



## Rights and Agency

Amartya Sen

*Philosophy and Public Affairs*, Vol. 11, No. 1 (Winter, 1982), 3-39.

Stable URL:

<http://links.jstor.org/sici?sici=0048-3915%28198224%2911%3A1%3C3%3ARAA%3E2.0.CO%3B2-X>

*Philosophy and Public Affairs* is currently published by Princeton University Press.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/pup.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

This paper is about three distinct but interrelated problems: (1) the role of rights in moral theory, (2) the characterization of agent relative values and their admissibility in consequence-based evaluation, and (3) the nature of moral evaluation of states of affairs.

First, it is argued in Sections I through III that both welfarist consequentialism (such as utilitarianism) and constraint-based deontology are fundamentally inadequate because of their failure to deal with certain important types of interdependences present in moral problems. This leads to an alternative approach, called goal rights systems, which incorporates, among other things, some types of rights in the evaluation of states of affairs, and which gives these rights influence on the choice of actions through the evaluation of consequent states of affairs. The formulation of goal rights in the form of rights to certain 'capabilities' is explored and contrasted with other approaches in Section IV.

Goal rights systems require consequential analysis, though they may or may not be fully consequentialist. Recently, consequentialism has been sharply criticized by Williams, Nozick, Nagel, and others for, among other things, its failure to accommodate agent relative values of the kind suggested by deontological reasons as well as reasons of autonomy and integrity. Different types of agent relativities (in particular, doer relativity, viewer relativity and self-evaluation relativity) are distinguished in this paper, and their interrelations are analyzed. Different kinds of agent relative values are shown in

Section V to involve different combinations of these agent relativity types.

Consequentialism is then shown in Section VI to accommodate some agent relative values, but not others. But even those agent relative values that are inconsistent with consequentialism are accommodated in consequence-based evaluation, admitting evaluator relativity in the evaluation of consequences themselves. Thus, goal rights, including capability rights, and other goals can be combined with deontological values (in the nonconstraint form), along with other agent relative considerations, in an integrated system.

The possibility of evaluator relativity in the evaluation of states of affairs raises a complex problem of second-order moral theory. Examining this problem using various interpretations of moral judgments of states, I present and explore the possibility of a "positional interpretation" of such judgments in Section VII. In the last Section I make some concluding remarks on positive and negative freedoms.

#### I. WELFARIST CONSEQUENTIALISM AND CONSTRAINT-BASED DEONTOLOGY

In the instrumental view rights are not valuable in themselves, but right-based rules, conventions, institutions, etc., are useful in pursuing other-right-independent-goals. The most commonly identified goals in the instrumental approach tend to be 'welfarist' goals,<sup>1</sup> with the goodness of states of affairs being judged entirely by the personal utility features of the respective states. One special case of welfarist evaluation is by far more common than others, and that is the case of utilitarian evaluation in which the goodness of a state of affairs is judged simply by the sum total of personal utilities in that state.<sup>2</sup> But other welfarist approaches exist, for example, judging states by the utility level of the worst-off individual in that state, as under a

1. For a discussion of the distinguishing features of welfarism, and a critique, see my "Utilitarianism and Welfarism," *Journal of Philosophy* 76 (September 1979).

2. When the population is a variable, one has to make the further distinction between 'classical' and 'average' utilitarianism. In this paper I shall not go into that issue.

variant of Rawls's "Difference Principle,"<sup>3</sup> or by some other method of distribution-sensitive aggregation of personal utilities.<sup>4</sup>

In contrast, in the constraint-based deontological view rights are treated as constraints on actions. These constraints must not be violated *even if* such violation would lead to better states of affairs. Violating rights is simply wrong. Unlike in the instrumental view, rights *are* given intrinsic importance, but unlike in "goal rights systems," to be presented later in this paper, rights *directly* affect judgments of actions—and only of actions—rather than being embedded first in the evaluation of states of affairs and then affecting the evaluation of actions through consequential links between actions and states. As Robert Nozick puts it, "Individuals have rights, and there are things no person or group may do to them (without violating their rights)."<sup>5</sup> Further, "Rights do not determine a social ordering but instead set the constraints within which a social choice is to be made, by excluding certain alternatives, fixing others, and so on."<sup>6</sup>

I shall now argue that both the welfarist instrumental approach (including, *inter alia*, the traditional utilitarian approach) and the deontological constraint-based approach are inadequate in important ways. Furthermore their respective inadequacies are related to a common ground shared by the two, despite sharp differences in other respects. The particular common ground is the *denial* that

3. John Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971). Rawls himself repudiates this *variant*, but it is nevertheless much used. See E.S. Phelps, ed., *Economic Justice* (Harmondsworth: Penguin Books, 1973); P.J. Hammond, "Equity, Arrow's Conditions and Rawls's Difference Principle," *Econometrica* 42 (1976); and C. d'Aspremont and L. Gevers, "Equity and Informational Basis of Collective Choice," *Review of Economic Studies* 46 (1977).

4. See my *On Economic Inequality* (Oxford: Basil Blackwell, 1973). Kevin Roberts has provided an axiomatization of the class of distribution-sensitive, utility-based social welfare functionals, satisfying symmetry, homotheticity, and additive separability; see his "Interpersonal Comparability and Social Choice Theory," *Review of Economic Studies* 47 (1980). See also C. Blackorby and D. Donaldson, "Utility vs. Equity: Some Plausible Quasi-orderings," *Journal of Public Economics* 7 (1977).

5. R. Nozick, *Anarchy, State and Utopia* (Oxford: Basil Blackwell, 1974), p. ix.

6. Nozick, *Anarchy, State and Utopia*, p. 166.

realization and failure of rights should enter into the evaluation of states of affairs themselves and could be used for consequential analysis of actions. Nozick's view that "rights do not determine a social ordering" is shared fully by welfarists in general and utilitarians in particular. Their ways part there, however, with the welfarist instrumentalist viewing rights in terms of their consequences for *right-independent* goals and the constraint-based deontologist reflecting rights *without consequential justification* as constraints on actions. State-evaluation independent of rights leaves a gap that cannot be adequately closed by either of these approaches.

Consider the constraint-based deontological approach first. It is of course obvious that this approach can hardly do justice to those rights associated with the so-called positive freedom.<sup>7</sup> But the problem is not confined to that, and is present even when the intrinsic value of positive freedom is disputed and the focus is chosen to be entirely on noncoercion and related issues of negative freedom. Even with negative freedom, *multilateral* interdependences can arise and undermine the rationale of the constraint-based deontological approach. The only way of stopping the violation of a very important liberty of one person by another may be for a third to violate some other, less important liberty of a fourth. To take a crude example, the only way of saving A from rape by B could be for C to arrive speedily at the spot in a car stolen from D, who is not a party to the rape but who does not want his car to be used for this purpose. The justification of C's action will require consequential analysis trading off the badness of violating D's right to the disposal of his own car against the badness of letting the rape occur. Since the constraint-based deontological view does not permit violation or realization of rights "to determine the social ordering,"<sup>8</sup> it is particularly inadequate in deal-

7. For the classic statement of the distinction between 'positive' and 'negative' freedom, see Isaiah Berlin, "Two Concepts of Liberty," in his *Four Essays on Liberty* (Oxford: Clarendon Press, 1969).

8. Nozick, p. 166. It is, however, possible to respond to problems of this kind by making the so-called 'constraints' nonconstraining under particular circumstances, though there is obviously a danger here of resorting to *ad hoc* solutions. (For an illuminating critique of some possible reasons for overriding right-based constraints, see Judith Jarvis Thomson, *Self-defense and Rights: The Lindley Lecture*, University of Kansas, 1976.) For example, it can be specified that if the badness of the state of affairs resulting from obeying the constraint exceeds

ing with such cases of *multilateral* interdependences, which can be easily accommodated however in a system of consequential evaluation.

The instrumental welfarist approach is well-armed with a consequential framework of moral evaluation. But since the evaluation of consequences is based ultimately on utility information only (non-utility information being valued just as a causal influence on—or as a surrogate for—utility information<sup>9</sup>), mental features (such as pleasures, happiness, desires, etc., depending on the particular interpretation of utility) rule the roost entirely. The losses of the victim and other sufferers are contrasted with the gains of the violators and other gainers entirely in terms of relative utility features. There might have been good utilitarian reasons for forcing men to fight wild animals in the Colosseum with the utility gain of the thousands of spectators outweighing the utility loss of the few forced men.

I shall return to these general issues later. In the next section, I illustrate the difficulties created by the two traditional approaches by taking up a particular example of a moral problem.

## II. AN ILLUSTRATIVE MORAL PROBLEM

Ali is a successful shopkeeper, who has quickly built up a good business in London since immigrating from East Africa. He is, however, hated by a small group of local racists, and a particular gang of them—I shall call them bashers—are, it happens, planning to bash Ali that evening in some secluded spot to which Ali will go alone. Donna, a

---

some 'threshold,' then the constraint may be overridden. Such a threshold-based 'constraint' system must rest ultimately on consequential analysis, comparing one set of consequences (badness resulting from obeying the constraint) with another (badness of violating the constraint itself, given by the threshold), and its distinguishing feature will be the particular *form* of the consequence-evaluation function. Compromises of this kind raise other problems, which I do not pursue here, but I should emphasize that I do not include such consequential analysis in the category of 'constraint-based deontological approach,' against which my criticism here is directed.

9. See my "On Weights and Measures: Informational Constraints in Social Welfare Analysis," *Econometrica* 43 (1977), and "Personal Utilities and Public Judgments: Or What's Wrong with Welfare Economics?" *Economic Journal* 89 (1979).

West Indian friend of Ali, has just come to know of the bashers' plan, and wants to warn Ali about it. But Ali has gone away for the day, and will go to that secluded spot without returning home. Donna does not know where Ali has gone nor the location of the planned bashing, but she does know that Ali has left a message on the desk of his business contact Charles about his movement. However, Charles is away for the day also, and cannot be contacted. Hence the only way of getting Ali's message is by breaking into Charles's room. Donna asked for the help of the police, who dismissed Donna's story as a piece of paranoiac fantasy. Donna knows that she can certainly frustrate the planned bashing by breaking into Charles's room, recovering the message, and warning Ali during the day. But she cannot do this without violating the privacy of Charles, who is, Donna also knows, a secretive man who will feel rather embarrassed at the thought of someone looking through his personal papers to find the message. Indeed, Donna also knows Charles, a self-centered egoist, well enough to be sure that he will be more disturbed by the violation of his own privacy than by the bashing of Ali. What should Donna do?

The long-term utilities of Ali and the ten people in the gang of bashers are given in table 1.

TABLE 1  
LONG-TERM UTILITIES

State of Affairs	Ali's Utility	Each Basher's Utility (10 Bashers)	Utility Sum Total	Minimum Utility	Inequality of Utilities
No bashing	15	5	65	5	more
Bashing	10	6	70	6	less

Notice that Ali will suffer a good deal more than any of the bashers will gain in utility terms, but the aggregate utility gain of the bashers

exceeds the utility loss of Ali. In terms of these utilities, however, Ali remains better off than the gang of poor, unemployed bashers, even though his suffering is large enough to make a substantial impact on his long-term utility total.<sup>10</sup>

Donna considers the utility information in viewing the problem first from the welfarist angle. In terms of the utilitarian objective, it is clear that the bashing up is doing more good than harm. There could be indirect effects, of course, but Donna finds that they won't be very serious in this case. There is so much fear of racial violence in that locality anyway, that one incident will not *add* significantly to the general sense of insecurity. Also, whether or not the bashing is prevented by warning Ali, the bashers will continue to go about their business as usual, and there will certainly not be any better chance of making the police take some action if Ali is *not* actually bashed up. Thus, there is nothing in these considerations to weaken the utility argument for the incident. Of course, there is Charles's utility also, but that will strengthen the case for no action by Donna, since breaking into his room to stop the bashing will reduce his utility.

Donna turns next to the welfarist version of Rawlsian difference principle. She finds that the utility level of the worst-off individuals will go up rather than down as a consequence of the bashing up. So this "maximin" view also favors doing nothing to stop the bashing up. Indeed, so will every welfarist criterion that responds positively to a *larger utility total, more equally distributed*. So Donna moves to indirect utilitarian (and more generally, indirect welfarist) reasoning. She can well believe that among the class of 'uniform' rules dealing with bashing up, the rule of not treating anyone thus, in any situation whatever, may receive much support from the point of view of utilitarian evaluation of outcomes. But clearly from the same point

10. Cardinal interpersonal comparability of utility is assumed in the table. None of the arguments will change if the numbers are all altered by applying some positive linear (affine) transformation, for example, multiplying each number by 10, or adding 100 to each. On the framework for measurability and comparability assumptions, see my *Collective Choice and Social Welfare* (San Francisco: Holden-Day, 1970; distribution taken over by North Holland, Amsterdam); and L. Gevers, "On Interpersonal Comparability and Social Welfare Orderings," *Econometrica* 45 (1979). Also K. Basu, *Revealed Preference of Governments* (Cambridge: Cambridge University Press, 1979).



of view that is, at most, a second best if choices are not necessarily confined to such uniform rules. Better still will be compliance with the no-bashing rule *except* in cases like this, in view of the net utility gain from this particular incident. Why should utilitarianism settle for such a second best by arbitrarily restricting choices to the class of rigid rules only?<sup>11</sup> Will it not be better from the utilitarian point of view to have a more flexible rule that permits bashing up in cases of the type described, thereby avoiding unnecessary sacrifice of utility? The justification of any policy—be it a rule, or an act, or something else—must rest ultimately on the ability to produce the best outcomes, judged by “outcomes utilitarianism.”<sup>12</sup>

However, it is possible that following such flexible rules is not feasible, and this is quite possibly a case when we should deal with ‘disposition’ as a variable. Henry Sidgwick<sup>13</sup> had seen in this an argument for going against act utilitarian reasoning, and recently this aspect of the problem has been thoroughly investigated from different perspectives in the works of Richard Hare, Robert Adams, John Harsanyi, and others.<sup>14</sup> Even if a particular act of bashing, or raping, or torturing improves the utility picture, given other things, this does not imply a utilitarian endorsement of that act if that act must go with a certain disposition that will typically cause harm.<sup>15</sup> The eschewal of that act will then be a necessary part of the suppression of that bad disposition.

Donna ponders over this indirect utilitarian reasoning, and becomes convinced that if she were to advise *the bashers* on what to do, from the utilitarian point of view, she would indeed argue for the

11. Cf. David Lyons, “Utility and Rights,” mimeographed (Cornell University, 1979); forthcoming in *Nomos* 24 (1982).

12. Cf. my “Utilitarianism and Welfarism,” pp. 464-67.

13. Henry Sidgwick, *The Method of Ethics*, 7th ed. (London: Macmillan, 1907).

14. See especially R. M. Hare, “Ethical Theory and Utilitarianism,” in *Contemporary British Philosophy*, ed. H. D. Lewis (London: Allen & Unwin, 1976); R. M. Adams, “Motive Utilitarianism,” *Journal of Philosophy* 73 (August 1976); J. Harsanyi, “Rule Utilitarianism and Decision Theory,” in *Decision Theory and Social Ethics*, ed. H. W. Gottinger and W. Leinfellner (Dordrecht: Reidel, 1978).

15. For a discussion of some of the difficulties with this general approach, see the ‘Introduction’ to *Utilitarianism and Beyond*, ed. A. Sen and B. Williams (Cambridge: Cambridge University Press, forthcoming).

removal of the disposition to bash up innocent people (including Ali). But Donna also recognizes that advising the bashers on what to do is not the exercise in which she is currently engaged, and her actions, whatever they are, are most unlikely to have any significant effect on the disposition of the bashers. Her moral problem concerns the issue of whether to break into Charles's room to collect the information that will permit her to warn Ali. There is no direct utilitarian case for her to break in, and it is not clear how bringing in the choice of dispositions is going to provide an argument for her to break in. Of course, if a disposition "to break into other people's rooms" were found to be a good disposition to cultivate, this would give her, in terms of disposition utilitarianism, a reason to break in. But she can hardly believe that it is likely that such a general disposition to break in, or even a disposition to break in for a perceived excellent cause, will be a good one to cultivate in terms of utilitarian evaluation of consequences. Clearly, what is needed in this particular case and in cases like this is a *discriminating* defense of breaking in that balances pros and cons, rather than a general disposition to break down the door. And in this case such a calculating defense of the act of breaking in is yet to be found within the utilitarian (and more generally, welfarist) approach.

Of course, a strong argument for breaking into Charles's room could have emerged if the violation of Ali's bodily integrity were given a force strong enough for it not to be outweighed by the countervailing utility advantage of the bashers. But the utilitarian and other welfarist methods of outcome evaluation do not permit this, as they insist on judging the strength of claims exclusively in terms of utility information only.

Despite this failure of welfarism (including utilitarianism) to give Donna a good ground for doing what her moral conviction tells her she should do, to wit, break into Charles's room and save Ali, she decides that she must stick by her conviction. How can a person's bodily integrity, his freedom to move about without harm, be outweighed by mere pleasure or desire-fulfilment of the bashers? By not stopping the bashers, she would rob Ali of one of his most elementary rights. With this thought in mind, Donna decides to turn now to constraint-based deontological approaches. And yes, she sees that there is indeed an in-

flexible "side constraint," in Nozick's terms, which is morally imposed on the bashers not to bash up Ali. However, this constraint does not affect Donna directly since she is not one of the bashers! There is nothing in that constraint-based deontological perspective that would require Donna to do anything at all.

The more Donna thinks about it, however, the more she feels convinced that she must really break into Charles's room and save Ali from bodily injury. Maybe she is not *required* to do anything, but surely she is *free* to? But, no, she isn't free to break into Charles's room since that deontological perspective also imposes a side constraint against the violation of *Charles's* rights. Since right violations and realizations do not enter the evaluation of states of affairs ("do not," as Nozick puts it, "determine a social ordering") and the violation of Ali's more important right cannot be used for consequential justification of infringing Charles's less important right, Donna's hands are tied. Indeed, Nozick repudiates such trade-offs (what he calls "utilitarianism of rights"<sup>16</sup>), and the constraint-based deontological approach, free from consequential analysis, offers nothing else.

To summarize the position, at the risk of some oversimplification, Donna can have a good case for breaking into Charles's room to save Ali if she can use a consequential analysis with nonwelfarist evaluation of consequences. Constraint-based deontology does not permit the former (namely, consequential analysis), while welfarist instrumentalism does not permit the latter (namely, nonwelfarist evaluation of consequences). It appears that to make room for her deeply held and resilient conviction that she must save Ali by breaking into Charles's room, Donna must reject both these traditional approaches and look for a new approach that is at once consequentialist and non-welfarist.

### III. GENERAL PROBLEMS AND SPECIAL EXAMPLES

The example illustrating the moral issues in the last section might be considered to be rather a special one, and so indeed it is. Does this

16. Nozick, *Anarchy*, pp. 28-29. On this question see Herbert Hart, "Between Utility and Rights," *Columbia Law Review* 79 (1979): 828-46.

undermine the whole line of my criticism of the two traditional approaches? I shall argue that it does not.

First, while the chosen example illustrates the problem, the nature of it could be anticipated on the basis of the general arguments presented in Section I. The failure to introduce fulfilment and non-realization of rights in the evaluation of states of affairs produces a lacuna that can scarcely be made good either by inflexible deontological constraints used without consequential evaluation or by trying to catch the importance of rights in terms of the metric of utilities.<sup>17</sup> It is not so much that we stumbled on some peculiar intuition that we did not expect and then proceeded to argue against the two traditional approaches; the example merely served the purpose of *illustrating* a difficulty diagnosed in general terms.

The idea of incorporating realization of rights in the evaluation of states of affairs themselves might also appear to be an *ad hoc* device to take care of problems of the kind illustrated by the chosen example. Further, it may be asked whether statements about states of affairs in terms of rights realization really make sense.

In fact, as it happens, statements about states of affairs involving such incorporation are perfectly natural in other contexts as well. Consider the remark that the period of Emergency Rule in India was a bad one because so many rights of so many people were violated. This is a statement about a state of affairs, and it does include the violation of rights directly in the evaluation of the state. Statements like this can be used in a consequential analysis of actions, and this will translate the importance of fulfilment and violation of rights from the evaluation of states of affairs to that of actions.

Second, I don't think the example is in any sense peculiar, but even if it were, this would not imply that the intuition from it should be

17. However, it is possible to go some distance in filling this lacuna by the approach of discounting, ignoring, or even devaluing 'bad' pleasures or desires, for example, those related to sadism (see Harsanyi, "Rule Utilitarianism and Decision Theory," p. 8). Although I shall not examine this alternative approach here, I believe it to be inadequate fundamentally; for example, the contrast between sadomasochistic fantasies (involving private reading) and cruel actions (involving right violations) may involve similar types of pleasures, but require quite different moral perspectives. Further, violation of rights need not necessarily involve 'bad' pleasures or desires.

dismissed. Indeed, in many of the common cases, intuitions based on quite different principles tend to run in the same direction, so that it is impossible to be sure of the basis of an overall judgment. The problem is comparable to that in statistical inference when different explanatory variables corresponding to alternative models move together (and are 'multicollinear'), making discrimination impossible. In order to do the discrimination, we choose examples such that different principles (for example, utilitarianism and the approach of incorporating rights in the evaluation of states of affairs) push us in different directions, even if such examples are somewhat less common than the ones in which either approach will recommend the same answer. It is methodologically wrong, therefore, to ignore the relevance of our intuitions regarding rather unusual examples which are brought for this reason into moral arguments. Hare's dismissal of what he sees as the "commonest trick"<sup>18</sup> and even Rawls's decision to postpone discussing "hard cases"<sup>19</sup> are on reflection rather less justified than they might first appear.

Third, although the example involving Ali, bashers, Charles, and Donna is a rather elaborate one, it had to be finely tuned to be usable *simultaneously* against welfarist instrumentalism and constraint-based deontology. It is, in fact, much easier to find different examples in which each of the two traditional approaches fails respectively to account for our deeply held moral convictions.

All this does not, of course, dispute the need to test one's intuition by seeking what John Rawls has called a reflective equilibrium.<sup>20</sup> Nor to dispute the importance of assessing moral beliefs at the 'critical' level, seeking principles that "would be arrived at by leisured moral thought in completely adequate knowledge of facts."<sup>21</sup> But there is nothing particularly illegitimate in illustrating difficulties with specific

18. See Hare, "Ethical Theory and Utilitarianism," p. 123. See also K. Ng, "Welfarism: A Defence against Sen's Attack," *Economic Journal* 91 (1981).

19. See John Rawls, "A Kantian Concept of Equality," *Cambridge Review*, February 1975, p. 96.

20. Rawls, *A Theory of Justice*, p. 20. Nor to subjecting them to what Bruce Ackerman calls 'dialogic test' (*Social Justice in the Liberal State* [New Haven: Yale University Press, 1980], p. 353).

21. Hare, "Ethical Theory and Utilitarianism," p. 123.

moral approaches by referring to resilient moral intuitions dealing with comparatively unusual cases, when these cases provide scope for discrimination the more 'usual' cases lack.<sup>22</sup>

#### IV. GOAL RIGHTS SYSTEMS AND CAPABILITIES

A moral system in which fulfilment and nonrealization of rights are included among the goals, incorporated in the evaluation of states of affairs, and then applied to the choice of actions through consequential links will be called a goal rights system. A few preliminary warnings may be called for to avoid misunderstanding.

First, the goal rights systems form a wide *class* rather than represent some unique moral position. There are many sources of difference, for example, what rights to include among the goals, in what form they are to be included, what nonright values (if any) are to be admitted, what weights to use, how should choice of actions be related to the evaluation of outcomes.

Second, although rights are included within the evaluation of states of affairs, there could be other things to which the evaluation of states of affairs is sensitive in a goal rights system. The crucial issue is the inclusion of fulfilment and nonfulfilment of rights—rather than the exclusion of nonright considerations—in the evaluation of states of affairs.

Third, while *sensitivity* of action-evaluation to consequences is essential for a rights-consequence system, such a system need not be fully consequentialist. This question is further pursued in Section VI.

Fourth, it may be worth emphasizing that although a goal rights system incorporates a goal-included view of rights, it does not reject the instrumental relevance of rights either.<sup>23</sup> Indeed, the violation of a right, say, freedom of speech, might be seen as making the outcome worse *both* because of the violation of that right itself and because of

22. See also my "Utilitarianism and Welfarism," pp. 475-79, and "Equality of What?" in *The Tanner Lectures on Human Values*, vol. 1, ed. S. McMurrin (Cambridge: Cambridge University Press, 1980).

23. See Thomas Scanlon, "Rights, Goals and Fairness," in *Public and Private Morality*, ed. Stuart Hampshire (Cambridge: Cambridge University Press, 1978) on the essentiality of the instrumental role of rights. See also Partha Dasgupta, "Decentralization and Rights," *Economica* 47 (1980).

the negative effect it has on other objectives, including quite possibly other right-based goals, for example, making it more difficult for the economically deprived to make claims on the state for corrective action.<sup>24</sup> Furthermore, even in a goal rights system there can be rights that are justified *entirely* in instrumental terms, for example the ticket checker's right to see your ticket during a train journey. The important point is that a goal rights system permits the inclusion of right-based considerations in the goals themselves (and thus permits its direct use in the evaluation of outcomes and consequences), but it does not deny the use of instrumental considerations as well.

I turn now to the question of how to incorporate goal rights in the evaluation of states of affairs. It is not unusual to think of rights as a relation between two parties *i* and *j*, for example, person *i* having the claim on *j* that he will do some particular thing for *i*.<sup>25</sup> There is, however, some advantage in characterizing goal rights as a relation not primarily between two parties but between one person and some 'capability' to which he has a right, for example, the capability of person *i* to move about without harm.<sup>26</sup> This rather blurs the distinction between rights that relate to so-called positive freedoms and those related to negative freedoms such as liberty and noncoercion. If all goal rights take the form of rights to certain capabilities, then a goal rights system may be conveniently called a capability rights system.

The capability perspective has some distinctive advantages. An important advantage over the two-person (or two-agent) relationship can be illustrated even in terms of the type of problem encountered

24. The suppression of the freedom of speech in India during the Emergency Rule was widely criticized—in my judgment correctly—on *both* types of grounds.

25. See, for example, W. N. Hohfeld, *Fundamental Legal Conceptions* (New Haven: Yale University Press, 1923); S. Kanger "New Foundations for Ethical Theory" in *Deontic Logic: Introductory and Systematic Readings*, ed. R. Helpinen (Dordrecht: Reidel, 1971); and the helpful critical account given by L. Lindahl, *Position and Change* (Dordrecht: Reidel, 1977). See also J. Raz, *The Concept of a Legal System* (Oxford: Clarendon Press, 1970; 2nd edition, 1980).

26. Cf. my "Equality of What?" pp. 215-20. The focus on capabilities also differs fundamentally from that on utilities in being concerned with (i) achievements rather than their mental assessments, and (ii) the opportunity sets and not only the chosen position. See also John Mackie, "Can There Be a Right-based Moral Theory," *Midwest Studies in Philosophy* 3 (1978), and John Bennett, "Goods, Needs, and Social Theory," mimeographed (Cornell University, 1978).

in the earlier example involving Donna's dilemma. Consider two different variations of that case. Denoting the original case as case A, consider variant B, in which the so-called bashers are not in fact planning to bash up Ali, but to injure him by planting a small explosive device on the spot where Ali is going to be that evening. After planting that bomb, the so-called bashers have actually disappeared—in fact emigrated from “rotting” England—and they will never really find out what happened; it is agreed that the bashers' utilities are independent of the success of the plot.<sup>27</sup> Donna can, as before, save Ali by warning him about this danger, but given the garbled message she has received (without any locations specified), she can only do this by breaking into Charles's room and finding that note from Ali about his own movements. The consequences on Ali and Charles are the same as in the original version of the problem.

In variant C, there are no bashers. But there is, in fact, an explosive device surviving from the last war at the spot where Ali is planning to be that evening. Donna has come to know of this without being able to decipher from the message where that spot is. Again, Donna can save Ali in exactly the same way as before, that is, by breaking into Charles's room and finding out about Ali's current whereabouts and warning him. The utility consequences on Ali and Charles are just as in variants A and B.

Now the question. If it is all right for Donna to break into Charles's room in variant B, is it all right for her to break into Charles's room in variant C? If we look at the violation of the right of Ali vis-à-vis the bashers, the two cases are quite different. In variant B, Ali's right of 'immunity' against the bashers is violated by them, whereas in variant C no such violation is involved. If, however, one takes the view—as seems plausible—that the bad outcomes Donna may be trying to prevent are equally bad in the two cases, then that view seems to undermine the consequential relevance of the particular perspective of *right violation by someone else*. It is the harm to Ali that one is con-

27. This need not be so in some variant of the desire fulfilment version of utility without an 'awareness requirement,' on which see Jonathan Glover, *Causing Death and Saving Lives* (Harmondsworth: Penguin, 1977), p. 63. See also James Griffin, "The Concept of Utility," mimeographed (Keble College, n.d.).



cerned with in the two variants, and that harm is the same. If the important issue is Ali's safety, his 'general right' to the capability of moving about without harm, then variants B and C must be taken to be exactly comparable.

It might, however, appear that in asserting this close moral similarity of Donna's decision problems in variants B and C we have overproved our case. While, on the one hand, that argument does push us in at least one respect further from the deontological constraint-based perspective (the formulation of rights is not only not in terms of constraints, it may not even be related to *violation* by other persons), but, on the other hand, it might seem to push us closer to the utilitarian (or, more generally, welfarist consequentialist) position. In terms of utility consequences, variants B and C are not distinguishable, and by treating them as involving closely similar moral problems for Donna, we might appear to be back in the welfarist ballpark. Indeed, in the particular contrast between variants B and C there is not a lot to choose between the capability framework and the welfarist framework. However, contrasting variants B and C is not the only thing of interest. In dealing with case A, welfarism seemed particularly inadequate, and that inadequacy remains despite the congruence of the general rights position and the welfarist positions in contrasting B and C.

In fact, the issues can be brought out more sharply by considering, in addition, a further contrast. Corresponding to case A, we now have a case D in which the utility picture is just as in case A, but the non-utility features involving rights are quite different. In particular, the so-called bashers aren't bashers really; in fact they work in Ali's shop. (Ali calls them bashers since he is liable to complain that they handle the merchandise clumsily.) The bashers suspect that they have been given a more dismal picture about the firm's financial position than is justified by the accounts, and they want to examine the accounts without Ali's permission in Ali's absence. Donna has come to know of this, and is considering warning Ali about this likely occurrence, which will have, Donna knows, had consequences for Ali's business plans. But she can warn Ali only by breaking into Charles's room, and the utilities are all as in case A (with "bashing up" being replaced by

“examining Ali’s business accounts without permission”). If it is all right for Donna to break into Charles’s room to warn Ali in case A, is it all right for her to do the same in case D? Since *all* the utility consequences are taken to be exactly the same, all welfarist moralities (including the utilitarian one) must treat cases A and D in exactly similar ways. On the other hand, the general right not to be bashed up is hardly matched in terms of seriousness by the right of not having one’s financial accounts examined and one’s business affected, even when the utility consequences are just the same. The ability to retain bodily safety is quite a different type of right from the ability to keep one’s financial accounts private. Donna could very well decide to break into Charles’s room in case A but not in case D if she is to be guided by some capability rights systems of the type being explored in this paper.

The constraint-based deontological view will, of course, treat cases A and D as similar, since *Charles’s* right must not be violated in either case. On the other hand, a consequentialist system with goal rights being specified in the traditional two-party form may well discriminate between A and D since the ‘immunity’ right of Ali against the bashers may be given much greater weight in case A (when bodily injury is involved) than in case D (when only privacy of accounts is at issue).

The contrasting positions of the four approaches are shown in table 2. To simplify matters, the two-person goal rights systems and capability rights systems are considered in their “exclusive” form, with no goals other than fulfilment of the respective rights and with full consequentialism.

## V. AGENT RELATIVITY

I turn now to the issue that Derek Parfit<sup>28</sup> and Thomas Nagel<sup>29</sup> have called agent relativity of values. “Agent relative values” are “specified by reference to the agent for whom they provide reasons.”<sup>30</sup> The ques-

28. Derek Parfit, “Prudence, Morality, and the Prisoner’s Dilemma,” *Proceedings of the British Academy for 1979* (London: Oxford University Press, 1981).

29. T. Nagel, “The Limits of Objectivity,” in *The Tanner Lectures on Human Values*, vol. I, ed. S. McMurrin (Cambridge: Cambridge University Press, 1980).

30. Nagel, “The Limits of Objectivity,” p. 119; also pp. 101-4, and 120-21.

TABLE 2  
MORAL CORRESPONDENCES

Can Donna Justifiably Intervene?	Direct Utilitarian Systems	Constraint-based Deontological Systems	Pure Two-person Goal Rights Systems	Pure Capability Rights Systems
Case A	No	No	Yes	Yes
Case B	Yes	No	Yes	Yes
Case C	Yes	No	No	Yes
Case D	No	No	Depends on specification	Depends on specification
Same answer in cases A and D?	Yes	Yes	No	No
Same answer in cases B and C?	Yes	Yes	No	Yes

tion of agent relativity is an important one in moral philosophy,<sup>31</sup> and it is of obvious relevance to the contrast between characterizing rights and duties in terms of deontological constraints and doing that through a consequential system. Indeed, Nagel has argued that "cer-

31. See also Bernard Williams, "A Critique of Utilitarianism," in *Utilitarianism: For and Against*, ed. J.J.C. Smart and B. Williams (Cambridge: Cambridge University Press, 1973) and his "Persons, Character and Morality," in *The Identities of Persons*, ed. A. Rorty (Berkeley: University of California Press, 1976); Jonathan Glover, *Causing Death and Saving Lives* (Harmondsworth: Penguin Books, 1977); John Mackie, *Ethics: Inventing Right and Wrong* (Harmondsworth: Penguin Books, 1977); Stuart Hampshire, ed., *Public and Private Morality* (Cambridge: Cambridge University Press, 1978); Charles Fried, *Right and Wrong* (Cambridge, MA: Harvard University Press, 1978); Nancy Davis, "Utilitarianism and Responsibility," *Ratio* 22 (June 1980).

tain ethical positions, those sometimes called consequentialist, admit only agent-neutral values."<sup>32</sup>

I would like to examine how agent relativity might fit into a goal rights system, but before getting on to that question, two preliminary issues have to be sorted out. First, what is agent relativity? In this Section I shall argue that agent relativity can be of very different types, which—for many purposes—require a discriminating analysis rather than being lumped together to be dealt with similarly. Second, what are the restrictions on admitting agent relative values that are imposed by consequence-based evaluation in general and consequentialism in particular? That question is taken up in Section VI.

An agent *does* things and also *views* actions and outcomes. Agent neutrality can be characterized in terms of the switching of the “doer” role or of the “viewer” role or of both together. The corresponding invariance requirements are quite different though, obviously, not independent. These yield three different conditions of neutrality.

The analysis presented here concentrates on a very simple case, which is not only easy to investigate but is also adequate for bringing out the distinctions and for categorizing the agent relative values examined by Nagel, Parfit, Williams, and others. There are two persons sharing the same, possibly agent relative morality, and at most one of them can do some act. But it can be either, *provided* he is not stopped by the other. The consequences—other than the identity of the agent doing the act—depend only on whether or not the action is performed, and not on who does the act or whether either has been stopped by the other.<sup>33</sup>

*Doer neutrality (DN)*: Person *i* may do this act if and only if person *i* may permit person *j* to do this act.

*Viewer neutrality (VN)*: Person *i* may do this act if and only if person *j* may permit person *i* to do this act.

*Self-evaluation neutrality (SN)*: Person *i* may do this act if and only if person *j* may do this act.

32. Nagel, “The Limits of Objectivity,” p. 119.

33. These are exacting requirements, but tend to be typically implicit in discussions on agent relativity.

Denoting  $A_i(i)$  as the assertion that person  $i$  may do the act himself, and denoting  $A_i(j)$  as the assertion that person  $i$  may permit  $j$  to do the act, we get the following entailment relations, with the appropriate universal quantifiers (for all  $i$  and  $j$ , for all acts and the corresponding consequences with the characteristics specified).

*Doer neutrality*  $A_i(i) \iff A_i(j)$ .

*Viewer neutrality*  $A_i(i) \iff A_j(i)$ .

*Self-evaluation neutrality*  $A_i(i) \iff A_j(j)$ .

How do these neutralities relate to each other? Two results are straightforward:

- (1) *DN*, *VN* and *SN* are bilaterally independent of each other (that is, none of them alone entails any of the others);
- (2) any two of *DN*, *VN*, and *SN* together entail the third.<sup>34</sup>

*Doer relativity* (*DR*), *viewer relativity* (*VR*) and *self-evaluation relativity* (*SR*) are defined as negations of *DN*, *VN*, and *SN* respectively. Taking 'may permit' as 'has no obligation to prevent', these conditions can be stated thus:

*Doer relativity* (*DR*): It is not the case that  $i$  may do this act if and only if he has no obligation to prevent  $j$  from doing this act.

*Viewer relativity* (*VR*): It is not the case that  $i$  may do this act if and only if  $j$  has no obligation to prevent  $i$  from doing this act.

*Self-evaluation relativity* (*SR*): It is not the case that  $i$  may do this act if and only if  $j$  may do this act.

In view of (1) and (2) above:

- (1\*) two of *DR*, *VR*, and *SR* can hold with or without the third;
- (2\*) exactly one of *DR*, *VR*, and *SR* cannot hold.

Hence, if any one type of relativity is satisfied, at least one other type of relativity will also obtain.

I turn now to the agent relative values discussed in the literature. Thomas Nagel identifies "two broad types of reasons that appear to be agent-relative in form":

34. To check (2), the following will do. From *SN* and *DN*:  $A_i(i) \iff A_j(j) \iff A_j(i)$ ; thus *VN*. From *SN* and *VN*:  $A_i(i) \iff A_j(j) \iff A_i(j)$ ; thus *DN*. Finally, from *DN* and *VN*:  $A_i(i) \iff A_i(j) \iff A_j(j)$ ; thus *SN*.

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*.

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 agent-neutral reasons for everyone to want it to be *the case* that no one is maltreated, but agent-relative reasons for each individual not to *maltreat others himself*, in his dealings with them (e.g., by violating their rights, breaking his promises to them, etc.). These I shall collect under the general, ugly, and familiar heading of deontology.<sup>35</sup>

I start with deontology. The agent relativity of this value is brought out nicely by Nagel's example, the analysis of which from the deontological point of view leads to the following conclusion:

. . . it seems that you shouldn't twist the arm of a small child to get its grandmother to do something, even if the thing is quite important—important enough so that it would not be reasonable to forego a comparable benefit in order to prevent someone else from twisting a child's arm.<sup>36</sup>

Notice that this involves only one *viewer* (namely, *you*) and two *doers* (namely, *you* and *someone else*). Thus directly this invokes *doer relativity* (DR). For reasons of consistency—see (2\*)—it must also lead to at least one other type of relativity. The reasoning makes it clear that the obvious candidate for this is *viewer relativity* (VR). Although that “someone else”—if he or she shared that deontological morality—could not think it okay to twist a child's arm (despite the stated benefits), he or she need not accept the obligation to stop another person (say, *you*) from doing the arm twisting (with similar other consequences).

Self-evaluation relativity is *not*, however, a part of this “deontological view,” and SN can indeed hold. It is *not* okay for you to do the

35. Nagel, “The Limits of Objectivity,” pp. 119-20.

36. *Ibid.*, p. 127.

twisting of the child's arm, and it is also *not* okay for someone else to do that.

Although the deontological argument is presented by Nagel in the form of requiring a *constraint* (to wit: "shouldn't twist the arm of a small child"), this is, of course, much stronger than negating doer neutrality, and his arguments really amount to a defense of doer relativity (negating *DN*) rather than a justification for doing this through the extremism of a constraint.<sup>37</sup> I have already argued against a constraint-based deontological morality in Sections I through III, but that was not an argument against doer relativity. In fact, deontological morality in the weak form of denying doer neutrality—and correspondingly viewer neutrality—can be accommodated *within* a goal rights system. The point will be further pursued in Sections VI and VII.

I turn now to what Nagel calls autonomy. This is a very broad category and includes rather different types of agent relative arguments, based on "desires, projects, commitments and personal ties." One type of argument relates to Bernard Williams's analysis of 'integrity' applied to cases like the following:

George, who has taken his Ph.D. in chemistry, finds it extremely difficult to get a job. . . . An older chemist who knows about the situation, says that he can get George a decently paid job in a certain laboratory, which pursues research into chemical and biological warfare. George says that he cannot accept this, since he is opposed to chemical and biological warfare. The older man replies that he is not too keen on it himself, come to that, but after all George's refusal is not going to make the job or the laboratory go away; what is more, he happens to know that if George refuses to take the job, it will certainly go to a contemporary of George's who is not inhibited by any such scruples and is likely if appointed

37. Similarly, Bernard Williams's analysis of the deontological issue in the form of Jim's dilemma ("A Critique of Utilitarianism," pp. 98-118) does not lead to treating deontological grounds as sufficient for imposing a constraint that *must be met*, but to treating them as important in moral calculation. Indeed, Williams comes to the conclusion, after balancing the pros and cons in Jim's decision problem, that overriding the deontological consideration is "probably right in this case" (p. 90).

to push along the research with greater zeal than George would. . . .  
What should he do?<sup>38</sup>

Williams argues that for George to be persuaded by the older chemist's argument "is to neglect the extent to which *his* actions and *his* decisions have to be seen as the actions and decisions which flow from the projects and attitudes with which he is most closely identified," and it "is, thus, in the most literal sense, an attack on his integrity" (pp. 116-17).

I am not concerned here with assessing the force of this example, or of Williams's arguments.<sup>39</sup> My concern here is solely with the types of agent relativity that will follow from Williams's arguments if they are accepted. This connection might look more immediate than it, in fact, is. It is, of course, true that although Williams's arguments will tend to support George's judgment that it will be wrong of him to take the job, they will not suggest a similar case—based on integrity—for his contemporary to refuse that employment. But this does not imply that self-evaluation relativity is involved in this example, since George and his contemporary do *not* share the same morality. Indeed, the consequences of George's taking the job can be seen to be quite different from those resulting from his contemporary's taking the job, if the violation of George's moral attitude is directly counted in, abandoning the rather limited system of consequence-accounting on which utilitarianism happens to rely. While this feature of the example can be modified by making George and his contemporary share the same morality, this will alter the nature of the example altogether, possibly affecting the moral judgments related to it.

On the other hand, there is nothing especially problematic about making the older chemist—call him Harry—share George's morality (he is not, we know, "too keen" on chemical and biological warfare anyway). With this modification, and further altering the story in

38. Williams, "A Critique of Utilitarianism," pp. 97-98.

39. An important issue is raised by Jonathan Glover, "It Makes No Difference Whether Or Not I Do It," *Proceedings of the Aristotelian Society*, Supplement (1975); on this see also Bernard Williams, "Utilitarianism and Moral Self-Indulgence," in *Contemporary British Philosophy*, ed. H. D. Lewis (London: Allen & Unwin, 1976).



such a way that the other consequences of Harry taking the job are identical with those of George taking it, we can now ask:

- (1) *Viewer relativity*: Is it not the case that George may take the job if and only if Harry has no obligation to stop George from taking it?
- (2) *Doer relativity*: Is it not the case that Harry may take the job if and only if he has no obligation to stop George from taking it?

The answers to these questions will depend on how integrity is treated.

I shall distinguish between "integrity respect," a general respect for integrity no matter whose, and "integrity responsibility," a personal responsibility for one's own integrity. With general integrity respect featuring in the shared morality of George and Harry, there is no need to require viewer relativity or doer relativity. No matter who takes the job and who views it, the same unfavorable picture of involvement in chemical and biological warfare, by someone who believes such involvement to be wrong, comes through. On the other hand, with integrity responsibility viewer relativity *is* involved: George must not take the job even though Harry has no obligation to stop him. And so is doer relativity: Harry must not take the job even though he has no obligation to stop George from taking it ("It's *your* decision, George").<sup>40</sup> Notice that in neither case is self-evaluation relativity involved, since George and Harry sharing the same morality and circumstances must evaluate their own respective acts of taking the job in the same way.

A somewhat similar interpretational dichotomy can be found in the other cases covered under the broad heading of "autonomy." Take the case of "personal ties."<sup>41</sup> It may be the case that parents should be

40. The concern with one's own integrity does not imply that the person must take integrity as a virtue on its own rights. As Bernard Williams argues, integrity is not "a disposition which itself yields motivations"; "it is rather that one who displays integrity acts from those dispositions and motives which are most deeply his" ("Utilitarianism and Moral Self-indulgence," p. 316).

41. See Parfit, "Prudence, Morality, and the Prisoner's Dilemma." Parfit links up the discussion with issues of 'co-ordination,' on which see also Hans Herzberger, "Coordination Theory," in *Foundations and Applications of Decision Theory*, vol. I, ed. C. Hooker, J. Leach, and E. McClennen (Dordrecht: Reidel, 1978); and Donald Regan, *Utilitarianism and Cooperation* (Oxford: Clarendon

specially concerned with their own children's welfare, and this will immediately entail self-evaluation relativity. It may be okay for me to do some act differentially benefiting my children over yours without it being okay for you to do this (benefiting my children over yours), given your responsibility to your own children. We know from the analytical result presented earlier (see (2\*)) that self-evaluation relativity must go with *at least one* of doer relativity and viewer relativity, and indeed it can be either.

Consider the distinction between (1) people having different aims related to their own kith and kin, for example, parents aiming differentially to further benefits to their own children ("tie aims," for short), and (2) everyone having the same aims but valuing acts directly responsive to ties, for example, everyone valuing the acts of parents helping their respective children ("tie respect," for short). In the former case there can be viewer relativity: *i* may do something favoring his own children vis-à-vis *j*'s children, but *j*—sharing the same tie-aimed morality—must try to stop *i* from doing this. This can go with doer *neutrality*, since it need not matter *how* the benefits to one's children are brought about, whether by one's own actions or by those of others.

In contrast, what is important in the case of tie respect is the belief that parents themselves should respond by taking actions benefitting their own children, and it is not so good if the same benefits were brought about by helpful actions of strangers. This implies doer relativity, and it can go with viewer neutrality, since in this case no one need specially value benefits going to one's own children. It can also involve self-evaluation relativity: Ed may do something helping his own children at the cost of Bill's children, while Bill may not do that thing helping Ed's children (not because Bill values benefits to Ed's children differently from the way Ed does, but both value the action of someone helping his own children *himself*).

It is, of course, possible to have a "mixed" case, that of "aimed tie respect," in which people have different aims favoring, say, their own children, but also value that the benefits to their own chil-

---

Press, 1980). The characteristic of being 'collectively self-defeating' is associated with viewer relativity without necessarily involving doer relativity.

dren are brought about by their own actions.<sup>42</sup> In this case *VR*, *DR*, and *SR* may all be involved. Ed may do something favoring his own children, while it could be the case that (1) Bill must try to stop Ed (viewer relativity), (2) Ed must try to stop Bill from doing the same thing helping Ed's children (doer relativity), and (3) Bill may refuse to do the same thing helping Ed's children (self-evaluation relativity).

This discussion has not, of course, begun to exhaust different types of agent relative reasons, but even with the few types considered here, quite a rich variety of contrasts have been found. Before moving in the next Section to the question of how these different types of agent relativity relate to the discipline of consequence-based evaluation, I present a summary picture of the different agent relative reasons in terms of *DR*, *VR*, and *SR* (see table 3).

TABLE 3  
MORAL REASONS AND AGENT-RELATIVITY TYPES

	Deontology	Tie Aims	Tie Respect	Aimed Tie Respect	Integrity Respect	Integrity Responsibility
Doer Relativity	Yes	No	Yes	Yes	No	Yes
Viewer Relativity	Yes	Yes	No	Yes	No	Yes
Self-evaluation Relativity	No	Yes	Yes	Yes	No	No

## VI. CONSEQUENCE-BASED EVALUATION AND CONSEQUENTIALISM

An outcome morality is a morality dealing with judging states of affairs. Consequentialism requires the hegemony of outcome morality over moral judgments concerning other variables as well, for ex-

42. Cf. Parfit, "Prudence, Morality, and the Prisoner's Dilemma," p. 562.

ample, actions. There are various strategic issues involved in the use of control variables to achieve the best consequences taking everything into account, and it is possible to show that no consequentialist strategy can be fully successful in getting the best possible outcomes if it concentrates only on a "single influence," whether it be acts, rules, motives.<sup>43</sup> However, in this paper I shall abstract from these issues, and concentrate on act consequentialism only, in situations in which act-based calculations are adequate for serving overall *consequentialist* goals.

Whether consequentialism can accommodate agent relative values will depend on the way outcome morality is characterized. In especially narrow formulations, consequences are defined *excluding* the actions that bring them about, and in this case agent relativity of the "doer relative" kind would be impossible to accommodate. But this is an arbitrary exclusion, and it is not imposed in broader characterizations of consequentialism.<sup>44</sup>

Another important issue concerns the nature of the outcome evaluation function. Within a *given* moral approach, must every person have the same outcome evaluation function  $G(x)$  irrespective of differences in their positions vis-à-vis actions, beneficiaries, and the like? That is indeed the standard assumption in the consequentialist literature, but it is worth inquiring whether that feature is essential or accidental, for example owing to the historical association of consequentialism with utilitarianism, which certainly does impose an evaluator-neutral perspective on outcome evaluation. There is, it seems to me, no compelling reason why a morality that is sensitive to the differences in the position of people vis-à-vis states (including actions that bring the states about) should not permit—indeed require—that different people evaluate the same state different-

43. On the relationship between outcome morality and various types of consequentialism (depending on the 'control' variables), see my "Utilitarianism and Welfarism," pp. 464-68. The internal difficulties of act consequentialism are brought out by the classic paper of Allan Gibbard, "Rule Utilitarianism: A Merely Illusory Alternative?" *Australasian Journal of Philosophy* 43 (1965). See also Thomas Schelling, *The Strategy of Conflict* (Cambridge, MA: Harvard University Press, 1960); Herzberger, "Coordination Theory"; John Harsanyi, "Rule Utilitarianism and Decision Theory," *Erkenntnis* 11 (1977); Regan, *Utilitarianism and Co-operation*; and Parfit, "Prudence, Morality, and the Prisoner's Dilemma."

44. See, for example, Williams, "A Critique of Utilitarianism," pp. 86-88.

ly. A morality that insists that after killing his wife Desdemona Othello must regard that *state of affairs* to be morally exactly as good or as bad as others would—and no worse than that—would seem to miss something about the nature of moral evaluation of states.<sup>45</sup> I would, therefore, not rule out evaluator-relative forms of outcome morality, with  $G_i(x)$  being the moral value of state  $x$  from the point of view of person  $i$ . This raises some serious interpretational problems, given the traditional wisdom on the nature of outcome morality,<sup>46</sup> which I shall consider in Section VII.

The term consequentialism is, however, strongly associated with a characterization of morality that would not permit evaluator relativity in the evaluation of outcomes, and must insist on  $G_i(x) = G_j(x)$  for all  $i$  and  $j$ , for all  $x$ . When that restriction is relaxed but otherwise the hegemony of outcome morality is maintained, I shall call the approach “consequence-based evaluation.” Consequentialism, then, is the special case of consequence-based evaluation in which the outcome morality is evaluator neutral. It requires not only that each person should judge his control variables (such as acts) ultimately in terms of their effects on the goodness of outcomes, but also that each person must judge the goodness of any given outcome in exactly the same way.

Consequentialism, thus characterized, cannot at all accommodate *viewer relativity* of action judgments, and thus cannot reflect “deontology,” “tie aims,” “aimed tie respect,” and “integrity responsibility” (see table 3). However, viewer relativity can be admitted within consequentialism by dropping the assumption stated earlier (and which is typically implicit in the treatment of consequentialism in the literature on agent relative values) that “the consequences—other than the

45. One aspect, among others, of the issue relates to what Bernard Williams describes as “moral philosophy’s habit . . . of treating person in abstraction from character” (“Persons, Character and Morality,” in *The Identities of Persons*, ed. A. Rorty [Berkeley: University of California Press, 1976], p. 215).

46. There are also problems of linguistic convention since ‘morally good’ is used in the form of a one-place predicate (for example, “a morally good state”), not a two-place one (for example, “a morally good state from the point of view of person  $i$ ”). The case for two-place formulations is discussed in Section VII.

identity of the agent doing the act—depend only on whether or not the action is performed, and not on . . . whether either has been stopped by the other.”<sup>47</sup> While this distinction is important in some cases and is worth introducing anyway, it provides a rather limited way of accommodating viewer relativity of action judgments (in the absence of evaluator relativity of outcome judgments).

However, “tie respect” and “integrity respect” do not at all require viewer relativity of action judgments. Are these types of moral reasons consistent with consequentialism? There is no difficulty in answering in the affirmative in the case of integrity respect, which—as we have already seen (in table 3)—does not even involve any of the specified types of agent relativity. Although that is not the case with tie respect, which involves doer relativity and self-evaluation relativity, consistency with consequentialism does, in fact, hold.

Consider state  $x$  as involving some benefit to  $i$ 's children and not to  $j$ 's children, brought about by  $i$ 's action, and state  $y$  as the same benefits to  $i$ 's children and not to  $j$ 's children brought about by  $j$ 's action. A tie-respecting evaluation function  $G(\cdot)$  will value these states differently, with  $G(x) > G(y)$ . A consequentialist morality can, consistently with this, make it okay for  $i$  to take this action, but not for  $j$  (self-evaluation relativity). Further, it can make it not right for  $j$  to take that action but fine for him to let  $i$  take it (doer relativity). Thus, agent relativity associated with tie respect is consistent with consequentialism.

Consequentialism cannot, however, make much room for agent relative action judgments associated with deontology, tie aims, aimed tie respect, and integrity responsibility. However, with evaluator-relative outcome evaluation function  $G_A(\cdot)$ , consequence-based evaluation can cover agent relativity of these types. To understand how this would work, it may be useful to consider the outcome evaluation function

47. As was noted earlier in footnote 33 these are “exacting requirements.” There are often good reasons to distinguish between a state in which person  $i$  did not do act  $A$  because he did not *try* to, and a state in which person  $i$  did not do act  $A$  because he was *stopped* by person  $j$ . Even with evaluator-relative outcome judgments, which—as will be presently discussed—gives much scope for viewer relativity of action judgments, the distinction between “not trying” and “being stopped” will retain its relevance.

$G_i(\cdot)$  in the 'separable' form, related to the evaluation of the goodness of the actions  $a$  and the evaluation of the goodness of the "rest of the state of affairs"  $b$  brought about by those actions:  $G_i(x) = V(z_i(a), y_i(b))$ , when  $x = (a, b)$ .<sup>48</sup>

When we are exclusively concerned with relativities based on *deontological* values, it is natural to think of  $y_i(b)$  as evaluator-neutral, but the part dealing with action evaluation  $z_i(a)$  has to be evaluator relative, depending on whether or not the evaluator is himself the agent. This will lead to doer relativity and viewer relativity. Self-evaluation relativity will, however, be avoided since  $z_i$  will have the same value when the evaluator—no matter who—is *himself* the relevant agent.

With integrity *responsibility* of the kind illustrated by the case of George and Harry, in which the action and the agency are crucial,  $z_i$  will be clearly evaluator relative, but  $y_i$  can be, consistently with that illustration, evaluator neutral. This will lead to doer relativity and viewer relativity of action judgments in much the same way as in the case of deontological values.

In the pure case of *tie aims* illustrated by the situation in which the benefits to one's own children have to be specially valued but not the agency of the action that brings about these benefits,  $z_i$  will be evaluator neutral whereas  $y_i$  will be evaluator relative. Doer neutrality of action judgments will follow from the irrelevance of the agency of the relevant actions ("I *may* do this to help my children if and only if I *may* permit you to do this to help my children"), but viewer relativity and self-evaluation relativity will be produced by the evaluator-relative character of  $y_i$ . In the case of *aimed tie respect*,  $z_i$  and  $y_i$  are both evaluator relative, and all the three types of agent relativity can be induced by consequence-based evaluation.

48. Separability is an exacting requirement and many functions do not have separable representations. For the issues involved in separability, see the helpful paper of W. M. Gorman, "Tricks with Utility Functions," in *Essays in Economic Analysis*, ed. M. J. Artis and A. R. Nobay (Cambridge: Cambridge University Press, 1976). It is worth emphasizing that the ability of evaluator-relative, consequence-based evaluation to accommodate these different types of agent relativities does *not* depend on whether or not  $G_i$  is separable. Separability is assumed here for expositional purpose only to make the exercise transparent. Also note that an additional "degree of freedom" can be introduced in the evaluator-relative form  $G_i(x)$  by permitting an *i*-variant  $V$  function, that is:  $V_i(z_i(a), y_i(b))$ .

## VII. EVALUATOR RELATIVITY AND THE POSITIONAL INTERPRETATION

The possibility of using consequence-based evaluation combined with an evaluator-relative outcome morality was entertained in the last Section, distinguishing that package from consequentialism. But the notion of an evaluator-relative outcome morality may appear to be deeply problematic, and it certainly requires scrutiny.

Evaluator relativity of a given outcome morality must not be confused with interpersonal differences of moralities. Indeed insofar as a particular person  $k$  entertains an evaluator-relative morality, we can represent that morality in the form  $G_i^k(x)$ , reflecting the moral value that in the opinion of person  $k$  (parameter) should be appropriately attached to state  $x$  (variable) by person  $i$  (variable). This could, of course, include the moral value that  $k$  himself should attach to state  $x$ :  $G_k^k(x)$ .<sup>49</sup> Since I am not concerned currently with contrasting different outcome moralities but with interpreting any given one, the superscript  $k$  is eschewed, and an outcome morality is represented in the two-variable form  $G_i(x)$ . This could stand for the moral beliefs of a particular person  $k$  or a possible moral theory (to be compared with, say, the utilitarian outcome morality).

It is easy to entertain the suspicion that an evaluator-relative outcome morality must be an internally contradictory one: how can  $i$  and  $j$  morally value a given state  $x$  differently *within* one consistent moral theory? The inadmissibility of evaluator relativity within one moral theory can be based on two propositions:

- (1) interpersonal differences of moral valuations of the same states must indicate contradictory moral beliefs;
- (2) no moral theory should endorse the holding of contradictory moral beliefs.

Evaluator relativity can be defended by denying either (1) or (2), or—of course—both.

49. Charles Taylor has pointed out to me the importance of clarifying that while  $G_i^k(x)$  is called the 'outcome morality' of  $k$ , it is doing two quite distinct jobs: (i) the specification of  $k$ 's own moral evaluation of states, and (ii) the specification of  $k$ 's views of *appropriate correspondences* between different people's moral evaluations of states.



Before disputing (1), which is my primary target, I should briefly comment on the possibility of justifying evaluator relativity via denying (2). There could be good *instrumental* reasons for a moral theory to require inconsistent moral beliefs. For example, in "two-level" moral theories, which have been extensively explored by Hare,<sup>50</sup> instrumental use of an evaluator-relative morality (for example, parents valuing especially highly the benefits to their own children) can be justified quite possibly by results evaluated from the point of view of a higher-level, evaluator-neutral moral theory (in Hare's own case, utilitarianism). Another possibility arises from the higher-level morality valuing people's working moralities and valuational methods themselves. The higher-level moral values might judge favorably the evaluator relativity shown by people with warm concern, say, for their own families.<sup>51</sup> Thus an essentially evaluator-neutral higher-level moral theory could directly encourage evaluator-relativity in less reflective valuations.<sup>52</sup>

Neither of these two lines of reasoning need dispute that evaluator-relativity involves contradictory beliefs, that is, (1) need not be disputed. But contradictory beliefs are taken to be fruitful either—directly—because of "valued valuations," or—indirectly—because of instrumental use, in each case disputing (2). I shall not pursue these possible lines here.

I turn now to (1). Much depends here on the *nature* of the exercise of morally evaluating states of affairs. One interpretation, which, again, I shall not pursue here, but which potentially can accommodate evaluator relativity, is the "preference" interpretation. John Harsanyi's concept of "ethical preferences" defined over social states is an example of the preference interpretation of moral valuation.<sup>53</sup> Different preferences about the same states need not involve con-

50. See Hare, "Ethical Theory and Utilitarianism."

51. I am grateful to Ronald Dworkin for helpful discussion of this possibility and for pointing out the relevance of valuational methods in judgments of 'character,' which could figure prominently in the higher-level morality. Some related issues have been explored in Bernard Williams in "Persons, Character and Morality" even though his own approach is rather different. It should be added that there are some interesting consistency problems in valuation-valuing evaluations, but I shall not pursue them here.

52. See, however, the 'Introduction' to *Utilitarianism and Beyond*, ed. A. Sen and B. Williams, on some difficulties with two-level moral theories.

53. J. C. Harsanyi, "Cardinal Welfare, Individualistic Ethics, and Interper-

tradictory beliefs. Of course, it might well be taken to be appropriate to impose evaluator neutrality as a second level moral condition. (Indeed, in the definition of "ethical preferences" Harsanyi incorporates something that may push us in that direction, to wit: ethical preferences must "express what he prefers in those possibly rare moments when he forces a special impartial and impersonal attitude upon himself.") But once moral valuations are seen as preferences of some kind, the door to evaluator relativity is wide open, e.g., *you* must disprefer a state in which *you yourself* twist a child's arm (corresponding to Nagel's example of deontology).

One feature of the preference interpretation of moral valuation of states of affairs is that it is very close to judging *actions* themselves, since preferences can be interpreted in terms of hypothetical acts of choosing between alternative states. Contradictory beliefs on values of states are replaced, on this interpretation of moral evaluation, with disparate attitudes to action choice. I leave the preference interpretation here and pursue other interpretations.

If moral valuations of a state by different persons are seen as alternative statements about the same characteristic of the state, then clearly evaluator relativity must involve *contradictory* statements. But *are they* really about the same characteristic despite that appearance? I would like to explore the possibility that they are coherently interpretable as "positional" statements, reflecting the view of the state from the position of the evaluator.

An analogy might be useful to bring out the distinction. Contrast the two following statements:

- (A) "Mount Everest is beautiful."
- (B) "From here, Mount Everest is beautiful."

Although an assertion and a denial of (A) by two persons respectively—taken literally—must be seen as contradicting each other, this is not necessarily the case with (B), since the two persons may be occupying different positions. If moral judgments about states of affairs are like aesthetic judgments of type (B), evaluator relativity of outcome morality need not imply holding contradictory views.

---

sonal Comparisons of Utility," *Journal of Political Economy* 63 (1955). See also K. J. Arrow, *Social Choice and Individual Values* (New York: Wiley, 1951).

I would like to emphasize that the distinction between (A) and (B) does not correspond to the objective-subjective dichotomy. On the objective interpretation of values, anyone standing at the same position (say, fifty miles due north of Mount Everest, or—in the moral case—being opposed to chemical and biological warfare) may well be facing the same objective value. There are many such objective “positional” judgments, for example regarding the relative brightness of stars or relativities of movements of bodies, which need not have anything to do with morals, values, or aesthetics: The issue of objectivity of values is left completely *open* in moving from interpretation (A) to interpretation (B).

Certainly (B) should not be interpreted as endorsing the view that beauty or moral value lies in the eyes of the beholder. Given the viewers occupying the same position regarding the states, the viewers may well be required by the given evaluator-relative morality to take the same view of the state. Being Desdemona’s husband, lover, and killer—whether or not Othello—or being the person twisting the child’s arm, may determine a person’s position in regard to a state with sufficient definiteness to admit little or no variation of the evaluation of the corresponding states within that moral theory, irrespective of the exact identity of the person.

It follows that some types of evaluator relativity would indeed be contradictory for an outcome morality related to a certain identification of distinct positions. For example, if the position of being the twister of the child’s arm is a crucial one for a particular moral theory, then two persons, differing in other respects but both *jointly* twisting the child’s arm, are not free to judge quite differently the resulting state by virtue of their being *different evaluators*. If evaluator-relativity is derived from the “positional interpretation,” then it is really *position relativity* that is admitted, which entails evaluator relativity only to the extent that evaluators differ from each other in their respective positions.

It should be noted that the aesthetic analogy in terms of which the positional concept of morality was explained is misleading in at least one respect. In the case of aesthetics it is possible in a natural way to move from statements of type (B) to those of type (A), since one can typically vary one’s position in pursuing aesthetic appreciation. Statement (A), that is, “Mount Everest is beautiful,” can be inter-

preted as: "there *exists* a position" (or perhaps "there exist many positions") "from which statement (B) is correct." When it is asserted that Renoir's *La Loge* is beautiful, the claim isn't that statement (B) is correct no matter what one's position is vis-à-vis the painting (it *isn't*, I suspect, beautiful from a distance of 2 inches, nor of 200 yards), but that one *can* choose positions from which (B) can be emphatically affirmed. (A) is obtained, in the aesthetic case, from 'existential' use of (B).

In the case of moral evaluation of states the position of the evaluator may not be left open in a similar way. Othello cannot cease to be Desdemona's husband, lover, and killer for the purpose of the evaluation of *that* state. Nor can George admire the state in which he, the pacifist, works for the war machine by evaluating the state as a no-nonsense hawk. And if you are actually twisting the child's arm, you have to evaluate *that* state from that position. It is not that you must lack the *ability* to imagine what it would be like to evaluate the state from a different position—indeed  $G_i^k(x)$  expresses just that for person  $k$ —but that one of the positions in that state is peculiarly *your own*. This contrasts with the aesthetic case, in which there is nothing to tie you to a certain spot in evaluating a picture.<sup>54</sup> It is this *positional fixity given the state* that makes translation of statements of type (B) into statements of type (A), through 'existential' search, that much less admissible in moral theory than in aesthetic theory.

Thus, moral evaluation of states is more firmly anchored in the positional interpretation than aesthetic evaluation is. And evaluator relativity need not involve any contradictory beliefs under the positional interpretation so long as it is subsumed by position relativity.

This does not, of course, deny the substantive possibility that some things may be valuable or disvaluable from every position (starvation or acute suffering no matter to whom it occurs, for example, being a moral disvalue for every evaluator). The positional interpretation of moral judgments of states *permits* categories of moral thoughts not admissible in more traditional formats. How divergent, substantively, the positional valuations may be is a different question. Indeed, the exact scopes of both agent-relative *action* judgments and position-relative *state* judgments remain open issues. However, although some scope for agent-relative action judgments is widely accepted, there

54. One is, as it were, a part of the 'picture' in the moral case.

has been no corresponding acceptance of the admissibility of position-relative state judgments. My main concern in this Section has been to assert that admissibility.

### VIII. NEGATIVE AND POSITIVE FREEDOMS

The necessity of goal rights was argued in Sections I through III on the basis of the inadequacy of moral systems that do not give right-based considerations any role in outcome judgments. These goal rights may or may not be formulated in terms of two-person right-duty correspondences, and they can take the form of capability rights and corresponding duties (Section IV).

Of the more traditional approaches to rights, there is no particular difficulty in accommodating "instrumental" rights in consequence-based evaluation, giving them a status *through* consequential reasoning despite noninclusion in the goals (Section IV). Issues of coordination and cooperation can, of course, be both important and complex, but they are really concerned with how to *use* consequence-based evaluation, bringing in strategic considerations in the assessment of consequences.

It is the deontological approach—important especially for consideration of "negative freedom"—that may appear to be difficult to accommodate within consequence-based evaluation. However, deontological values can, in fact, be accommodated within consequence-based evaluation through evaluator-relative outcome moralities (Sections VI and VII). This would yield *doer* relativity and viewer relativity of *action* judgments of the kind imposed by deontology, but will operate through consequence-based reasoning.

In a goal rights system with consequence-based evaluation,<sup>55</sup> negative freedom can figure in at least three distinct ways:

- (1) *as parts of goal rights* (for example, as capability rights considered jointly with positive freedoms or as negative rights incorporated in the goals in the form of, say, noncoercion,

55. A goal rights system may or may not be consequence-based, since the dividing line is merely sensitivity to consequences (see Section IV) rather than the hegemony of outcome evaluation on the judgment of control variables. However, *consequence-based* goal rights systems form an important subclass, with a defined structure.

- though the merit of the latter approach was disputed in Section IV);
- (2) *for instrumental reasons* (through the influence of negative freedom on the fulfilment and nonfulfilment of other goals);
  - (3) *through evaluator-relative values in the outcome morality* (for example, 'deontological' disvalue to an evaluator of a state in which he has done some 'bad' act).

This differs from the deontological constraint-based approach in at least two respects:

- (i) admitting the possibility of these rights being *overridden* for other benefits—related to other rights or nonright goals—if they are sufficiently strong; and
- (ii) entailing duties of "third parties," who can help or hinder, requiring the evaluation of duties through consequential analysis of interdependences (possibly in an evaluator-relative form).

Positive freedoms can receive support through reasons of types (1) and (2). They can also receive support through reasons of type (3) in some cases, for example, through the aims of *i*, giving a special weight on *j*'s positive freedoms (connected with 'autonomy' reasons). However, the particular case of 'deontological' reasons working against 'bad' acts by the evaluator will apply to negative freedoms in a way that it would not to positive freedoms.

In this sense, negative freedom may well be valued in "one more way" than positive freedom. However, that does not, of course, imply that negative freedoms on balance will be more important than positive freedoms. On the approach outlined in this paper, the relative importance of different types of rights and duties must depend ultimately on the comparative analysis of consequences and their evaluations.

For extremely helpful comments on earlier drafts, I am most grateful to Derek Parfit, and also to James Griffin, Isaac Levi, John Mackie, Donald Regan, Bernard Williams, and the Editors of *Philosophy & Public Affairs*. I have also benefited greatly from discussions with Michael Bratman, G. A. Cohen, Ronald Dworkin, Hans Herzberger, Ted Honderich, Stig Kanger, Saul Kripke, John Perry, John Rawls, Patrick Suppes, Charles Taylor, and Richard Wollheim.