

Agent-Relativity and Deontology

Thomas Nagel

Thomas Nagel, "Agent-Relativity and Deontology," *The View From Nowhere*: New York: Oxford University Press, 1986), pp, 164-85.

1. Three Kinds of Agent-relativity

In this chapter I want to take up some of the problems that must be faced by any defender of the objectivity of ethics who wishes to make sense of the actual complexity of the subject. The treatment will be general and very incomplete. Essentially I shall discuss some examples in order to suggest that the enterprise is not hopeless.

The discussion will revolve around the distinction between agent-relative and agent-neutral values, I won't try to set forth a full ethical theory, even in outline, but I will try to say something in this chapter and the next about the central problem of ethics: how the lives, interests, and welfare of others make claims on us and how these claims, of various forms, are to be reconciled with the aim of living our own lives. My assumption is that the shape of a moral theory depends on the interplay of forces in the psychic economy of complex rational beings, (I shall not say anything about aesthetic values, whose relation to human interests is obscure, though they are revealed to us by the capacity of certain things outside us to command our interest and respect.)

There is one important component of ethics that is consequentialist and impersonal. [S]ome kind of hedonistic, agent-neutral consequentialism describes a significant form of concern that we owe to others, Life is filled with basic pleasures and pains, and they matter. Perhaps other basic human goods, such as health and survival, have the same status, but let me put that aside for the moment. I want now to examine other objective reasons that complicate the picture. Ethics is concerned not only with what should happen, but also independently with what people should or may do. Neutral reasons underlie the former; but relative reasons can affect the latter. In philosophical discussion, the hegemony of neutral reasons and impersonal values is typically challenged by three broad types of reasons that are relative in form, and whose existence seems to be independent of impersonal values.

The first type of reason stems from the desires, projects, commitments, and personal ties of the individual agent, all of which give him reasons to act in the pursuit of ends that are his own. These I shall collect under the general heading of reasons of autonomy (not to be confused with the autonomy of free will).

The second type of reason stems from the claims of other persons not to be maltreated in certain ways. What I have in mind are not neutral reasons for everyone to bring it about that no one is maltreated, but relative reasons for each individual not to maltreat others himself, in his dealings with them (for example by violating their rights, breaking his promises to them, etc.). These I shall collect under the general, ugly, and familiar heading of deontology. Autonomous reasons would limit what we are obliged to do in the service of impersonal values. Deontological reasons would limit what we are permitted to do in the service of either impersonal or autonomous ones.

The final type of reason stems from the special obligations we have toward those to whom we are closely related: parents, children, spouses, siblings, fellow members of a community or even a nation. Most people would acknowledge a noncontractual obligation to show special concern for some of these others - though there would be disagreement about the strength of the reasons and the width of the net. I'll refer to them as reasons of obligation, even though they don't include a great many obligations that are voluntarily undertaken. I mention them here only for completeness and won't discuss them in detail. I have less confidence here than with regard to the other two categories that in ordinary thought they resist agent-neutral justification.

I am not sure whether all these agent-relative reasons actually exist. The autonomous ones and perhaps the obligatory ones are fairly intelligible; but while the idea behind the deontological ones can I think be explained, it is an explanation which throws some doubt on their validity. The only way to find out what limits there are to what we may or must do in the service of impersonal values is to see what sense can be made of the apparent limits, and to accept or reject them according to whether the maximum sense sense is good enough.

Taken together, autonomous, obligatory, neutral, and deontological reasons cover much of the territory of unreflective bourgeois morality. Common sense suggests that each of us should live his own life (autonomy), give special consideration to certain others

(obligation), have some significant concern for the general good (neutral values), and treat the people he deals with decently (deontology). It also suggests that these aims may produce serious inner conflict. Common sense doesn't have the last word in ethics or anywhere else, but it has, as J. L. Austin said about ordinary language, the first word: it should be examined before it is discarded.

Attempts have been made to find room for some version of all three types of apparent exception to impersonal ethics in a more complex impersonal system, using developments of consequentialism like rule-utilitarianism and motive-utilitarianism. A recent example is Hare's two-level version of utilitarianism in *Moral Thinking*. And T. M. Scanlon offers a consequentialist but nonutilitarian justification of deontological rights in "Rights, Goals, and Fairness." I shall not try to show that these reductions of the agent-relative to the agent-neutral fail, since I believe they are partly correct. They just aren't the whole truth. I shall try to present an alternative account of how the exceptions might make sense independently. My aim is to explain what it is that eludes justification in neutral terms. Since this is most conspicuous with regard to autonomy and deontology, I shall concentrate on them. The account in both cases depends on certain discrepancies between what can be valued from an objective standpoint and what can be seen from an objective standpoint to have value from a less objective standpoint.

2. Reasons of Autonomy

Not all the sources of subjective reasons are as simple as sensory pleasure and pain. I believe that the most reasonable objectification of the value that we all recognize in our own encounter with these experiences is an impersonal one. Difficult as it may be to carry out, each of us has reason to give significant weight to the simple sensory pleasure or pain of others as well as to his own. When these values occur in isolation, the results can be demanding. If you and a stranger have both been injured, you have one dose of painkiller, and his pain is much more severe than yours, you should give him the painkiller - not for any complicated reasons, but simply because of the relative severity of the two pains, which provides a neutral reason to prefer the relief of the more severe. The same may be said of other basic elements of human good and ill.

But many values are not like this. Though some human interests (and not only pleasure and pain) give rise to impersonal values, I now want to argue that not all of them do. If I have a bad headache, anyone has a reason to want it to stop. But if I badly want to climb to the top of Mount Kilimanjaro, not everyone has a reason to want me to succeed. I have a reason to try to get to the top, and it may be much stronger than my reason for wanting a headache to go away, but other people have very little reason, if any, to care whether I climb the mountain or not. Or suppose I want to become a pianist. Then I have a reason to practice, but other people have little or no reason to care if I practice or not. Why is this?

Why shouldn't the satisfaction of my desire to climb the mountain have impersonal value comparable to the value it has for me - just like the elimination of my headache? As it happens, you may have to put up with severe altitude headaches and nausea to get to the top of a mountain that high: it has to be worth it to you. Why doesn't the objectification of these values preserve the relation among them that exists in the perspective of the climber? This problem was originally formulated by Scanlon. He makes a strong case against the view that the satisfaction of preferences as such provides the raw material for ethics - the basis of our claims to the concern of others. The impersonal value of things that matter to an individual need not correspond to their personal value to him. "The fact that someone would be willing to forgo a decent diet in order to build a monument to his god does not mean that his claim on others for aid in his project has the same strength as a claim for aid in obtaining enough to eat" (Scanlon (1), pp. 659-60).

There are two ways in which a value may be conditional on a desire: the value may lie either outside or inside the conditional, so to speak. In the former case, a person's having X if he desires X has neutral value: satisfaction of the desire has objective utility that everyone has reason to promote. In the latter case, if a person desires X, his having X has relative value for him: susceptibility to the value is conditional on having the desire, and satisfaction of the desire does not have impersonal utility.

It isn't easy to state a general rule for assigning desires to one category or the other. I have claimed that sensory experiences which we strongly like or dislike simply in themselves have agent-neutral value because of those desires. Such immediate likes and dislikes, not resulting from any choice or underlying reason, are very different from the desires that define our broader aims and ambitions. The former result in mental states that are transparently good or bad, because the attitude of the subject is decisive. The latter require more complicated evaluation.

Most of the things we pursue, if not most of the things we avoid, are optional. Their value to us depends on our individual aims, projects, and concerns, including particular concerns for other people that reflect our relations with them; they acquire value only because of the interest we develop in them and the place this gives them in our lives, rather than evoking interest because of their value. When we look at such desires objectively, from outside, we can acknowledge the validity of the reasons they give for action

without judging that there is a neutral reason for any of those things to be done. That is because when we move to the objective standpoint, we leave behind the perspective from which the values have to be accepted.

The crucial question is how far the authority of each individual runs in determining the objective value of the satisfaction of his own desires and preferences. From the objective standpoint we see a world which contains multiple individual perspectives. Some of the appearances of value from within those perspectives can just be taken over by the objective self. But I believe that others must remain essentially perspectival--appearances of value only to the subject, and valid only from within his life. Their value is not impersonally detachable, because it is too bound up with the idiosyncratic attitudes and aims of the subject, and can't be subsumed under a more universal value of comparable importance, like that of pleasure and pain.

Anyone may of course make the ends of another person his own, but that is a different matter: a matter of personal sympathy rather than of objective acknowledgment. So long as I truly occupy the objective standpoint, I can recognize the value of one of these optional ends only vicariously through the perspective of the person who has chosen it, and not¹ in its own right.

This is true even if the person is myself. When I regard my life from outside, integration of the two standpoints cannot overcome a certain form of detachment. I can't directly appreciate the value of my climbing Mount Kilimanjaro just because I want to, as I appreciate the value of my being adequately fed and clothed. The fact that I want to, viewed from outside, has none of the importance of wanting to, experienced from within. I can see a reason here only through the perspective of TN [Thomas Nagel], who has chosen an optional goal which adds to the values operating within his life something beyond the reasons that simply come at him independently of his choices. I cannot see it except as a value for him, and I cannot therefore take it on without qualification as an impersonal value.

While this seems to me true, there is a natural way to dispute it. I have acknowledged that in the case of sensations, a strong desire or aversion can confer agent-neutral value, and it doesn't require that I have the desire or even fully understand it. Even if, for example, I don't mind the sound of squeaking chalk, I can acknowledge that it is impersonally bad for someone who hates it to be subjected to that sound. The impersonal badness attaches not to the experience conceived merely as a certain sound, but to someone's *having an experience he hates*. The evident awfulness is enough. Now someone might ask, why shouldn't a comparable impersonal value attach to someone's *having (or doing) something he wants* - whatever the desire is? Even if I can't objectively identify with the desire, and therefore can't assign any value to the achievement as such, why can't I judge it to have impersonal value under this more complex description? This would be the universal value under which one could objectively favor all preference-satisfaction.

It isn't easy to make the case convincingly, but I don't believe there is such a universal value. One reason is that the personal projects we are talking about generally involve things happening in the world outside our minds. It seems too much to allow an individual's desires to confer impersonal value on something outside himself, even if he is to some extent involved in it. The impersonal authority of the individual's values diminishes with distance from his inner condition. We can see this clearly, I think, in the limiting case of a personal desire for something which will never impinge on his consciousness: posthumous fame, for example. If someone wants posthumous fame, he may have a reason to do what he thinks will achieve it but one cannot see it as anything but a good *for him*. There is no agent-neutral value whatever in the realization of his hope: the only reason anyone else could have for caring about it would be a specific personal concern for him and his ambitions.

On the other hand, the more a desire has as its object the quality of the subject's experience, and the more immediate and independent of his other values it is, the more it will tend to generate impersonal as well as personal reasons. But to the extent that it transcends his own experience, the achievement of a typical personal project or ambition has no value except from the perspective of its subject -- at least none in any way comparable to the value reasonably placed on it by the person whose ambition it is. (I am assuming here that we can abstract from any intrinsic value the achievement may have which does not depend on his interest at all -- or else that we are dealing with projects whose actual value, whatever it is, derives entirely from the interest of the subject.) Whereas one clearly can find value in the occurrence/nonoccurrence of a sensory experience that is strongly liked/disliked for itself, whether or not one has or even empathizes with the reaction. To put it in a way that sounds paradoxical: the more subjective the object of the desire, the more impersonal the value of its satisfaction.

If this is right, then a certain amount of dissociation is inevitable when we bring the two standpoints together. From within I am directly subject to certain agent-relative reasons. From without all I can do is to acknowledge the reasonableness for the person I am of being motivated by those reasons - without being motivated by them myself, qua objective self. My objectivity shows up in the acknowledgment that these relative reasons are examples of something general, and could arise for any other agent with optional goals of his own. From a point of view outside the perspective of the ambition to climb Kilimanjaro or become a pianist, it

is possible to recognize and understand that perspective and so to acknowledge the reasons that arise inside it; but it is not possible to accept those reasons as one's own, unless one occupies the perspective rather than merely recognizing it.

There is nothing incoherent in wanting to be able to climb Kilimanjaro or play all the Beethoven piano sonatas, while thinking that impersonally it doesn't matter whether one can do this. In fact one would have to be dotty to think it did matter impersonally. It doesn't even matter much impersonally that *if* someone wants to play all the Beethoven sonatas by heart, he should be able to. It matters a little, so that if he is incapable of achieving it, it might be better if he didn't want to --leaving aside whatever value there may be in the ambition itself. The neutral values of pleasure and pain come into effect here. But even that is a rather weak neutral value, since it is not the neutral correlate of the agent-relative reasons deriving directly from the ambition, whose object is not pleasure. If an interest is developed by the agent himself through his choices and actions, then the objective reasons it provides are primarily relative.

Any neutral reasons stemming from it must express values that are independent of the particular perspective and system of preferences of the agent. The general values of pleasure and pain, satisfaction and frustration, fill this role to some extent, as I have said, though only to the extent that they can be detached from the value of the object of desire whose acquisition or loss produces the feeling, (This, incidentally, explains the appeal of hedonism to consequentialists: it reduces all value to the impersonal common denominator of pleasure and pain.) But what there is not, I believe, is a completely general impersonal value of the satisfaction of desires and preferences. The strength of an individual's personal preferences in general determines what they give him reason to do, but it does not determine the impersonal value of his getting what he wants. There is no independent value of preference-satisfaction per se, which preserves its force even from an impersonal standpoint.

3. Personal Values and Impartiality

This may seem harsh, and if we left it at that, it would be. For if agent-neutral reasons derived only from pleasure and pain, we would have no reason to care about many fundamental aspects of other people's welfare which cannot easily be given a hedonistic interpretation -- their freedom, their self-respect, their access to opportunities and resources that enable them to live fulfilling lives.

But I believe there is another way in which these things can be seen as having impersonal value -- without giving carte blanche to individual preferences. These very general human goods share with the much more specific goods of pleasure and freedom from pain a characteristic that generates neutral reasons. Their value does not have to be seen through the particular values of the individual who has or lacks them, or through the particular preferences or projects he has formed.² Also, though they do not involve solely the contents of consciousness, such goods are very "close to home": they determine the character of life from the inside, and this lends authority to the value placed on them by the subject. For both these reasons, when we contemplate our own lives and those of others from outside, the most plausible objectification of these very general goods is not agent-relative.

From the objective standpoint, the fundamental thing leading to the recognition of agent-neutral reasons is a sense that no one is more important than anyone else. The question then is whether we are all equally unimportant or all equally important, and the answer, I think, is somewhere in between. The areas in which we must continue to be concerned about ourselves and others from outside are those whose value comes as close as possible to being universal. If impersonal value is going to be admitted at all it will naturally attach to liberty, general opportunities, and the basic resources of life, as well as to pleasure and the absence of suffering. This is not equivalent to assigning impersonal value to each person's getting whatever he wants.

The hypothesis of two levels of objectification implies that there is not a significant reason for something to happen corresponding to every reason for someone to do something. Each person has reasons stemming from the perspective of his Own life which, though they can be publicly recognized, do not in general provide reasons for others and do not correspond to reasons that the interests of others provide for him. Since the relative reasons are general and not purely subjective, he must acknowledge that the same is true of others with respect to him. A certain objective distance from his own aims is unavoidable; there will be some dissociation of the two standpoints with respect to his individual concerns. The ethical results will depend on the size of the impersonal demands made on him and others by the actual circumstances, and how strongly they weigh against more personal reasons,

One difficult question is whether such a two-tier system implies a significant limit to the degree to which ethics requires us to be impartial between ourselves and others.³ It would imply this if the agent-relative reasons coming from our personal aims were simply added on to the neutral reasons derived from more universal values. For then I would be permitted to pursue my personal projects in preference to the impersonal good of others just as I can pursue those projects in preference to my own health, comfort, etc.; and I wouldn't have to sacrifice myself in return for the furtherance of their personal projects -- only for their

impersonal good. So it looks as though each person's agent-relative reasons would give him a margin of protection against the claims of others - though of course it could be overridden by sufficiently strong impersonal reasons.

However, there is some reason to doubt that the result will be this straightforward. In weighing our agent-relative reasons against the impersonal claims of others, we may not be able to use the same standards we use within our own lives. To take Scanlon's example again: just as we have more reason to help someone get enough to eat than to help him build a monument to his god -- even if he is willing to forgo the food for the monument -- so he may have more reason to help feed others than to build the monument, even if he cannot be faulted for starving himself. In other words, we have to give basic impersonal goods more weight when they come from other people's needs than when they compete with personal reasons within our own lives.

I am not sure of the best account of this, or how far it would go toward requiring impartiality. Full impartiality would seem to demand that any tendency toward self-favoritism on the basis of personal reasons be offset by a corresponding decrease in the weight given in one's interpersonal decisions to impersonal reasons deriving from one's own basic needs -- so that one's total is not increased, so to speak. All reasons would have to be weighted so that everyone was equally important. But I don't know whether a credible system of this kind could be described, at any rate for the purposes of individual decision making. It seems more likely that interpersonal impartiality, both among others and between oneself and others, would have to be defined in terms of agent-neutral values, and that this would leave room for some partiality toward oneself and one's personal concerns and attachments, the extent of it depending on the comparative importance of relative and neutral reasons in the overall system. A stronger form of impartiality, if one is required, would have to appear at a higher level, in the application of practical reason to the social and political institutions that provide a background to individual choice.

There is one objection to this approach which ought to be mentioned, though probably few people would make it. I have claimed that a neutral objectification of the bulk of individualistic subjective reasons does not make sense. But of course that doesn't entail that a relative objectification is correct instead. There is a radical alternative: it could be that these reasons have no objective validity at all, relative or neutral. That is, it might be said by an uncompromising utilitarian that if there isn't a neutral reason for me to climb Kilimanjaro or learn the Beethoven sonatas -- if it wouldn't be a good thing in itself, if the world wouldn't be a better place for my getting to the top of the mountain or being able to play the sonatas -- then I have no reason of any kind to do those things, and I had better get rid of my desire to do them as soon as possible. I may not, in other words, accord more personal value to anything in my life than is justified by its impersonal value.

That is a logically possible move, but not a plausible one. It results from the aim of eliminating perspective from the domain of real value to the greatest possible extent, and that aim is not required of us by objectivity, so far as I can see. We should certainly try to harmonize our lives to some extent with how we think the world should be. But there is no necessity, I now believe, to abandon all values that do not correspond to anything desirable from an impersonal standpoint, even

If there are, objectively, both relative and neutral reasons, this raises a problem about how life is to be organized so that both can be given their due. One way of dealing with the problem is to put much of the responsibility for securing impersonal values into the hands of an impersonal institution like the state. A well designed set of political and social institutions should function as a moral buffer to protect personal life against the ravenous claims of impersonal good, and vice versa. I shall say a bit more about the relation between ethics and political theory later.

Before leaving the subject of autonomy, let me compare what I have said with another recent treatment of the relation between personal and impersonal values in ethical theory: Samuel Scheffler's *The Rejection of Consequentialism*. He proposes an "agent-centred prerogative", which would permit each individual to accord extra weight to all of his interests in deciding what to do, above that which they contribute to the neutral value of the total outcome of his actions, impersonally viewed.

More specifically, I believe that a plausible agent-centred prerogative would allow each agent to assign a certain proportionately greater weight to his own interests than to the interests of other people. It would then allow the agent to promote the non-optimal outcome of his own choosing, provided only that the degree of its inferiority to each of the superior outcomes he could instead promote in no case exceeded, by more than the specified proportion, the degree of sacrifice necessary for him to promote the superior outcome. (p. 20)

This proposal is different from mine but not strictly incompatible with it. Scheffler does not make the distinction I have made between those interests and desires that do and those that do not generate impersonal values. He is not committed to a particular method of ranking the impersonal value of states of affairs, but his discussion suggests that he believes the satisfaction of most types of human preferences could be counted in determining whether one state of affairs or outcome was impersonally better than

another. But whether or not he would accept my distinction, one could accept it and still formulate the proposal of an agent-centered prerogative; for that proposal describes a limit on the requirement always to produce the impersonally best outcome, which is independent of how the comparative impersonal value of outcomes is determined. It might be determined not by all interests but only by some. Then the prerogative would allow an individual to give those interests extra weight if they were his. The trouble is that on the autonomy view I have put forward, he may already have some unopposed reasons which favor himself, arising from those desires whose satisfaction yields personal but not impersonal value. Perhaps it's going too far in moral indulgence to add to these a further prerogative of favoring himself with respect to the fundamental goods and evils whose impersonal value is clear.

An alternative position, which combines aspects of Scheffler's and mine, might be this. The division between interests that give rise to impersonal values and interests that don't is not sharp; it is a matter of degree. Some interests generate only relative reasons and no neutral ones; some generate neutral reasons that are just as strong as the relative ones; but some generate both relative reasons and somewhat weaker neutral ones. An individual is permitted to favor himself with respect to an interest to the degree to which the agent-relative reason generated by that interest exceeds the corresponding agent-neutral reason. There is no uniform prerogative of assigning a single proportionately greater weight to the cure of one's headaches, the realization of one's musical or athletic ambitions, and the happiness of one's children.

A variable prerogative of this kind would accord better than a uniform prerogative with Scheffler's account of the motivation behind it: the wish to give moral significance to the personal point of view by permitting morality to reflect the way in which concerns and commitments are naturally generated from within a particular point of view. If some interests are more dependent on a particular normative point of view than others, they will more naturally resist assimilation to the unifying claims of impersonal value in the construction of morality. All this emerges from the attempt to combine subjective and objective standpoints toward action and its motives.

On the other hand, even after such adjustments there will still be claims of impersonal morality that seem from an individual point of view excessive, and it may be that the response to this will have to include a more general agent-centered prerogative.

4. Deontology

Let me turn now to the obscure topic of deontological constraints. These are agent-relative reasons which depend not on the aims or projects of the agent but on the claims of others. Unlike autonomous reasons, they are not optional. If they exist, they restrict what we may do in the service of either relative or neutral goals.

They complicate an already complicated picture. If there are agent-relative reasons of autonomy that do not give rise to agent-neutral interpersonal claims, then the claims of others must compete with these personal reasons in determining what one should do. Deontological constraints add further agent-relative reasons to the system - reasons not to treat others in certain ways. They are not impersonal claims derived from the interests of others, but personal demands governing one's relations with others.

Whatever their explanation, they are conspicuous among the moral appearances. Here is an example to focus your intuitions.

You have an auto accident one winter night on a lonely road. The other passengers are badly injured, the car is out of commission, and the road is deserted, so you run along it till you find an isolated house. The house turns out to be occupied by an old woman who is looking after her small grandchild, There is no phone, but there is a car in the garage, and you ask desperately to borrow it; and explain the situation, She doesn't believe you, Terrified by your desperation she runs upstairs and locks herself in the bathroom, leaving you alone with the child, You pound ineffectively on the door and search without Success for the car keys. Then it occurs to you that she might be persuaded to tell you where they are if you were to twist the child's arm outside the bathroom door. Should you do it?

It is difficult not to see this as a dilemma, even though the child's getting its arm twisted is a minor evil compared with your friends' not getting to the hospital. The dilemma must be due to a special reason against doing such a thin., Otherwise it would be obvious that you should choose the lesser evil and twist the child's arm. Common moral intuition recognizes several types of deontological reasons --limits on what one may do to people or how one may treat them. There are the special obligations created by promises and agreements; the restrictions against lying and betrayal; the prohibitions against violating various individual rights, rights not to be killed, injured, imprisoned, threatened, tortured, coerced, robbed; the restrictions against imposing certain sacrifices on someone simply as means to an end, and perhaps the special claim of immediacy, which makes distress at a distance so different from distress in the same room, There may also be a deontological requirement of fairness, of evenhandedness or equality in

one's treatment of people, (This is to be distinguished from an impersonal value thought to attach to equality in the distribution of benefits, considered as an aspect of the assessment of states of affairs.)

In all these cases it appears that the special reasons, if they exist, cannot be explained simply in terms of neutral values, because the particular relation of the agent to the outcome is essential. Deontological constraints may be overridden by neutral reasons of sufficient strength, but they are not themselves to be understood as the expression of neutral values of any kind. It is clear from the way such reasons work that they cannot be explained by the hypothesis that the violation of a deontological constraint has high negative impersonal value. Deontological reasons have their full force against your doing something -- not just against its happening.

For example, if there really are such constraints, the following things seem to be true. It seems that you shouldn't break a promise or tell a lie for the sake of some benefit, even though you would not be required to forgo a comparable benefit in order to prevent someone else from breaking a promise or telling a lie. And it seems that you shouldn't twist the arm of a small child to get its grandmother to do something, even something important enough so that you would not be required to forgo a comparable benefit in order to prevent someone else from twisting a child's arm. And it may be that you shouldn't engage in certain kinds of unfair discriminatory treatment (in an official role, for example) even to produce a good result which you would not be required to forgo in order to prevent similar unfairness by others.

Some may simply deny the plausibility of such moral intuitions. Others may say that their plausibility can be subtly accounted for in terms of impersonal values, and that they appear to involve a fundamentally different type of reason for action only if they are inadequately analyzed. As I have said, I don't want to take up these alternative accounts here. They may provide the best hope of rationally justifying something that has the rough shape of a set of deontological restrictions; but offered as complete accounts they seem to me essentially revisionist. Even if from that point of view they contain a good deal of truth, they do not shed light on the independent deontological conceptions they are intended to replace. Those conceptions still have to be understood, even if they will eventually be rejected.

Sometimes, particularly when institutions and general practices are involved in the case, there is a neutral justification for what looks initially like an agent-relative restriction on action. And it is certainly a help to the acceptance of deontological constraints that general adherence to them does not produce disastrous results in the long run. Rules against the direct infliction of harm and against the violation of widely accepted rights have considerable social utility, and if it ceased to be so, those rules would lose much of their moral attractiveness.

But I am convinced that a less indirect, nonstatistical form of evaluation is also at work in support of deontological constraints, and that it underlies the central, most puzzling intuitions in this area. This is what would produce a sense of dilemma if it turned out that general adherence to deontological restrictions worked consistently contrary to impersonal utility. Right or wrong, it is this type of view that I want to explore and understand. There is no point in trying to show in advance that such dilemmas cannot arise.

One reason for the resistance to deontological constraints is that they are formally puzzling, in a way that the other reasons we have discussed are not. We can understand how autonomous agent-relative reasons might derive from the specific projects and concerns of the agent, and we can understand how neutral reasons might derive from the interests of others, giving each of us reason to take them into account. But how can there be relative reasons to respect the claims of others? How can there be a reason not to twist someone's arm which is not equally a reason to prevent his arm from being twisted by someone else?

The relative character of the reason cannot come simply from the character of the interest that is being respected, for that alone would justify only a neutral reason to protect the interest. And the relative reason does not come from an aim or project of the individual agent, for it is not conditional on what the agent wants. Deontological restrictions, if they exist, apply to everyone: they are mandatory and may not be given up like personal ambitions or commitments.

It is hard to understand how there could be such a thing. One would expect that reasons stemming from the interests of others would be neutral and not relative. How can a claim based on the interests of others apply to those who may infringe it directly or intentionally in a way that it does not apply to those whose actions may damage that same interest just as much indirectly? After all, it is no worse for the victim to be killed or injured deliberately than accidentally, or as an unavoidable side-effect of the dangerous rescue operation. In fact the special features of action that bring these reasons into effect may not add to the impersonal badness of the occurrence at all. To use an example of T. M. Scanlon, if you have to choose between saving someone from being murdered and saving someone else from being killed in a similar manner accidentally, and you have no special relation to either of them, it seems that your choice should depend only on which one you're more likely to succeed in saving. Admittedly

the wickedness of a murder is in some sense a bad thing; but when it is a matter of deciding which of them there is more reason to prevent, a murder does not seem to be a significantly worse event, impersonally considered, than an accidental or incidental death. Some entirely different kind of value must be brought in to explain the idea that one should not kill one person even to prevent a number of accidental deaths: murder is not just an evil that everyone has reason to prevent, but an act that everyone has reason to avoid.

In any case, even if a murder were a worse event, impersonally considered, than an accidental death, this could not be used to explain the deontological constraint against murder. For that constraint prohibits murder even if it is necessary to prevent *other murders* -- not only other deaths.

There is no doubt that ideas of this kind form an important part of common moral phenomenology. Yet their paradoxical flavor tempts one to think that the whole thing is a kind of moral illusion resulting either from innate psychological dispositions or from crude but useful moral indoctrination. Before debunking the intuition, however, we ought to have a better grasp of what it is. No doubt it's a good thing for people to have a deep inhibition against torturing children even for very strong reasons, and the same might be said of other deontological constraints. But that does not explain why we find it almost impossible to regard it as a merely useful inhibition. An illusion involves a judgment or a disposition to judge, and not a mere motivational impulse. The phenomenological fact to be accounted for is that we seem to apprehend in each individual case an extremely powerful agent-relative reason not to harm an innocent person. This presents itself as the apprehension of a normative truth, not just as a psychological inhibition. It needs to be analyzed and accounted for, and accepted or rejected according to whether the account gives it an adequate justification.

I believe that the traditional principle of double effect, despite problems of application, provides a rough guide to the extension and character of deontological constraints, and that even after the volumes that have been written on the subject in recent years, this remains the right point of convergence for efforts to capture our intuitions.⁴ The principle says that to violate deontological constraints one must maltreat someone else intentionally. The maltreatment must be something that one does or chooses, either as an end or as a means, rather than something one's actions merely cause or fail to prevent but that one doesn't aim at.

It is also possible to foresee that one's actions will cause or fail to prevent a harm that one does not intend to bring about or permit. In that case it does not come under a deontological constraint, though it may still be objectionable for neutral reasons. The precise way to draw this distinction has been the subject of extensive debate, sometimes involving ingenious examples of a runaway trolley which will kill five people unless you. . . . , where the dots are filled in by different ways of saving the five, all of which in some way involve one other person's death. I won't try to draw the exact boundaries of the principle. Though I say it with trepidation, I believe that for my purposes they don't matter too much, and I suspect they can't be drawn more than roughly: my deontological intuitions, at least, begin to fail above a certain level of complexity. But one point worth mentioning is that the constraints apply to intentionally permitting as well as to intentionally doing harm. Thus in our example there would be the same kind of objection if with the same end in view you permitted someone else to twist the child's arm. You would have let it happen intentionally, and that would be different from a failure to prevent such an occurrence because you were too engaged in doing something else, which was more important.

5. Agents and Victims

So far this is just moral phenomenology: it does not remove the paradox. Why should we consider ourselves far more responsible for what we do (or permit) intentionally than for consequences of action that we foresee and decide to accept but that do not form part of our aims (intermediate or final)? How can the connection of ends and means conduct responsibility so much more effectively than the connection of foresight and avoidability?

It is as if each action produced a unique normative perspective on the world, determined by intention. When I twist the child's arm intentionally I incorporate that evil into what I do: it is my deliberate creation and the reasons stemming from it are magnified and lit up from my point of view. They overshadow reasons stemming from greater evils that are more "faint" from this perspective, because they do not fall within the intensifying beam of my intentions even though they are consequences of what I do.

That is the picture, but can it be correct? Isn't it a normatively distorted picture?

This problem is an instance of the collision between subjective and objective points of view. The issue is whether the special, personal perspective of agency has legitimate significance in determining what people have reason to do - whether, because of this perspective, I can have sufficient reason not to do something which, considered from an external standpoint, it would be better if I did. That is, things will be better, what happens will be better, if I twist the child's arm than if I do not. But I will have done

something worse. If considerations of what I may do, and the correlative claims of my victim against me, can outweigh the substantial impersonal value of what will happen, that can only be because the perspective of the agent has an importance in practical reasoning that resists domination by a conception of the world as a place where good and bad things happen whose value is perspective-free.

I have already claimed that the dominance of this neutral conception of value is not complete. It does not swallow up or overwhelm the relative reasons arising from those individual ambitions, commitments, and attachments that are in some sense chosen. But the admission of what I have called autonomous reasons does not imply the possibility of deontological reasons.⁵ The two are very different. The peculiarity of deontological reasons is that although they are agent-relative, they do not express the subjective autonomy of the agent at all. They are demands, not options. The paradox is that this partial perspectival respect for the interests of others should not give way to an agent-neutral respect free of perspective. The deontological perspective seems primitive, even superstitious, by comparison: merely a stage on the way to full objectivity. How can what we *do* in this narrow sense be so important?

Let me try to say where the strength of the deontological view lies. We may begin by considering a curious feature of deontological reasons on which I have not yet remarked. Intention appears to magnify the importance of evil aims by comparison with evil side-effects in a way that it does not magnify the importance of good aims by comparison with good side-effects. We are supposed to avoid using evil means to produce a good end, even though it would be permissible to produce that good end by neutral means with comparably evil side-effects. On the other hand, given two routes to a legitimate end, one of which involves good means and neutral side-effects and the other of which involves neutral means and equally good side-effects, there is no reason to choose the first route. Deontological reasons tell us only not to aim at evil; they don't tell us to aim at good, as a means. Why should this be? What is the relation between evil and intention, or aiming, that makes them clash with such force?

The answer emerges if we ask ourselves what it is to aim at something, what differentiates it from merely producing the result knowingly.

The difference is that action intentionally aimed at a goal is guided by that goal. Whether the goal is an end in itself or only a means, action aimed at it must follow it and be prepared to adjust its pursuit if deflected by altered circumstances - whereas an act that merely produces an effect does not follow it, is not guided by it, even if the effect is foreseen.

What does this mean? It means that to aim at evil even as a means, is to have one's action guided by evil. One must be prepared to adjust it to insure the production of evil: a falling-off in the level of the desired evil becomes a reason for altering what one does so that the evil is restored and maintained. But the essence of evil is that it should repel us. If something is evil, our actions should be guided, if they are guided by it at all, toward its elimination rather than toward its maintenance. That is what evil *means*. So when we aim at evil we are swimming head-on against the normative current. Our action is guided by the goal at every point in the direction diametrically opposite to that in which the value of that goal points. To put it another way, if we aim at evil we make what we do in the first instance a positive rather than a negative function of it. At every point, the intentional function is simply the normative function reversed, and from the point of view of the agent, this produces an acute sense of moral dislocation.

If you twist the child's arm, your aim is to produce pain. So when the child cries, "Stop, it hurts!" his objection corresponds in perfect diametrical opposition to your intention. What he is pleading as your reason to stop is precisely your reason to go on. If it didn't hurt you would twist harder, or try the other arm. There may be cases (e.g. of justified punishment or obloquy) when pain is not intrinsically evil, but this is not one of them: the victim is innocent. You are pushing directly and essentially against the intrinsic normative force of your goal, for it is the production of his pain that guides you. It seems to me that this is the phenomenological nerve of deontological constraints. What feels peculiarly wrong about doing evil intentionally even that good may come of it is the headlong striving against value that is internal to one's aim.

I have discussed a simple case, but naturally there can be complications. One is the possibility of someone volunteering to be subjected to some kind of pain or damage, either for his own good or for some other end which is important to him. In that case the particular evil that you aim at is swallowed up in the larger aim for deontological purposes. So the evil at which we are constrained not to aim is *our victim's* evil, rather than just a particular bad thing, and each individual has considerable authority in defining what will count as harming him for the purpose of this restriction.⁶

All this still leaves unsettled the question of justification. For it will be objected that if one aims at evil as a means only, then even if several people's interests are involved one's action is really being guided not by evil but by overall good, which includes a balance of goods and evils. So when you twist the child's arm, you are guided by the aim of rescuing your injured friends, and the good of

that aim dominates the evil of the child's pain. The immediacy of the fact that you must try to produce evil as a subsidiary aim is phenomenologically important, but why should it be morally important? Even though it adds to the personal cost to you, why should it result in a prohibition?

I don't believe there is a decisive answer here. The question is whether to disregard the resistance encountered by my immediate pursuit of what is evil for my victim, in favor of the overall value of the results of what I do. When I view my act from outside and think of it as resulting from a choice of the impersonally considered state of the world in which it occurs, this seems rational. In thinking of the matter this way, I abstract my will and its choices from my person, as it were, and even from my actions, and decide directly among states of the world, as if I were taking a multiple choice test. If the choice is determined by what on balance is impersonally best, then I am guided by good and not by evil.

But the self that is so guided is the objective self, which regards the world impersonally, as a place containing TN and his actions, among other things. It is detached from the perspective of TN, for it views the world from nowhere within it. It chooses, and TN, its instrument, or perhaps one could say its agent, carries out the instructions as best he can. *He* may have to aim at evil, for the impersonally best alternative may involve the production of good ends by evil means. But he is only following orders.

To see the matter in this light is to see both the appeal of agent-neutral, consequentialist ethics and the contrary force of agent-relative, deontological ethics. The detached, objective view takes in everything and provides a standpoint of choice from which all choosers can agree about what should happen. But each of us is not only an objective self but a particular person with a particular perspective; we act in the world from that perspective, and not only from the point of view of a detached will, selecting and rejecting world-states. So our choices are not merely choices of states of the world, but of actions. Every choice is two choices, and from the internal point of view, the pursuit of evil in twisting the child's arm looms large. The production of pain is the immediate aim, and the fact that from an external perspective you are choosing a balance of good over evil does not cover up the fact that this is the intrinsic character of your action.

I have concentrated on the point of view of the agent, as seems suitable in the investigation of an agent-relative constraint. But there is also something to be said about the point of view of the victim. There too we encounter problems having to do with the integration of the two standpoints, and further support for the analysis. Moral principles don't simply tell agents what they may and may not do. They also tell victims what sort of treatment they may and may not object to, resist, or demand. If I were justified in killing one innocent person to save five others, then he would have no right to object, and on a fully consequentialist view he would have no right to resist. The other five, by contrast, would have the right to object if I didn't kill him to save them. A thoroughly impersonal morality would require that victims as well as actors be dominated by impersonal, agent-neutral values in their judgments about how others treat them.

But this seems an excessive demand to make of individuals whose perspective on the world is inherently complex and includes a strong subjective component. Of course none of the six people in this dilemma wants to die, but only one of them is faced with me trying to kill him. This person is not permitted, on a purely agent-neutral consequentialist view, to appeal for his life against my deliberate attempt to take it from him. His special position as my victim doesn't give him any special standing to appeal to me.

Of course the deontological position has a parallel feature. On a deontological view, the five people I could save by killing the one cannot appeal to me for their lives, against my refusal to save them. (They may appeal against their killers, if that's the nature of the death threat, but not against me.) But this does not make the two positions symmetrical, for there is a difference. The deontological constraint permits a victim always to object to those who aim at his harm, and this relation has the same special character of normative magnification when seen from the personal perspective of the victim that it has when seen from the personal perspective of the agent. Such a constraint expresses the direct appeal to the point of view of the agent from the point of view of the person on whom he is acting. It operates through that relation. The victim feels outrage when he is deliberately harmed even for the greater good of others, not simply because of the quantity of the harm but because of the assault on his value of having my actions guided by his evil. What I do is immediately directed against his good: it doesn't just in fact harm him.

The five people I could save by killing him can't say the same, if I refrain. They can appeal only to my objective acknowledgment of the impersonal value of their lives. That is not trivial, of course, but it still seems less pressing than the protest available to my victim - a protest he can make not to them but to me, as the possessor of the life I am aiming to destroy.

This merely corroborates the importance of the internal perspective in accounting for the content of deontological intuitions. It does not prove the correctness of those intuitions. But it confirms that a purely impersonal morality requires the general suppression of the personal perspective in moral motivation, not only in its rejection of relative reasons of autonomy but also in its refusal to

accept agent-relative deontological restrictions. Such restrictions need not be absolute: they can be thought of as relative reasons with a certain weight, that are among the sources of morality but do not exhaust it. When we regard human relations objectively, it does not seem irrational to admit such reasons at the basic level into the perspective of both agents and victims.

Notes

1. [On pp. 152-3 of *The View from Nowhere*, Nagel explains the distinction between agent-neutral and agent-relative reasons as follows: 'If a reason can be given a general form which does not include an essential reference to the person who has it, it is an agent-neutral reason. For example, if it is a reason for anyone to do or want something that it would reduce the amount of wretchedness in the world, then that is a neutral reason. If on the other hand the general form of a reason does include an essential reference to the person who has it, it is an agent-relative reason. For example, if it is a reason for anyone to do or want something that it would be in his interest, then that is a relative reason.' -Ed.]
2. This is the rationale behind the choice of primary goods as the common measure of welfare for distributive justice in Rawls (1). See Rawls (2) for a much fuller treatment. That essay, Scanlon, and the present discussion are all treatments of the "deep problem" described in Rawls (1), pp. 173-5. Dworkin's defense of resources rather than welfare as the correct measure of equality is also in part a response to this problem.
3. Impartiality should not be confused with equality. Nothing I say here bears on the question of how much equality is required in the allocation of what has impersonal value. Absolute impartiality is consistent with a denial that equality should be an independent factor at all in settling distributive questions.
4. A good statement of a view of this type is found in Fried.
5. This is emphasized by Scheffler, who has a cautiously skeptical discussion of deontological constraints under the heading of "agent-centred restrictions."
6. The same seems to apply even when informed consent is impossible, as when we cause suffering or damage to a young child for its own greater good -- though here there may be a residual inhibition: if we imagine in the case described that the child's safety depends on getting the car keys, it doesn't altogether remove the revulsion against twisting his arm to get them.

References

- Fried, Charles. *Right and Wrong*. Cambridge, MA: Harvard University Press, 1978.
- Rawls, John. (1) *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.
(2) "Social Unity and Primary Goods," in A. Sen and B. Williams (eds.) *Utilitarianism and Beyond*. Cambridge University Press, 1982.
- Scanlon, T.M. "Preference and Urgency," *Journal of Philosophy*, 1975.
- Scheffler, S. *The Rejection of Consequentialism* (Oxford: Oxford University Press, 1982).