect for innocent human life suggests something further. Perhaps most people would agree that, if what we are comparing are not very different possible worlds, we cannot now, nor can we in a likely future, do the consequentialist calculation to a reliable degree of probability. But some persons say that all that consequentialism needs is a smaller-scale, more manageable calculation about certain changes from the *status quo*. But the case of lessening respect for innocent human life is a change from the *status quo*, and not more manageable for that.

A fairly common defence of consequentialism is that any relevant limitation in human understanding and will for which there is adequate empirical evidence can simply be incorporated into the calculus. We should ask: which dispositions, rules, and principles would, if they were dominant in our society, have best consequences, given agents with such-and-such limitations in understanding and will, over society as a whole and in the long run? But

this is no answer to the doubts. It just makes the already dubious calculation even more difficult to do.

Another common reply is to cite the distinction between a decision procedure (the way we should actually decide what to do) and a criterion of right and wrong (what in the end settles what is the right thing to do). It does not matter, the reply goes, that we often cannot calculate consequences to a sufficient degree of probability in deciding what to do; we do not have to; instead, we are right to follow ingrained dispositions or well-established rules or the like. But that is too weak a defence of consequentialist calculation. There are epistemic constraints on a criterion of right and wrong too: a 'criterion' largely beyond our capacity to apply cannot serve as a criterion; it will not perform its function of sanctioning our decision procedure.

There are several other ethics-shaping human limitations.[39] Let me mention just one more. Not every action is within normal human motivational capacity. If persons are, as surely we should want them to be, capable of love, affection, and deep commitment to particular persons, institutions, careers, and causes, then certain actions will be beyond their motivational reach. One cannot enter into and exit from these commitments at will—say, as calculations of consequences might demand. When the ship goes down, I shall save my own child rather than a larger number of unknown children. On the most plausible interpretation of '*ought* implies *can*', I cannot ignore my own child and save the others, so it is not the case that I ought to do so. Many philosophers believe that my lack of obligation to save the other children is grounded not in motivation but in morality: I am duty-bound to give each person his or her due, and much more care is due from me to my own child than to other children. We are dealing here, they say, with a moral limit, not a motivational one. But we are also dealing with a motivational limit. Motivational limits are a reality, and, because '*ought* implies *can*', these limits have moral consequences.

There are, of course, weighty replies to what I have claimed about the limits of human motivation. Let me give just a sample of them. Human motivation, it will correctly be said, is plastic. For one thing, motivation can be enlarged by knowledge. Charities know that a single photograph that brings home the reality of a famine can spur many of us to reach for our cheque-books. Still, greater knowledge does not render characteristic human agents able to meet any demand that any ethics might impose. Famine relief workers in the field, well motivated and with ample understanding of the suffering they see all around them, do not generally sacrifice themselves and those they love to the

point where a further sacrifice would exceed a further benefit to the starving. It is true that there are often good impartial-maximizing reasons for those aiding to have more than those aided—for one thing, they must be able to carry on helping—but relief workers generally do not sacrifice themselves to that point either.

Perhaps, then, the need is for a more inspiring ethics. There certainly are dreary, narrow, depressing ethics. An exciting ethics—perhaps certain religious ethics or the Platonic vision of the Good that Iris Murdoch thought could inspire a prisoner in a concentration camp to take a stranger's place in the queue for the gas chamber—would help us to rise to ethics' greatest demands. The trouble, to my mind, is that the sorts of ethics that can so revolutionize motivation are not plausible, and the sorts that are plausible cannot so revolutionize motivation. But cannot unreachable goals still play an important ethical role? They stretch us, and most of us are undeniably less benevolent than we could and should be. But it is an oxymoron to speak of my adopting what I accept is an 'unreachable goal'. Whatever grunts and groans you come upon me making, I cannot seriously tell you that I am *trying* to jump unaided a hundred metres into the air.

But cannot the right education or rigorous training enlarge motivation? Do we not regularly see how military training can turn quite ordinary persons into ones willing to die for their country? During the Cultural Revolution, the Chinese government managed to produce the Red Guard, some of whom became free enough from earlier patterns of behaviour to turn their dissident or bourgeois or merely learned parents over to the police. But this sort of training does not succeed widely or for long. The fanatical personality of the Red Guard and their most ardent collaborators was unstable. Some of the young among them turned up years later in the tents in Tiananmen Square.

We cannot understand much about human rights until we know a fair amount about moral norms in general. I have only most roughly sketched a normative ethics that is teleological but not consequentialist. It is teleological somewhat in the way that Aristotle's ethics is: the only values used in the derivation of moral principles are the ends of human life, but more enters the derivation than simply these ends. Besides the personhood ground for human rights, there is also the practicalities ground, and although some practicalities come down to consideration of quality of life, not all do. For example, we have to fix boundaries for the right to security of person. The boundaries that any particular society chooses will be to some degree arbitrary. They will not be chosen because we can calculate to a sufficient degree of probability that

this would be the maximizing place to fix them; we can eliminate extreme options, but there will remain very many alternatives in a large middle range which we simply cannot rank reliably. What is more, the rule 'Don't deliberately kill the innocent' is based on the great value that we attach to human life, but not only on that. It is a rule shaped considerably by human limitations. We adopt the rule, which, given the high value of human life, we regard as demanding *strict* respect. The strictness of that respect will manifest itself in various ways, importantly in our demanding that any exception to the rule have an exceptionally strong justification, as perhaps there is for carefully circumscribed euthanasia. The justification in the case of euthanasia would probably be in terms of good consequences, or largely in those terms: the benefit to those in need of a good death would outweigh the costs elsewhere in society. But good consequences do not capture the element of *policy* in the rule 'Don't deliberately kill the innocent'—the policy of making an exception only when it is especially strongly justified. And the package made up of the rule and its exceptions does not get its authority from its maximizing good consequences. We do not know whether it does. There are very many other packages made up of a rule about respect for life and its exceptions that we cannot rank as to good outcomes. At times, the only moral life open to us involves *respecting* values, not *promoting* them. By 'respecting' the value of human life, for example, I mean primarily, but not solely, not oneself taking innocent life; by 'promoting' life, I mean bringing about its preservation as much as possible by any means open to one.[40] We must come to terms with how certain limits to human nature determine limits to moral obligation.

From the eighteenth century to the present, most philosophers, dazzled by the success of natural scientists, pre-eminently Newton, went in search of highly systematic theory. Moral philosophers sought the reduction of all our varied moral thought to one principle, or to a small number of them. Kant, consciously inspired by Newton, stopped there: moral obligation, he thought, could be reduced to a single a priori principle, the Categorical Imperative. Hume, Adam Smith, and the Utilitarians went further; they looked not just for high system but for empirical system. And that is roughly where we are now: committed to moral 'theories' at a Newtonian level of abstraction, and proposing to assess them with a coherence or reflective equilibrium test that is effective in the natural sciences only because of the presence of features that are absent in ethics.[41]

I think that we have no choice but to take a highly practical turn in ethics, not just to ensure that our abstract principles are adequate to our practice,

but also to accommodate the ways in which our practice—our human nature with all its limitations and the needs of our actual societies—determine the content of our principles. My proposed turn to the practical makes central to ethics both following rules and training dispositions—for example, turning to rules of the nature of 'Don't deliberately kill the innocent'. This rule is little like a Newtonian principle; it has major elements of policy in it. We follow it not because doing so is best for everyone impartially considered; we just follow it. It is practical not just because it guides practice, but also because it is shaped by the sort of practice possible for agents like us. We live the kind of moral life open to us. I say 'my proposed turn', but it is a turn taken a long time ago by common-sense ethics and the law, neither of which has needed a highly abstract, systematic morality behind it in order to be authoritative and effective. What I approve of here is not the present content of common-sense ethics, which leaves a lot to be desired, but its lack of pretension to system.

To return to our subject: these human limitations inevitably shape human rights. The rule 'Don't deliberately kill the innocent', was, historically, the first component of the right to life. In the seventeenth century, indeed, the right to life was largely seen as little more than the supposedly negative right not to have one's life taken without due process. If thousands of citizens in Argentina and Chile would feel more secure if a radical reformer were made to 'disappear', the conflict between their welfare and the reformer's right to life cannot be settled just by consulting effects on quality of life. We are unlikely to be able to calculate them well enough, and, in any case, they would not adequately capture the rule. We simply must not deliberately kill the innocent unless the case before us falls under an especially strongly justified exception, and that some middle-class Argentines or Chileans, rightly or not, would feel somewhat more secure is certainly not one of them.

That there are such moral rules complicates the resolution of conflicts. Sometimes we can resolve conflicts involving human rights by deciding the severity of effects on personhood, or on the quality of life. But when a moral rule derived not only from the quality of life enters consideration, then we cannot. Moral deliberation must then take place largely on the common-sense level at which it occurs in ordinary life. We usually think in terms of 'murder' (i.e. 'deliberately killing the innocent'), 'the parent–child relation', 'a right to life', 'a right to free expression', and so on. And although the resolution of conflict puts pressure on us to rise to language of greater abstraction, not all of this everyday vocabulary can be left behind.

3.6 A STEP BEYOND INTUITION

Human rights are resistant to trade-offs, but not completely so. The strongest version of this non-absolutist view is that only something on the order of a catastrophe, such as a nuclear holocaust, can outweigh a human right. This, as we have seen, is Robert Nozick's view.[42] I have proposed various considerations that should enter the resolution of conflicts. My proposal makes it seem that the detention of a smallish number of suspected terrorists for only a few months would be justified if it were to avert a serious threat of the nuclear destruction of half of Manhattan. Yet the destruction of half of Manhattan, for all its terribleness, is well short of nuclear holocaust. On my proposal, exceptions to human rights are unlikely to be quite so exceedingly rare. But near-absolutists are likely to complain that I have simply failed to introduce into the scales precisely the crucial consideration: respect for persons. But I have not failed to introduce it: I have spelled out the values that the long human rights tradition attaches to our status as persons. What have I, or the tradition, left out? One possible answer is that, though I have included *an* interpretation of respect for persons, I have not included the Kant-like one that most deontologists have in mind. I say 'Kant-like', because Kant is an absolutist, and we are now addressing strong forms of non-absolutism. The near-absolutist, Kant-like interpretation must therefore have sufficient richness to show where the turning point is and why it is where they put it. I cannot find a plausible one. My response is short of conclusive, and I shall return to it presently.

The near-absolutist view is a special case of what I shall call the 'common' version of the non-absolutist view. That version holds that an exception is justified, not by a simple surplus of value over a human right, but only by a sufficiently great surplus, where sufficiency may be fixed at catastrophe or somewhere short of it. But what are the two scales appealed to in this common version? We have, it appears, to be able to tell when the competing value *just* exceeds the human right, and then, on a new scale, when the excess is enough. Where is the conceptual complexity that will allow the construction of those two scales? Our dim and undeveloped idea here is, I should guess, that the first comparison appeals to a scale of well-being, and the second to a scale of moral importance—the first to quality of life, the second to normative weight. But that cannot be right. Take an example. We

decide that the relief of such-and-such a degree of suffering justifies the loss of thus-and-so aspects of autonomy. If earlier I described our thoughts about this case accurately, the bridging notion here is gain and loss to the quality of human life: the gain of the relief is greater, in the case I described, than the loss of the fairly minor elements of autonomy. This makes it seem to be the first of the two supposed measurements. In terms of quality of life, we judge that the relief of suffering just exceeds the loss of autonomy. But it just exceeds it in terms that constitute justification: this is a gain that justifies the loss. But to be justified is (simply) to be justified. The idea of a *sufficient* surplus of justification is nonsense.

So this makes it seem that, contrary to our first impression, we must have here not the first but the second of the two supposed measurements. Are there any materials, then, out of which to construct the first scale? It might be suspected that my reintroducing the example of the conflict between autonomy and suffering, along with my earlier remarks about it, is question-begging. My earlier explanation of this conflict, it might be thought, already revealed my teleological drift; quality of life comparisons are relatively well understood and the mark of the modern economic mind. But to go along with the teleological drift may be simply to assume that deontology is wrong. So let us keep looking a little longer for that other scale. A judgement to the effect that the relief of suffering justifies the partial loss of autonomy, we are now hypothesizing, must be employing the final scale, the scale on which the welfare consideration overrides the right. We need, then, the materials for constructing the initial scale, the scale on which the welfare *just* exceeds the right. But there are none.

Let me, therefore, drop this two-point model of the non-absolutist view; it may be obscuring the central point in deontology: that the right is often prior to the good. Is not the requirement of a *sufficient* surplus of good simply a representation of this independent weight of the right, and can this not be done without the two points? On this suggestion, a deontologist's judgements of when we reach sufficiency may be ground-floor and intuitive, but I am in no position to object to that feature of them. So may be the judgements that I have to make to resolve conflict. But think again of the example of conflict between autonomy and suffering. Suppose we weigh solely the good at stake for the persons involved, leaving the independent deontological weight of the human right aside for the moment. On the one hand there is the suffering of the people in the neighbouring country, and on the other our loss of autonomy in our not having a say in the decision.

One can imagine being able intuitively to judge in this particular case that the relief of suffering just exceeds the loss of autonomy, again basing one's decision solely on the goods for the persons involved. But there still seems to be no need for a further judgement; we already have here the justification of the trade-off.

To try to avoid this conclusion, we might say instead that the judgement that I have just imagined already incorporates the independent deontological weight of our right to autonomy. But this would make the trade-off point using this Kant-like interpretation of respect for persons the same in this case as the trade-off point using my interpretation of it. It would not, though, show that the points will be the same in all cases; whether they will be the same depends upon what the Kant-like interpretation actually says. I cannot claim unproblematic clarity for my interpretation of what is at stake, but deontology has always had to struggle with the obscurity of the idea 'respect for persons' when it must give us reasons not just about right and wrong but also, as in our present cases, about when the consideration 'respect for persons' is just outweighed, in action-justifying terms, by quality of life. We hope for more content to the Kant-like idea of 'respect for persons' so that we can know what is going on in this judgement. The mere fact that people are prepared to make these rankings hardly shows that the rankings are rational; the human psyche finds paths to rankings apart from rationality.

Perhaps one might get this more Kant-like interpretation that we are looking for from contemporary Kant-inspired contractualism. But there is Judith Jarvis Thomson's powerful challenge to contractualism to contend with:

> For my own part, I cannot bring myself to believe that what *makes* it wrong to torture babies to death for fun (for example) is that doing this 'would be disallowed by any system of rules for the general regulation of behaviour which no one could reasonably reject as a basis for informed, unforced general agreement.' My impression is that explanation goes in the opposite direction—that it is the patent wrongfulness of the conduct that explains why there would be general agreement to disallow it.[43]

T. M. Scanlon, at whose version of contractualism this objection was aimed, responds:[44]

> The contractualist formula that Thomson quotes is intended as an account of what it is for an act to be wrong. What *makes* an act wrong are the properties that would

make any principle that allows it one that it would be reasonable to reject (in this case, the needless suffering and death of the baby).

Scanlon reasonably distinguishes here what it *is* for an act to be wrong from what *makes* it wrong. This leaves a great deal of moral thought to be conducted in terms of particular right-making and wrong-making features of acts (e.g. suffering and death) and not by pondering whether the act fits the criterion of what wrongness *is*.

In the case we are considering, what make the act right or wrong are our loss of autonomy and the suffering of the people in the neighbouring country. They are also what make it right or wrong to a particular degree. In thinking about this case, we would consider how intense and widespread the suffering is and how large-scale or small-scale, and long-term or short-term, the loss of autonomy is. This suggests that the non-absolutist deontologist following this contractualist interpretation of respect for persons and an agent following my interpretation would not only arrive at the same conclusion, but arrive at it for the same reasons.

It is a commonplace that human rights are particularly hard for utilitarianism to explain. I agree, but would want to add, what is not a commonplace, that they are no less hard for deontology to explain. And we need an explanation not just of right–welfare conflicts, which is what I have been discussing and which the model of two scales was, I believe, designed to fit. We also need explanation of right–right conflicts, right–justice conflicts, and probably yet more.

3.7 SOME WAYS IN WHICH HUMAN RIGHTS RESIST TRADE-OFFS

Teleology can fairly readily explain why human rights are not *too* resistant to trade-offs. But can it explain why, in the first place, they are *resistant*?

I have suggested that in different situations a human right can be under threat to different degrees. So clearly can welfare. To resolve conflicts between them, we look for how much under threat each is. Once one deliberates in terms of 'how much', we are well on the way to wanting to minimize loss of personhood or maximize its protection. And once on the way to minimization and maximization, the fear understandably arises that we are on the way to the sort of moral mathematics that will justify, say, killing one person to save five.

What I have suggested, though, is not any kind of utilitarianism or consequentialism, but a kind of teleology. Although utilitarianism and consequentialism are forms of teleology, they restrict the test of right and wrong to the production of as much good as rationality requires—maximizing, say, or, on a different view of rationality, satisficing. Teleology allows yet other ways of basing the right on the good.[45] To explain this further, let me turn to various ways in which, consistently with teleology, human rights can be resistant to trade-offs.

One way comes from the great value we attach to our personhood. That we attach such great value to it, especially once we are above a minimum acceptable level of material provision, which virtually all people in the First World are, means that it has a general resistance to trade-offs with welfare. Once above the minimum acceptable level, it takes some unusually large amount of welfare to outweigh personhood. This remark echoes talk about a 'sufficient surplus', but it does not require two scales.

A second, particularly striking way that personhood resists trade-offs is through discontinuities. Some values—an obvious case being our status as persons—are such that no amount of certain other values can ever equal or surpass them. The value that we attach to personhood is unchallengeable by these other values. This remark echoes what is sometimes said in support of there being absolute human rights; but it does not really support it.

A third way I discussed a short while ago. I appealed to a distinction between 'respecting' and 'promoting' goods.[46] Respecting goods, as well as promoting them, can be a teleological position; both positions can hold that the good is basic in the moral structure and the right derived from it. My example earlier was that life must be respected, and that one must simply follow the norm, 'Don't deliberately kill the innocent'—follow it because that is the only moral life available to the likes of us, though one might also adopt the policy that exceptions will be allowed only so long as the case for them is especially convincing. This talk of requiring an *especially* convincing case echoes the talk of a *sufficient* surplus of value, but it is not the same point. Talk of an especially convincing case introduces an epistemic scale, not another moral one. It is the statement of a policy—an openly conservative policy—for what to do when something as important as human life is at stake and our calculations of the goods at stake are altogether too shaky and incomplete and badly conceptualized for us to be willing to live by.

3.8 REPRISE

As my argument in this chapter, more than most in philosophy, moves step by step, it may be helpful to summarize it in that form.

1. We need to supply existence conditions for human rights.
2. The existence conditions for human rights in particular will have to have some substantive evaluative elements, in order to express the value to be attached to human status.
3. Those substantive evaluative elements then impose material constraints on the content of human rights: for example, the personhood account says that human rights protect not just anything that we rationally desire or that benefits us, but rather our status as normative agents.
4. Normative agency can be divided into parts: namely, autonomy, minimum provision, and liberty.
5. Each part of normative agency can be threatened by other persons, and threatened to different degrees: one can lose a little, or a lot, of one's autonomy, material provision, or liberty.
6. These degrees are determined by effect on one's personhood: for example, a minor liberty is one the loss of which detracts from one's personhood to a fairly small extent; a major liberty is one the loss of which detracts from one's personhood substantially.
7. When we can resolve conflicts involving human rights, we do so in the case of right–right conflicts by appeal to their effects on one's personhood; in the case of right–welfare conflicts, to those and to effects on welfare; in the case of right–justice conflicts, to those and to the weight of justice involved; and so on—though sometimes, in all of these cases, by appeal to a moral rule.
8. These resolutions need a bridging concept, or concepts, in order to allow us to determine the weight of conflicting considerations in sufficiently similar terms for us to be able to decide which is the weightier. The bridging concept need not itself be a substantive value; it could, for example, be the concept of prudential value itself. In the case of some conflicts there may be no bridging concept.
9. In the case of right–right conflicts, the bridging notion is loss/gain in personhood. In the case of right–welfare conflicts, the bridging notion must be broader: loss/gain in quality of life.

10. In both cases, the loss/gain has to be interpersonally comparable.
11. All of these loss/gain considerations are relevant to the resolution of the conflicts involving human rights that I have in particular considered, viz. right–right and right–welfare conflicts; so too are certain moral rules. There are no other relevant considerations.
12. Near-absolutists have no adequate explanation of the turning point they must posit. The more moderate (and more common) deontologists have no adequate explanation of how we can identify the two points they posit: a *simple* surplus of good and a *sufficient* surplus to override a human right.

4

Whose Rights?

4.1 THE SCOPE OF THE QUESTION

Human infants are not normative agents. Neither are human foetuses, nor the severely mentally handicapped, nor sufferers from advanced dementia. Do none of them, then, have human rights?[1] Perhaps we should not be so fixated on the award of the label 'human rights'. Do none of them at least have certain general moral rights simply in virtue of being human—only analogous to human rights, it may be, but for all practical purposes just as good?

4.2 POTENTIAL AGENTS

I shall concentrate on the case of children. My conclusion about them can, I believe, be extended to the other difficult marginal cases, which I shall return to at the end of the chapter.

Normal human adults have a kind of natural equality: they are all equally normative agents; they all cross the threshold into the class of such agents. This means that, when it comes to human rights, distinctions between man and woman, black and white, highly educated and little educated—in fact, all distinctions but for agent and non-agent[2]—are irrelevant. But it is not clear how to regard the distinction between infant and adult. As John Locke succinctly puts it, '*Children*, I confess are not born in this full state of *Equality*, though they are born to it.'[3] That the normal natural destiny of infants is to become agents must itself be a reason for an especially high concern for them. What makes infanticide in general a great wrong? A major part of the answer, but not all, must be that to deny an infant the whole of a possibly happy, productive, and rewarding life is an enormous deprivation. The mere potential of an infant must confer moral weight on it.

How much weight? And is it the kind of weight that allows an inference to human rights? A potential agent is a being having the power to become an agent, a being whose agency is in a latent or undeveloped state (*in posse*).[4] There is the obvious difficulty that, on this meaning, a foetus, a zygote, an embryo, even a sperm and an egg on course for fertilization would all then be potential agents, and so would all have moral weight and perhaps even a right to life. Many persons find this—I think correctly—close to a *reductio ad absurdum*. Another difficulty is that this definition does not capture our strong intuition that an infant has moral weight simply in virtue of its potential. But our intuition is nothing so simple as that all agents *in posse* have moral weight. That would include too much: a sperm and egg on course for fertilization and the rest. Our intuition is that the potential of an infant, in particular, to become an agent gives *it* moral significance. That intuition is widespread, and often accompanied by the intuition that a late foetus too has moral significance, though perhaps less than an infant but more than an early foetus or embryo. So, if we want to capture these intuitions, we should have to narrow the sort of potential that we see as conferring moral significance. It would be the potential of, particularly, an infant, or a late foetus, or ... (wherever we thought the line should come). The moral significance of potentiality, then, would depend upon not only what it is potential for, but also what it is the potential of. We must therefore be able to tell the moral weight of an infant, a late foetus, an early foetus, and so on. Without these further restrictions, potentiality for agency does not confer anything even approaching a right to life. So, at least, these strong intuitions go, and I shall assume that they are sound.

There are further restrictions needed on potentiality. A new-born baby with a certain serious cerebral deficit will not develop into an agent, let us say, unless stem cells are planted in the brain to remedy the deficit, a treatment not yet available.[5] Would such a baby have anything akin to the moral weight of a person or an agent just because it has the potential to become one? If not, we should add to this 'potentiality' the requirement that it be realizable with resources available at the time needed.

We also have a strong intuition that an infant is morally significant just by being a member of a species, *Homo sapiens*, a characteristic example of which is an agent. It is a belief I share. This acknowledges the moral significance of a severely mentally handicapped infant or an infant with *spina bifida* who will die within a few months of birth. Think of an extreme case of those thereby accorded moral significance: an anencephalic baby. Though

there is good reason to accord anencephalic babies great respect, a reason having to do precisely with our common membership of the species, it is not a reason to accord them, despite their condition, the value attaching to normative agency or even to potential agency on the restricted interpretation just sketched.

The United Nations, in its Convention on the Rights of the Child (1989), stresses not children's potentiality, but their vulnerability.[6] It attributes a raft of rights to children, but the only ones that it makes sense to attribute to infants, as distinct from children, are the well-established right to life and two less well-established, more questionable rights proposed in the Convention: namely, a right to the protection and care necessary for well-being and a right to development 'to the maximum extent possible'.[7] Certainly children's vulnerability imposes substantial obligations on us not imposed by those able to look after themselves. But one must not run together a justification of an obligation and a justification of a right. There are obligations, including highly important ones, that are not correlative to a human right. Also one must be alert to the difference between moral human rights and legal human rights—the first established on moral grounds, the second established on the broader grounds that concern the law, among which morality may well figure. United Nations agencies tend to speak of 'rights', not explicitly of 'human rights', and their aim is to draw up standards of treatment—in this case, of children—that they think all governments should guarantee. There is no gainsaying that aim, despite the doubts one may have about certain of their proposed rights. Do we really want to recognize a duty to ensure a child's development to the maximum extent possible? That would mean developing to the last degree every single potential talent and ability a child has, which seems a thoroughly dubious policy for raising sane and healthy children, let alone a duty we have to a child, let alone a claim that a child can make on us by right.[8]

Their vulnerability of itself does not establish that children are bearers of human rights. Too many things are vulnerable: plants, sperm, foetuses, and so on. And, as with potentiality, we have to have independent understanding of the value of plants, sperm, foetuses, and so on, before we can reach any moral conclusion. If a thing is not itself worth preserving, its vulnerability gives us no reason to protect it. So neither the idea of 'potentiality' nor that of 'vulnerability' on its own helps us much, though, to my mind, the strong intuitions behind them must still, in some form or other, be accommodated by an adequate ethics.