
THE JOURNAL OF PHILOSOPHY

VOLUME LXXXII, NO. 4, APRIL 1985

WELL-BEING, AGENCY AND FREEDOM THE DEWEY LECTURES 1984*

MORAL INFORMATION

THE main aim of these lectures is to explore a moral approach that sees persons from two different perspectives: *well-being* and *agency*. Both the “well-being aspect” and the “agency aspect” of persons have their own relevance in the assessment of states and actions. Each aspect also yields a corresponding notion of freedom.

A second objective is to examine a set of metaethical issues, making use of an “informational” approach to moral analysis which focuses on the admissibility and use of different types of information in moral valuation (by imposing a set of “invariance restrictions” on moral judgments). The two objectives are closely related in the sense that the substantive moral analysis draws on the metaethical arguments.

The first lecture is concerned primarily with metaethical issues, the second with well-being, and the third with agency.

I. INFORMATION AND INVARIANCE

Informational analysis can be used to bring out the content, scope, and limitations of different moral principles. Each moral principle needs some types of information for its use, and—no less impor-

* These John Dewey Lectures were presented at Columbia University on September 17, 18, and 19 of 1984. They constitute the fifth series of John Dewey Lectures, which were established in 1967 to honor the late John Dewey, who had been from 1905 to 1930 a professor of philosophy at Columbia.

In preparing these lectures for publication, I have benefited greatly from extensive comments of John Bennett, G. A. Cohen, Ronald Dworkin, Richard Hare, Isaac Levi, Tom Nagel, T. M. Scanlon, Ian White, and Bernard Williams. I have also had helpful suggestions from Bruce Ackerman, Kenneth Arrow, John Broome, Partha Dasgupta, James Griffin, Peter Hammond, Ted Honderich, Susan Hurley, Hidé Ishiguro, Stig Kanger, John Mackie, Sidney Morgenbesser, Derek Parfit, and John Rawls.

tantly—"rules out" direct use of other types of information. In their latter role, moral principles impose "informational constraints," demanding that certain types of information should not be allowed to influence the moral judgments to be made.

Informational constraints are typically imposed implicitly—entailed by other conditions—rather than stated explicitly.¹ But some principles do take the explicit form of informational constraints. For example, typical requirements of "universalizability" of moral judgments are direct informational constraints. The necessity to make "similar" judgments in "similar" circumstances rules out the use of all types of information not included in the notion of "similarity" of circumstances.

The basic form of an informational constraint is that of an invariance requirement: if two objects x and y belong to the same *isoinformation set* I , then they must be treated in the same way J . For all $x, y: x, y \in I \Rightarrow J(x, y)$. Such an invariance requirement is specified in a particular context, involving (among other things) the characterization of "objects" (x, y , etc.) and that of being treated "in the same way" (J). The invariance requirement takes the form of specifying, in that context, a class of isoinformation sets of objects, and any two objects in a given isoinformation set I must be treated in the same way.² An invariance restriction is based on a reading—explicit or implicit—of some types of information as being relevant for the moral judgment in question, and others not so. It asserts that any difference between two objects x and y belonging to the same isoinformation set is *irrelevant* in the current context.

A few examples may help to illustrate the diversity of ways in which informational constraints may be used through the specification of invariance requirements. When Henry Sidgwick claimed that "if a kind of conduct that is right (or wrong) for me is not right (or wrong) for some one else, it must be on the ground of some difference between the two cases, other than the fact that I and he are dif-

¹ Implicit uses of informational conditions are examined and illustrated in my "Informational Analysis of Moral Principles," in Ross Harrison, ed., *Rational Action* (New York: Cambridge, 1979).

² The role of informational constraints in the form of invariance restrictions in social choice theory is examined in my *Collective Choice and Social Welfare* (San Francisco: Holden-Day, 1970; republished, Amsterdam: North-Holland, 1979), hereafter, CCSW, and "On Weights and Measures: Informational Constraints in Social Welfare Analysis," *Econometrica*, XLV (1977): 1539–1572, reprinted in *Choice, Welfare and Measurement* (Oxford: Blackwell, and Cambridge, Mass.: MIT Press, 1982), hereafter, CWM. See also Claude d'Aspremont and Louis Gevers, "Equity and the Informational Basis of Collective Choice," *Review of Economic Studies*, XLIV (1977): 199–209, and Kevin Roberts, "Interpersonal Comparability and Social Choice Theory," *Review of Economic Studies*, XLVII (1980): 421–439.

ferent persons,"³ he used a powerful informational constraint to rule out taking note of personal identity, as such, in making these judgments of conduct. Other requirements of "universalization" of moral judgments take the form of similarly general informational constraints.⁴

Another example—rather more specialized than Kantian and other requirements of universalization—is Kenneth Arrow's condition of "the independence of irrelevant alternatives" in social choice.⁵ Here the "objects" x and y stand respectively for alternative n -tuples of individual preference orderings $\{R_i\}$ and $\{R_i^*\}$. The two are seen as belonging to the same isoinformation set, in the context of choice from a given subset S of social states, if and only if each individual's preference ordering of the states in S is the same in the first case as in the second (no matter how the two preferences differ regarding states *not* in S). The interpretation of $J(x,y)$, i.e., of x and y being treated in the same way, is that the *same* social choice is made from S in the two cases.⁶ The *independence condition* demands that, in the context of choice from the set S , all information *other than* the individual rankings of the elements of S by each person must be regarded as irrelevant and such information must not influence social choice over S .

Another interesting example is that of the Pareto principle, which is widely used in welfare economics. This can be defined in rather different forms, each imposing some informational constraint. One particular form—that of the *Pareto indifference rule*—is *purely* informational, and this demands that, if everyone is indifferent between two states x and y , then x and y are socially exactly as good as each other. Here the objects are states; the interpretation of $J(x,y)$ is that state x and state y are socially equally good; and the isoinformation sets are based on the congruence of x and y in individual preference rankings. The informational constraint imposed by this invariance restriction rules out the use of any information other than the placing of the respective states in individual preference orderings in the special case in which everyone happens

³ *The Method of Ethics* (London: Macmillan, 7th ed., 1907), Book III, ch. XII, p. 379.

⁴ Different types of universalization requirements are carefully distinguished by John Mackie, *Ethics: Inventing Right and Wrong* (Harmondsworth: Penguin Books, 1977), ch. 4. See also Richard Hare, *Moral Thinking: Its Levels, Method and Point* (New York: Oxford, 1981), ch. 6.

⁵ *Social Choice and Individual Values* (New York: Wiley, 1951; 2nd ed., 1963).

⁶ There can, however, be different interpretations of what "social choice" stands for. Some alternative ones have been explored and distinguished from one another in my "Liberty and Social Choice," this JOURNAL, LXXX, 1 (January 1983): 5-28. The independence condition can be correspondingly defined for each of the distinct interpretations.

to be indifferent between x and y . Given the congruence of two states in individual preference rankings, no *other* information can have any influence on the ranking of the two states in terms of social goodness, according to the Pareto indifference rule. Indeed, we do not have to know anything else at all about the two states to be able to make this judgment, and any other information would, in this context, be redundant and useless.⁷

II. UNIVERSALIZABILITY AND INVARIANCE

The informational constraints imposed by requirements of universalizability used in the literature have been justified in several different ways. They can be taken to be substantive moral principles, on which we have to exercise judgment, or as being necessitated by the very nature of ethics.⁸ In his *Language of Morals*,⁹ Richard Hare has powerfully argued that universalizability in the form of the need to make identical judgments in identical circumstances is entailed by the meaning of moral language. It can, however, be argued that if this claim is accepted, then it would follow that the so-called "Hume's law" (to wit: the impossibility of deriving moral propositions from factual premises), endorsed by Hare, among others, must be, strictly speaking, false.¹⁰ If two actions are factually identical, then universalizability (seen by Hare as a necessity due to the meaning of moral language) will require that they be *morally* equivalent.

In fact, this can also be seen as an aspect of the theory of identity, whereby if $x = y$, then $f(x) = f(y)$. If the "theory of valuation," to use John Dewey's term,¹¹ indicates that moral values must be seen as a function f of factual information, then factual identity ($x = y$) must entail moral equivalence: $f(x) = f(y)$.¹² This does not depend on the nature of the function $f(\cdot)$, except that it *is* a function.

The relevance of this line of reasoning does, however, depend on

⁷ Though the Pareto principle is usually taken to be "mild" and "innocuous," its cutting power is indeed quite severe, because of the power of the informational constraint in excluding the use of nonpreference information in judging the social goodness of states of affairs. It is a demanding condition on its own and becomes remarkably exacting when combined with other conditions (such as Arrow's independence condition and "unrestricted domain," or a condition of personal liberty). These issues are discussed in my "Personal Utilities and Public Judgments: Or What's Wrong with Welfare Economics?" *Economic Journal*, LXXXIX (1979): 537-558, reprinted in CWM; and in "Utilitarianism and Welfarism," this JOURNAL, LXXVI, 9 (September, 1979): 479-487.

⁸ Sidgwick took "the Kantian maxim" to be "certainly fundamental, certainly true, and not without practical importance" (*op. cit.*, p. xvii).

⁹ New York: Oxford, 1952; rev. ed., 1961.

¹⁰ This proposition was stated and defended in my "Hume's Law and Hare's Rule," *Philosophy*, xli (January 1966): 75-79. Richard Hare's response to this point can be found in *Moral Thinking* (1981), pp. 233/4.

¹¹ *Theory of Valuation* (Chicago: University Press, 1939).

¹² It is sometimes argued that to say that two actions (or two states) are morally

the interpretation of the idea of "identity" in dealing with universalizability of moral judgments. There is an exacting sense in which it can be said that two factual circumstances are never exactly identical. The notion of identity used in the application of universalizability is based on some informational constraint, ruling out some differences as irrelevant. Richard Hare, in fact, concentrates on different cases being "identical in their *universal properties*".¹³ There is an important problem in deciding whether the informational constraints that rule out the relevance of what are seen as *nonuniversal* properties must not themselves be substantive moral requirements—not necessitated by the language of morals itself. What is at issue is the *scope* of the principle of universalizability (seen as a necessity due to the nature of moral language).¹⁴ I shall not pursue this question further here. The theses presented in these lectures do not depend on taking a particular position on this question.

The possible conflict between the demands of Hume's law and the thesis that universalizability is necessitated by the nature of moral language, may well be a rather esoteric question. But in focusing on the need to relate moral valuation functionally to factual circumstances, the issue of universalizability does relate to the more general question of the "objectivity" of moral beliefs.¹⁵ The question of ob-

equally good is not in itself a moral judgment. According to this view, "x is morally good" is a moral judgment, but "x and y are morally equally good" is not, since it does not specify whether either of them is "morally good." (I am grateful to Ted Honderich for forcefully raising this eminently discussable question, when responding to a version of this paper at a meeting of the Philosophical Society in Oxford.) I believe that such a view is hard to defend, since moral equivalence must be saying *something* about moral matters. In fact, there is a strong case for taking the relation "at least as good as" as the primitive concept in goodness judgments (as, say, in Arrow's framework, *op. cit.*, ch. 2). From this relation, that of being "equally good" can follow through conjunction. It can even be argued that a statement such as "This action is good" can be fully understood only in terms of some comparison (usually implicit) with *other possible actions*. The idea of being "equally good" belongs solidly to that inescapably comparative framework of moral judgment.

¹³ *Moral Thinking* (1981), p. 108; italics added.

¹⁴ The exact proposition put forward in my "Hume's Law and Hare's Rule" (1966) was, in fact: "Either Hare's principle of universalizability is empty of content, or it conflicts with Hume's Law" (79).

¹⁵ See Thomas Nagel, *Mortal Questions* (New York: Cambridge, 1979), Essay 14, and "The Limits of Objectivity," in S. McMurrin, ed., *Tanner Lectures on Human Values*, vol. I (Salt Lake City: Utah UP, 1980). See also Simon Blackburn, "Moral Realism," in J. Casey, ed., *Morality and Moral Reasoning* (London: Methuen, 1971); David Wiggins, "Truth, Invention, and the Meaning of Life," *Proceedings of the British Academy*, LXII (1976); Donald Davidson, "Psychology as Philosophy," in *Essays on Action and Events* (New York: Oxford, 1980); John McDowell, "Noncognitivism and Rule-following," in S. H. Holtzman and C. M. Leich, eds., *Wittgenstein: To Follow a Rule* (London: Routledge & Kegan Paul, 1981); Susan Hurley, "Objectivity and Disagreement," Honderich, ed., *Ethics and Objectivity* (London: Routledge & Kegan Paul, forthcoming).

jectivity will be taken up later on in this lecture (in section VII), but what is important to stress here is the need, in this context, to examine the invariance requirements that must be imposed on moral beliefs. Objectivity of moral beliefs must require some types of invariance, but the necessity to make identical moral judgments in *exactly* identical circumstances is perhaps too “minimal” to take us very far into that question. It will be argued later that what may be called *authorship invariance* is indeed required by objectivity of moral beliefs, but this can coexist with the possibility that moral valuations of different persons (sharing exactly the *same* moral system) may vary with the different positions that they respectively occupy in the given states of affairs. Invariance conditions involving informational constraints will have to be formulated differently in that context (section VII).

III. UTILITARIANISM AND INFORMATION

It is manifest that the actual availability of information can affect the possibility of using specific moral approaches. For example, Lionel Robbins's celebrated attack on utilitarian welfare economics—an attack which had a profound influence on the direction of welfare economics for several decades—centered mostly on the unavailability of information regarding interpersonal comparison of utility, making utilitarianism allegedly unworkable.¹⁶ That attack was rather wrong-headed, and it was based on a failure to grasp the nature and the evidential basis of utility comparisons.¹⁷ But certainly the more difficult it is to get utility information, the more limited is the scope of application of utilitarianism. Other moral approaches are similarly constrained by the availability of information.

The dependence of the usability of moral principles on the availability of information is obvious enough. What is perhaps less obvious is the dependence of the usability of information on the particular moral principles adopted. This question of *implicit* informational constraints was briefly discussed earlier (in section I). A substantive moral system can indeed be substantially—though not

¹⁶ *An Essay on the Nature and Significance of Economic Science* (London: Macmillan, 1932), and “Interpersonal Comparisons of Utility,” *Economic Journal*, XLVIII (1938).

¹⁷ See particularly I. M. D. Little, *A Critique of Welfare Economics* (New York: Oxford, 2nd ed., 1957); Richard C. Jeffrey, “On Interpersonal Utility Theory,” this *JOURNAL*, LXVIII, 20 (Oct. 21, 1971): 647–656; Ilmar Waldner, “The Empirical Meaningfulness of Interpersonal Utility Comparisons,” *ibid.*, LXIX, 4 (Feb. 24, 1972): 87–103; and my “Interpersonal Comparisons of Welfare,” in M. Boskin, ed., *Economics and Human Welfare* (New York: Academic Press, 1979), reprinted in my CWM. See also the papers of Donald Davidson and Allan Gibbard to be included in a volume of essays edited by J. Elster and A. Hylland, provisionally called *Foundations of Social Choice Theory*, to be published by Cambridge University Press, and an unpublished paper by Robert Nozick, “Interpersonal Comparisons of Utility.”

completely—characterized by specifying the informational constraints it imposes.¹⁸

Take simple act utilitarianism, which can be factored into: (1) *act consequentialism* (the goodness of an act is given by the goodness of its consequent states of affairs), (2) *welfarism* (the goodness of a state of affairs is given by the goodness of the utility information regarding that state), and (3) *sum ranking* (the goodness of utility information is given by the sum total of the utilities in question).¹⁹ Other types of utilitarianism, such as rule utilitarianism, motive utilitarianism, etc., would require the replacement of act consequentialism by different consequentialist requirements such as rule consequentialism, motive consequentialism, etc. The common building block of different types of utilitarian consequentialism is given by the combination of some consequentialist form with *outcome utilitarianism*, which is produced by the union of welfarism and sum ranking, making the goodness of states of affairs be given by the utility sum total in that state.

These requirements impose a set of informational constraints, in the form of invariance restrictions linked to specific information types. The isoinformation sets are given by the factored conditions individually or jointly. For example, in judging states of affairs, *welfarism* would put two states in the same isoinformation set if they are identical in terms of utility (i.e., if they are mapped onto the same point in the n -dimensional utility space, in an n -person society). If *sum ranking* is also imposed, this requires further that two states must belong to the same isoinformation set whenever they have the same utility total (even if they are mapped onto different points in the n -dimensional utility space). Finally, additionally imposing a minimal (direction-neutral) version of *act consequentialism* will put two acts (even though they are different acts) in the same isoinformation set if the two states of affairs they respectively yield have the same utility total. And that, of course, takes us more or less all the way to act utilitarianism (all we need further is a “directional” axiom that says the more the better). In assessing act utilitarianism (and any other moral system), it is important to examine both what information it regards as appropriate and what information it constrains out as irrelevant and unusable.

IV. PLURALISM: PRINCIPLES AND INFORMATION

The informational parsimony of utilitarianism reflects its apparently “monist” rather than “pluralist” informational format. This

¹⁸ This is discussed in my “Informational Analysis of Moral Principles” (1979).

¹⁹ The rationale of this factorization is discussed in my “Utilitarianism and Welfarism” (1979). See also the “Introduction” to Sen and Williams, eds., *Utilitarianism and Beyond* (New York: Cambridge, 1982).

appearance of monism is somewhat deceptive because utility can be sensibly seen as a "plural" notion itself,²⁰ and I shall have more to say on this in the second lecture.

The issue of pluralism has been discussed a fair amount, and more specifically, the question has been seen as important in examining alternatives to utilitarianism. Is it necessary that a plausible alternative to utilitarianism be itself monist?

The idea of pluralism vis-à-vis monism not only is ambiguous, but also is sometimes viewed in a way that has no essential relation to the feature of plurality. A common approach seems to take the form of identifying the use of plurality of principles with *intuitionism*.²¹ Pluralism is a claim about the *form* of moral structures, whereas intuitionism—unless radically redefined—is a claim about how the moral structure may be *derived and supported* (i.e., whether by intuition only). The two things are not essentially related. The suggestion that if there are many ultimate principles (and they are not in a hierarchy), then the balance among them *must be* struck by intuition, is, of course, completely baseless. John Rawls, who does restrict the term 'intuitionism' to pluralist moral theories, does at least make clear that "a conception of justice can be pluralist without requiring us to weigh its principles by intuition."²²

Even when the confusing identification of pluralism with intuitionism has been rejected, there remain at least two different ways of defining pluralism. One is in terms of plurality of principles (I shall call this *principle pluralism*), and the other in terms of plurality of informational variables (to be called *information pluralism*).

It is probably not particularly worth while to try to define a metric on the space of principles (including "combining" principles) to be able to count them precisely and to be "really" sure how many principles there are in a particular moral structure. The genuine worry that seems to motivate concern about principle pluralism lies elsewhere and is easy to see; to wit, that a nonintegrated moral structure could leave many moral issues unresolved. In an often-quoted passage, John Stuart Mill put the problem thus:

There must be some standard by which to determine the goodness or badness, absolute and comparative, of ends or objects of desire. And

²⁰ I have tried to discuss different aspects of this plurality of utility in "Plural Utility," *Proceedings of the Aristotelian Society*, LXXXI (1980/81): 193–215. See also Martha C. Nussbaum, "Plato on Commensurability and Desire," *Proceedings of the Aristotelian Society*, LXXXIII (1984): 50–80.

²¹ See J. D. Urmson, "A Defence of Intuitionism," *Proceedings of the Aristotelian Society*, LXXV (1974/75): 111–119, where Urmson contrasts: (i) "mononomic," (ii) "hierarchical," and (iii) "intuitionist" moral structures.

²² *A Theory of Justice* (1971), p. 34.

whatever that standard is, there can be but one: for if there were several ultimate principles of conduct, the same conduct might be approved by one of those principles and condemned by another; and there would be need for some more general principle as umpire between them.²³

The argument clearly has some force.²⁴ It is, however, interesting to note that here Mill is not disputing the case for having a principle-pluralist structure with a combining principle (what he calls an "umpire"), but rather asserting only the need for such a principle for a principle-pluralist structure. He is concerned to show that if that is how pluralism is to be dealt with, then it would again be true that there is "but one" ultimate standard. One could dispute this—the role of the "umpire" does not obliterate the continuing role of the plural principles in the final ranking of actions, states, etc. But not much really hangs on how such a system is described.

V. PLURALISM AND INCOMPLETENESS

The really important questions regarding pluralism seem to be the following. First, the scope of combining principles may be limited,²⁵ the "diversity of goods" may be too deep.²⁶ The conflicting values may be unreconcilable and may be hard to integrate, as Karl Marx argued in his well-known *Critique of the Gotha Programme*, referring to the German Workers Party's simultaneous loyalty to the principles of "equal distribution" and the worker's right to get "the undiminished proceeds of labor."²⁷ On a different type of problem, it has been argued, by Isaiah Berlin in particular, that elimination of value conflicts inevitably implies some loss of value, and perhaps even that "we cannot conceive of a situation in which it was true both that all value-conflict had been eliminated, and that there had been no loss of value on the way."²⁸

²³ *A System of Logic*, Book VI, ch. XII, s. 7. See also Hillel Steiner, "Reason and Intuition in Ethics," *Ratio*, xxv, 1 (June 1983); 59-68.

²⁴ One point to note, though, is that it cannot be assumed that the existence of "several ultimate principles of conduct" *must* necessarily lead to conflict. Whether it would or not would depend inter alia on the overlap of the domains of the respective ultimate principles. Conditions of compatibility of principles are discussed in my "Informational Analysis of Moral Principles" (1979).

²⁵ See Bernard Williams, *Moral Luck* (New York: Cambridge, 1981), especially Essay 5 ("Conflicts of Values").

²⁶ See Charles Taylor, "The Diversity of Goods," in Sen and Williams, *op. cit.*, in which see also Stuart Hampshire, "Morality and Convention." For a different aspect of the issue of diversity and commensurability, see Nussbaum, *op. cit.*

²⁷ *Critique of the Gotha Programme* (1875; New York: International Publishers, 1966).

²⁸ The quoted sentence is Bernard Williams' statement of one of Berlin's arguments; *Moral Luck*, p. 71. See Berlin, *Concepts and Categories* (New York: Oxford, 1980).

If value conflicts do remain unresolved and if conflicting principles are not "umpired" together (to use Mill's phrase) by some balancing principle, then the over-all ranking may well be incomplete (if such a combined ranking is sought).²⁹ The "intersection" of the different criteria will, of course, yield a partial order, and although that "intersection partial order" may be extendable by weights that are not fully determined (and could vary over a wide range), the "maximal" partial order, given everything, may still be quite seriously incomplete. Then there may not, quite possibly, exist any "best" action (or "best state"). Is that incompleteness an embarrassment for a principle-pluralist moral approach? That is, indeed an important question, to be faced.

Second, if we do happen to resolve the problem of incompleteness by some combining principle, leading to what Mill takes to be a unique "standard," would that integrated system be *informationally* monist, in the sense of taking only one type of information (like utility) to be intrinsically valuable?

I take up the second question first. The answer is straightforward: informational monism is a completely different issue and is not required for an integrated, complete structure. Indeed, if integration is brought about by a combining principle "umpiring" together various demands based on various informational bases, it would *have to be* information-pluralist. Even if we are entirely persuaded by the quoted argument from Mill, there would still be no particular reason for us to look for some single informational base of the kind that, say, utility allegedly provides. There is no argument in all this for expecting that moral goodness must be ultimately decidable by counting the units of some homogeneous nonmoral quantity (as utility is taken to be), rather than having to balance the relative importance of different considerations that conflict.³⁰ It is not so much that information pluralism can be defended, as that there is no special need for a defense.

I turn now to the first question—the alleged need for completeness. In a recent paper, "Moral Struggle,"³¹ Isaac Levi has followed the lead of John Dewey in discussing the importance of conflicts of principles that may or may not get resolved, and has addressed the

²⁹ On the "disparity between the fragmentation of value and the singleness of decision," see T. Nagel, "The Fragmentation of Value," in *Mortal Questions*. See also R. Dworkin, "No Right Answer?" in P. M. S. Hacker and J. Raz, eds., *Law, Morality and Society: Essays in Honour of H. L. A. Hart* (New York: Oxford, 1977).

³⁰ See also Sen and Williams, "Introduction," pp. 16–18.

³¹ Mimeographed, Philosophy Department, Columbia University, 1984; this is in fact the first chapter of a forthcoming book.

problem of decision making under those circumstances. To quote Dewey and Tufts:

The struggle is not between a good which is clear to him and something else which attracts him but which he knows to be wrong. It is between values each of which is an undoubted good in its place but which now get in each other's way. He is forced to reflect in order to come to a decision. Moral theory is a generalized extension of the kind of thinking in which he now engages.³²

Decision making with unresolved conflicts is part of a more general problem, viz., decision making with *incomplete* rankings.³³ We may begin by noting that incomplete rankings may have to be faced even with *one* ultimate principle of valuation. The most obvious reason for this is the limitation of informational availability. For example, partial comparability of utilities—in *interpersonal* correspondences—will make the utilitarian ranking a partial ordering.³⁴ It can also be argued that partial comparability of utilities may be caused not only by the actual limitation of factual knowledge, but also by the intrinsic nature of interpersonal utility comparisons. Utilities may not, in this view, be things that can, even in principle, be fully compared between persons, leading to an exact one-to-one correspondence of the utility measures (levels, units, etc.) of different persons.³⁵ The incompleteness of the utilitarian ranking would, on this view, be unremovable through richer factual knowledge. In these circumstances, a pure utilitarian would have to *assert* the incompleteness of moral ranking as the correct moral position.

In fact, depending on the status of the incompleteness, it is possible to distinguish between (1) open incompleteness, and (2) closed incompleteness. The former is exemplified by a partial ordering that can still be extended, e.g., by more information (as in the case of utilitarianism with contingent partial comparability), or by the use of some other supplementary principle (e.g., extending the in-

³² J. Dewey and J. H. Tufts, *Ethics* (New York: Holt, 1932), p. 175.

³³ More generally, decision making with incomplete *valuation structures*, which may or may not be binary.

³⁴ See my "Interpersonal Aggregation and Partial Comparability," *Econometrica*, xxxviii (1970), reprinted in CWM; and *Collective Choice and Social Welfare* (1970), ch. 7*. See also C. Blackorby, "Degrees of Cardinality and Aggregate Partial Orderings," *Econometrica*, xliiii (1975); B. J. Fine, "A Note on 'Interpersonal Aggregation and Partial Comparability,'" *ibid.*; and K. Basu, *Revealed Preference of Governments* (New York: Cambridge, 1979).

³⁵ See my *Collective Choice and Social Welfare* (1970), ch. 7, and CWM, Essays 9 and 12.

complete ranking given by the Pareto criterion by some supplementary "distributional" judgment), or by moving toward the resolution of "moral struggles" (as in the case of conflicts of principles). The partial ordering stays, as it were, "ready" to be extended.

In contrast, for "closed" incompleteness, the absence of completeness is *assertive*, and there is no room for extending the partial order. In stating that x and y "cannot" be ranked vis-à-vis each other in terms of moral goodness, a substantive and definite moral assertion is being made. Though the demands of practical choice may force us to choose one or the other, there is—on this view—no additional moral criterion that can be used to rank the unranked pairs *in terms of moral goodness*.

Assertive incompleteness, like assertive completeness, has the effect of closing a substantive issue, declaring further search to be pointless. There is obviously an important problem here of decidability. It is also arguable that assertive incompleteness should never be "accepted," because that may discourage moral inquiry.³⁶ This methodological point (regarding search and research) might well be supportable, but it does not, of course, rule out the real possibility of assertive incompleteness.

When an incomplete ranking refers not to moral goodness as such but to some specific moral quality, e.g., justice, then the completion of that partial ranking by some *other* criterion may give a combined, complete view of moral goodness, but this would do nothing to complete the ranking of justice itself, which would continue to remain assertively incomplete. It could then be the case that while, say, state x is more just than y , states x and z simply cannot be compared in terms of justice (i.e., neither is more just than the other, nor can they be seen as equally just; they are simply non-comparable as far as justice is concerned). And that would continue to be so even if the incompleteness could be removed in the over-all ranking, or in decision making, by the supplementary use of some *other* criterion.

Yielding complete orders cannot be an a priori requirement of the legitimacy of a moral principle, or even of an entire moral approach with its full structure of principles. It is not a matter of getting metamoral passmarks. It does not determine a substantive moral approach, but rather *gets determined* by such an approach, which could yield completeness of valuations or not (possibly even *asser-*

³⁶ The strategic role of openness in this sense is discussed in my note, "The Nature and Classes of Prescriptive Judgments," *Philosophical Quarterly*, xvii, 66 (January 1967): 46-62.

tively not). If the appropriate moral approach produces the result that x and y are morally noncomparable, then decision making has to reckon with that.

It is often pointed out that in real life we have to choose something or other—we must decide, we must act. This is certainly so, and we must end up doing something or other, if only nothing, by default. If our actions are to be morally guided, then our incomplete moral valuations must be used to eliminate some alternatives, while permitting others.³⁷ We must not act like a moral counterpart of Buridan's ass—needlessly starving by default, unable to figure out which of the nonstarving alternatives happened to be the best.³⁸ Intelligent moral choice demands that we not choose—explicitly or by default—an alternative that we can see is morally inferior to another feasible alternative. But this does not require that the chosen alternative be seen to be “best” in that set of feasible alternatives, since there may be no best alternative at all, given the incompleteness of our moral ranking.

VI. STATES, ACTIONS, WELFARISM, AND CONSEQUENTIALISM

One source of informational pluralism is the possibility that acts themselves may have value or disvalue, distinct from the valuation of states resulting from such acts. This is an important issue in assessing consequentialism and in examining agent-relative moral approaches.

In examining states of affairs, we have to be sure of what they “contain.” That actions are part of the respective states of affairs is sometimes denied, especially to draw a sharp contract between consequentialist and deontological theories. It is hard to defend this exclusion. A state of affairs in which Brutus kills Caesar is not just one in which Caesar has expired. It is one in which the killing of Caesar by Brutus (and others) figures. There is, of course, also the question as to how we choose to *describe* the state.³⁹ But what the

³⁷ The specification of “permissibility” based on an incomplete ranking of moral goodness may, of course, be used to generate a complete binary relation of its own, but the *derived* binary relation of being “equally permissible” will typically lack some of the standard properties of a preference relation, such as transitivity (see my CCSW, pp. 48/9).

³⁸ I interpret Buridan's ass's problem as the inability to rank the two haystacks, rather than—as in the more standard version—viewing them as equally good, in which case *both* haystacks would have been “best” (and either could have been unproblematically chosen in terms of binary optimality). See my “Behaviour and the Concept of Preference,” *Economica*, xi. (1973); reprinted in CWM, pp. 61/2.

³⁹ See Frederic Schick, “Under Which Description?” in Sen and Williams. See also Davidson, “Logical Form of Action Sentences,” in *Essays on Actions and Events* (1980). I have tried to discuss some of the issues involved in description in “Description as Choice,” *Oxford Economic Papers*, xxxii (1980); reprinted in CWM.

state of affairs includes is really a matter of what *in fact* happened. Caesar could have expired in many different ways, and the fact that he got knifed by Brutus must be a *part* of that state, no matter how we *choose* to describe it in one particular context or another. If we choose to leave out this fact in the description of the state of affairs, that is just a decision to be silent about one part of that state.⁴⁰

It is, of course, true that utilitarianism *values* states of affairs in an informationally limited way, attaching no intrinsic importance to nonutility information about states, and thus *inter alia* ignoring actions too—except as variables for causal analysis (in linking utility to objects that yield utility), or perhaps as surrogates for utility (when utility information is hard to get). This feature of utilitarianism is welfarism, and it is a feature that is shared by some other approaches, such as utility-based “maximin” (seeing the goodness of a state as the utility level of the worst-off individual in that state.)⁴¹

It is important to distinguish between the respective roles of welfarism and consequentialism in making the evaluation of *actions* informationally restricted. Welfarism filters out all information other than utility information in the evaluation of *states of affairs*, and it is precisely because of that informational constraint that consequentialism has, in this compound system, the effect of ignoring everything other than utilities (in the consequent states) in *evaluating actions*. Had welfarism not been additionally imposed, consequentialism could have coexisted with taking note of such things as the values and disvalues of actions *through* the valuation of the states, which include these actions.

VII. AGENCY, POSITION, AND OBJECTIVITY

If the action information is not blocked out by welfarism or some other informational constraint) in evaluating states of affairs, the relation of the evaluator and his or her *agency* role in the state has to be considered in the moral assessment of states. It is often thought that although, in deliberating on the correctness of one's own *actions*, one must take ample note of one's responsibility for

⁴⁰ See Charles Taylor's argument for taking out actions from the assessment of consequent states (“The Diversity of Goods,” 1982). On this, see my “Evaluator Relativity and Consequential Evaluation,” *Philosophy and Public Affairs*, XII, 2 (Spring 1983): pp. 128–130.

⁴¹ Inspired by Rawls's Difference Principle, this “maximin” criterion—and its lexicographic variant (“leximin”)—have been much used in welfare economics; see, for example, E. S. Phelps, ed., *Economic Justice* (Harmondsworth: Penguin, 1973); Sen, *On Economic Inequality* (New York: Oxford, 1973); James Meade, *The Just Economy* (London: Allen & Unwin, 1976); A. B. Atkinson, *Social Justice and Public Policy* (Brighton: Wheatsheaf, 1983). Rawls's own Difference Principle is defined, however, not in terms of utility, but in terms of an index of “primary goods.”

these actions, that issue does not arise in the evaluation of states. But since states do include one's own actions (albeit *inter alia*), that contrast is not sharp. One is responsible for, as it were, a part of the state, and as far as one's own actions in that state are concerned, one is responsible for them in a straightforward way. If some act—such as knifing a friend—is a bad thing to do and if the agent who did the knifing was responsible for that act, then that special badness must be a part of the over-all evaluation, *made by the knifer*, of that state. It could, of course, still be the case that, despite the badness of that action, everything considered, knifing Caesar produced a better state of affairs even from the evaluative position of Brutus (“Not that I loved Caesar less, but that I loved Rome more”). However, Brutus cannot evaluate the state, of which his killing Caesar is a part, without taking note of his own responsibility for that act.

A consistent moral approach may, therefore, require that the moral valuation function be *position-relative*. Given Brutus's role in the knifing of Caesar, Brutus cannot view the resulting state—in which that knifing was also a part—in the same way as, say, Appian, the historian, completely uninvolved in the matter, can. The difference can exist even when Appian and Brutus share exactly the same moral approach, and it arises because of the objective differences in their positions vis-à-vis the state in question. A position-relative valuation function builds in the parameter of the assessor's position *within* the valuation function itself.⁴² Consider the following imaginary remark of Appian to Brutus: “If I were you, Brutus, I should have evaluated the state to be even worse than I do now, not having been involved in the knifing of Caesar myself.” The statement is of the form $f(x,a,b) < f(x,a,a)$, in which $f(x,a,b)$ is the valuation of the state x by a (Appian) in the position of b (Brutus), and $f(x,a,a)$, the valuation made by Appian from his own position. Notice that the relativity in question links not to the *authorship* (the second variable), but to the position from which the evaluation is being made (the third variable), i.e., whether a 's position or b 's. The authorship is immaterial for “position relativity” of a moral system.

Indeed, that the approach is not subject-relative (and does not include “subjectivism” *within* the system) can be guaranteed by insisting on an invariance requirement regarding authorship. “Authorship invariance” requires that the moral valuations of a

⁴² See my “Rights and Agency,” *Philosophy and Public Affairs*, xi, 1 (Winter 1982): 3–39, and the exchange with Donald Regan in the same journal next year: Regan, “Against Evaluator Relativity: A Response to Sen,” *ibid.*, xii, 2 (Spring 1983): 93–112.

state in a given moral system must not vary with the person making the judgment, even though it can vary with the *position from which* the valuation is to be made. Despite Appian's assertion (stated earlier) that $f(x,a,b) < f(x,a,a)$, which shows position-relativity, Appian would be bound by the requirements of authorship invariance to assert $f(x,a,b) = f(x,b,b)$.⁴³ "You, Brutus, should evaluate the state in the same unfavorable way as I should if I were in your position."

The possibility of combining position relativity with authorship invariance is also the reason why positionality of moral valuation is perfectly consistent with objectivity of moral values. Moral valuation can be position-relative in the same way as such statements as "The sun is setting." The truth of that statement varies with the position of the person, but it cannot vary from person to person among those standing in the same position. The aspect of objectivity related to authorship invariance, i.e., to $f(x,a,b) = f(x,b,b)$, for all x , a , and b , has to be, I believe, a part of what Thomas Nagel calls objectivity in "the primary sense"—that of "beliefs and knowledge."⁴⁴

VIII. CONCLUDING REMARKS

This lecture has been concerned entirely with general—primarily methodological—issues. The use of informational constraints in the form of invariance requirements has been examined, both in an explicit form (e.g., in the requirement of universalizability) and in implicit ways (e.g., in utilitarianism, through the features of welfarism, sum ranking, and consequentialism). I have also argued against ruling out principle pluralism, information pluralism, and incompleteness in valuations (such as partial orderings of acts, states, or rules). The case for position relativity in moral valuation has been outlined, and it has been argued that position relativity is consistent with authorship invariance, and indeed with the corresponding aspect of objectivity in moral beliefs. The importance of the informational perspective suggests the need for improving the standard informational accounts of states of affairs. These issues will figure rather prominently in the substantive moral discussion to be presented in the next two lectures.

⁴³ A "subjectivist" ethical approach can deny this invariance condition. Indeed, it can even insist that *everyone* must accept the possibility of this denial. This involves authorship invariance of a different type altogether, operating at a "meta" level (invariably denying the requirement of authorship invariance of moral valuation). I am grateful to Thomas Nagel for raising this question.

⁴⁴ "The Limits of Objectivity" (1980), p. 77.

WELL-BEING AND FREEDOM

THOMAS M. SCANLON argues that “utilitarianism derives much of its appeal from alleged difficulties about the foundations of rival views.”¹ He also suggests that “the wide-spread influence of utilitarian principles” derives from the “attractiveness” of what he calls “philosophical utilitarianism,” which is “a particular philosophical thesis about the subject matter of morality, namely the thesis that the only fundamental moral facts are facts about individual well-being” (108). This “philosophical utilitarianism” can be taken to provide a “foundation” for utilitarianism in a rather limited sense.² It specifies an “informational” basis that is indeed highly attractive, but this does not, of course, indicate that there is no further need for justification of this informational basis. However, it does provide an interesting and important way of looking at utilitarianism, and provides a focus for critical assessment. That will indeed be the point of departure of this lecture. But since the term ‘philosophical utilitarianism’ is a little misleading (even nonutilitarian approaches can be consistent with “the thesis that the only fundamental moral facts are facts about individual well-being”), I shall call the use of this informational foundation by the more descriptive (though much less elegant) name ‘well-being as informational foundation’, or WAIF for short.³

Defending utilitarianism with the help of WAIF raises some immediate questions. (1) Is the thesis of this informational foundation really believable? Does it survive critical examination? (2) Do utilitarian principles really follow from WAIF? In fact, I shall split up the last question into two parts: (2a) Is well-being best seen as utility? (2b) Do utilitarian principles follow from WAIF if well-being is seen as utility?

I. WELL-BEING AND AGENCY

I start with the first question: Is the well-being foundation believable? It would, of course, be altogether amazing if moral goodness had nothing to do with well-being. It is easy to argue that well-being is fundamentally important. If we were tied to informational

¹“Contractualism and Utilitarianism,” in Sen and Williams, *op. cit.*, p. 103.

²Scanlon himself goes on to propose a rival foundational structure, of a procedural variety (involving a “contractual account”), motivated by his diagnosis that “what a successful alternative to utilitarianism must do, first and foremost, is to sap this source of strength [of utilitarianism] by providing a clear account of the foundations of nonutilitarian moral reasoning” (103).

³This acronym is not, I fear, very helpful (though the association of loneliness may not be altogether inappropriate for the information monism of WAIF).

monism, that elementary thought would have already pushed us much of the way to WAIF, since there would be “room” for only one thing at the top. But, as was argued in the last lecture, insisting on informational monism is a crude prejudice. By recognizing the importance of well-being, we do not close the door on the possible importance—intrinsic importance—of other things. So the question is not whether well-being is an intrinsically important variable for moral analysis, but whether it is uniquely so.

People have aspects other than well-being. Not all their activities are aimed at maximizing well-being (nor do their activities always contribute to it), no matter how broadly we define well-being within the limits of that general concept. There are goals other than well-being, and values other than goals. “Human beings,” as John Dewey said in *Theory of Valuation*, “are continuously engaged in valuations” (58). It can be argued that something so central to human life—indeed to our being persons—cannot fail to be intrinsically relevant to moral analysis.

Indeed it is useful to distinguish clearly between the “well-being aspect” and the “agency aspect” of a person. John Rawls’s “Kantian constructivism,” in his Dewey lectures,⁴ started at the agency end—with “persons characterized as rational agents of construction.” Persons were seen as “having the moral power to have a conception of the good”; the agency role—broadly construed—came first, and other things followed. Although I shall not follow Rawls’s route of Kantian constructivism, I would argue that the conception of “persons” in moral analysis cannot be so reduced as to attach no intrinsic importance to this agency role, seeing them ultimately only in terms of their well-being.

There is a particular sphere in which such an agency role may be especially important, and that is the person’s own life. Various concepts of “autonomy,”⁵ and “personal liberty,”⁶ relate to this special role of agency in personal life, going well beyond considerations of well-being. The moral foundation of well-being is informationally extremely restrictive, and the agency aspect is much too crucial to leading a life for it to be intrinsically of no moral importance.

It is important to emphasize that in pointing to the informa-

⁴ “Kantian Constructivism in Moral Theory: The Dewey Lectures 1980,” this JOURNAL, LXXVII, 9 (September 1980): 515–572.

⁵ See Jonathan Glover, *Causing Death and Saving Lives* (Harmondsworth: Penguin Books, 1977) for some interesting—and important—conflicts involving autonomy and other values, including well-being.

⁶ Discussed in my CCSW (1970) and in “Liberty and Social Choice,” this JOURNAL, LXXX, 1 (January 1983): 5–28.

tional limitation of WAIF, in that it gives no fundamental place to the agency aspect of a person, the assumption is *not* being made that the agency and well-being aspects of a person are unrelated. For an integrated person it is likely—possibly even inevitable—that the person's well-being will be influenced by his or her agency role. This does not, however, imply that the well-being information itself could capture the important features of agency, or act as its informational surrogate. In fact, some types of agency roles, e.g., those related to fulfilling obligations, can quite possibly have a negative impact on the person's well-being. Even when the impact is positive, the importance of the *agency aspect* has to be distinguished from the importance of the *impact of agency on well-being*.

II. WELL-BEING AND UTILITY

Is well-being best seen as utility? This is not an easy question to answer, for various reasons. The most immediate reason is that there are different interpretations of utility, in particular (1) happiness, (2) desire fulfillment, (3) choice.⁷

The *choice* view of utility, it is fair to say, is more popular among economists than among philosophers. In this view utility is seen as a real-valued (that is, numerical) representation of a person's choice behavior, i.e., what the person chooses from each subset of the set of alternatives (or would choose if such choices were to arise). This view of utility is, of course, purely "ordinal." Attempts to get "cardinal" utility from the choice approach are deeply hampered by the necessity of having to assume a much more elaborate framework, with much stronger requirements of consistency *and* the willingness to accept the influence of arbitrary elements in the numbering system (e.g., that of one's attitude to risk if cardinality is constructed from choices over lotteries). The extension to interpersonal comparisons is also very problematic, since we do not actually face the choice of becoming someone else, and it is hard to decide what importance to attach to answers that may be given when such choices are posed hypothetically. The natural domain of the choice interpretation of utility is in providing non-interpersonally-comparable ordinal valuations, as under Paul Samuelson's "revealed preference" approach.⁸

⁷ See James Griffin, "Modern Utilitarianism," *Revue Internationale de Philosophie*, xxxvi.3, 141 (1982): 331-375, for an illuminating survey of different notions of utility and their use in utilitarian theory and for a critical assessment of some anti-utilitarian arguments, from a firm (but undogmatic) utilitarian perspective. See also J. C. B. Gosling, *Pleasure and Desire* (New York: Oxford, 1969).

⁸ See *Foundations of Economic Analysis* (Cambridge, Mass.: Harvard, 1947). Some of these issues are discussed in my CWM (1982).

Even in this limited form, a serious problem with this approach arises from the fact that a person's choice behavior may fail to have any binary representation, or it may be representable by a binary relation that is not transitive (as the utility relation is required to be).⁹ But perhaps a more basic problem with this approach comes from the fact that a person's choice may be guided by a number of motives of which the pursuit of personal well-being is only one. The well-being motivation may well be dominant in some choices, but not in others. Moral considerations may, inter alia, influence a person's "commitment."¹⁰ The mixture of motivations makes it hard to form a good idea of a person's well-being on the basis of choice information only.¹¹

The other two views of utility—happiness and desire fulfillment—undoubtedly have some plausibility. However, happiness has two basic problems in its claim to stand for well-being. First, as it is interpreted in the utilitarian tradition, happiness is basically a mental state, and it ignores other aspects of a person's well-being. If a starving wreck, ravished by famine, buffeted by disease, is made happy through some mental conditioning (say, via the "opium" of religion), the person will be seen as doing well on this mental-state perspective, but that would be quite scandalous. Second, as a mental-state concept, the perspective of happiness may give a very limited view of other mental activities. There are mental states

⁹ The respective conditions for binary representability and transitive binariness (and other—intermediate—structures), are identified and analyzed in Arrow, "Rational Choice and Orderings," *Economica*, xxvi (1959): 121–127; my "Choice Functions and Revealed Preference," *Review of Economic Studies*, xxxviii (1971): 307–317, reprinted in CWM; Hans Herzberger, "Ordinal Preference and Rational Choice," *Econometrica*, xli (1973): 187–237; K. Suzumura, *Rational Choice, Collective Decisions, and Social Welfare* (New York: Cambridge, 1983).

¹⁰ This issue, among others, is discussed in my "Choice, Orderings and Morality," in S. Körner, ed., *Practical Reason* (Oxford: Blackwell, 1974), and "Rational Fools: A Critique of the Behavioural Foundations of Economic Theory," *Philosophy and Public Affairs*, vi (1977): 317–344, both reprinted in CWM.

¹¹ See Thomas Nagel, *The Possibility of Altruism* (New York: Oxford, 1970); my "Behaviour and the Concept of Preference," *Economica*, xi (1973): 241–259, reprinted in CWM; Stephan Körner, *Experience and Conduct* (New York: Cambridge, 1974); Harvey Leibenstein, *Beyond Economic Man* (Cambridge, Mass.: Harvard, 1976); Tibor Scitovsky, *The Joyless Economy* (New York: Oxford, 1976); Kurt Baier, "Rationality and Morality," *Erkenntnis*, xi, 2 (August 1977): 197–223; Thomas Schelling, *Micromotives and Macrobehavior* (New York: Norton, 1978); Jon Elster, *Ulysses and the Sirens* (New York: Cambridge, 1979); Albert Hirschman, *Shifting Involvements* (Princeton, N.J.: University Press, 1982); Howard Margolis, *Selfishness, Altruism and Rationality* (New York: Cambridge, 1982); George Akerlof, "Loyalty Filters," *American Economic Review*, lxxiii (1983); Michael McPherson, "Economics: On Hirschman, Schelling, and Sen," *Partisan Review*, li (1984): 236–247; Derek Parfit, *Reasons and Persons* (New York: Oxford, 1984).

other than being just happy, such as stimulation, excitement, etc., which are of direct relevance to a person's well-being.¹² Furthermore, mental activities involve *valuation* of one's life—a reflective exercise—and the role of valuation in the identification of the person's well-being obviously cannot be seen in terms merely of the happiness that such reflection itself creates. It is hard to avoid the conclusion that although happiness is of obvious and direct relevance to well-being, it is inadequate as a representation of well-being.

Consider, now, the *desire* approach. There is some possibility of doubt about whether this is also a mental-state view of well-being. There can be, of course, hardly anything more "mental" than desires. But as James Griffin has recently argued in his important survey of "modern utilitarianism," "the desire account does not require that satisfaction of desire translates itself in every case into experience" (338, *op. cit.*). Griffin sees the desire approach as one in which welfare lies in "the state of the world," rather than in "the state of mind."

I believe that Griffin underestimates the extent to which a desire account has to be concerned with "the state of mind." An account of well-being that is to serve as a basis for a utilitarian calculus must be able to present a cardinal and interpersonally comparable view of utility. Neither of these types of information can be obtained just by checking whether the person's desires have been realized in the state of the world. The metric that will be needed for an informationally adequate view of utility cannot be obtained from the observation of objects of desire. The strength of desire must come into the picture, and the desires of different people have to be compared. Thus, the desire account of well-being is also, in an important sense, a mental-state account, though it is not a *purely* mental-state account (given the need to observe the objects of desire *as well*).

Is the importance of desire mainly *evidential*, viz., that it gives evidence—*possible* evidence—of value, or does desire have a "value-giving" role? That is, is it the case that the activity of desiring *makes* the objects of desire have value? This question can be posed in the general context of value as such, and also in the context of the value of well-being in particular.

I begin with the general context. I would argue that the primary

¹² See Scitovsky, *op. cit.*, and "The Desire for Excitement in Modern Society," *Kyklos*, xxxiv (1981).

role of desire is evidential and that the immediate linkage is with what the person regards as valuable. Compare the two following statements:

- (I) I desire x because x is valuable for me.
 (II) x is valuable for me because I desire x .

The former statement is intelligible and cogent in a way the latter clearly is not. Valuing something is a good reason for desiring it, but desiring something is not an obvious reason for valuing it. Of course the latter can still make sense in some special way, which can be explained, e.g., "Having first formed the desire for x , I have come to value x , because being frustrated in my desire will be disappointing and disvaluable."¹³ Though such—and other possible—special considerations do give desiring something more than a *purely* "evidential" role, it is hard to escape the conclusion that the main connection between desire and value is the evidential connection articulated in (I).¹⁴

If this line of reasoning is accepted, then the next step is to investigate the connection between what a person regards as valuable and the value of the person's well-being. That connection will be weakened precisely by the fact that a person may well value things other than personal well-being. Indeed, the agency aspect of a person and the power to form and pursue a conception of the good may well take a person's desires in a direction different from personal well-being, and the evidence for the person's valuation may not fully translate into evidence for his or her well-being. Nevertheless, there is a serious connection here, and the evidential importance of desires in reflecting a person's well-being may well be quite substantial.

III. INTERPERSONAL COMPARISONS OF DESIRES

The evidential role of desires is, however, deeply problematic in the context of *interpersonal* comparisons. Comparative intensities of desire may be a very dubious guide indeed to well-being intensities in comparing one person with another, since these intensities are influenced by many contingent circumstances that are arbitrary for

¹³ Cf. "Well-being and distress are to be understood as two kinds of human experience, those in which desires are satisfied and those in which desires are frustrated" [Ted Honderich, "The Question of Well-being and the Principle of Equality," *Mind*, xc, 360 (October 1981): 481-504, p. 481].

¹⁴ In third-person contexts, desire fulfillment may plausibly be seen as having much strategic importance. The fact that you desire something may be a good reason for me to value it. But this too may have an evidential base. In treating your desire as evidence of what you value, I may value your desire fulfillment as a *consequence* of my valuing your success in getting what you yourself value.

well-being comparisons. Our reading of what is feasible in our situation and station may be crucial to the intensities of our desires, and may even affect what we dare to desire. Desires reflect compromises with reality, and reality is harsher to some than to others. The hopeless destitute desiring merely to survive, the landless laborer concentrating his efforts on securing the next meal, the round-the-clock domestic servant seeking a few hours of respite, the subjugated housewife struggling for a little individuality, may all have learned to keep their desires in line with their respective predicaments. Their deprivations are gagged and muffled in the interpersonal metric of desire fulfillment. In some lives small mercies have to count big.

It may be asked whether this problem cannot be dealt with in the same way as some utilitarians have proposed in "purifying" desires of "imperfection." Hare demands "perfectly prudent" preferences—"fully informed and unconfused." Harsanyi wants that, and *more*; to wit, to "exclude all antisocial preferences, such as sadism, envy, resentment, and malice."¹⁵ These devices raise questions about how we may decide on what "purifications" are appropriate and justifiable within the general logic of utilitarianism, but they may well be able to deal sensibly with some problems, particularly those associated with factual ignorance and thoughtless decisions.

Can such purification procedures eliminate the problem under discussion, viz., the *circumstantial* contingency of desires? It is not clear how they could. Desiring is a part of living; it plays a strategic role in making our wants credible, our aspirations viable. To ask what one would desire in unspecified circumstances—abstracting from the concreteness of *everyone's* life—is to misunderstand the nature of desire and its place in human life. Of course, we can *pretend* to answer this question. Since all this is imaginary anyway, we need not live in the fear of being proved wrong. This can be done by making—explicitly or by implication—some simple assumption, e.g., that our desires would be in line with what Scanlon¹⁶ has called "an objective criterion" of well-being, appealing to a certain "consensus" of values about the content of well-being. But if that is

¹⁵ "Ethical Theory and Utilitarianism," in Sen and Williams, *op. cit.*, p. 28; Harsanyi, "Morality and the Theory of Rational Behaviour," in Sen and Williams, p. 56. See also James Mirrlees, "The Economic Uses of Utilitarianism," and Peter Hammond, "Utilitarianism, Uncertainty and Information," both in Sen and Williams, *op. cit.*

¹⁶ "Preference and Urgency," this JOURNAL, LXXIII, 19 (Nov. 6, 1975): 655-669. Scanlon is, however, primarily concerned with *moral* urgency rather than with getting a concept of well-being that may be used for other purposes as well.

what we are going to do, we could just as easily have started from the objective criterion itself, and "founded" it on the consensus of values on well-being, rather than having the imaginary exercise of *counterfactual desiring*. Since, according to the arguments presented earlier in this lecture, the importance of desiring is primarily "evidential" anyway, that imaginary exercise is particularly pointless. There is little merit in justifying interest in desire information on the ground that it gives evidence of valuation and then going on to "obtain" the "purified" desire information through the knowledge (or assumption) of a consensus of valuations.

IV. EQUALITY AND MORAL VALUATION

I turn now to question (2b): Do utilitarian principles follow from WAIF if well-being is taken as utility? It was argued in the first lecture that utilitarianism is really a combination of welfarism, sum ranking, and some—direct or indirect—version of consequentialism. Therefore, the question can be seen as being about the derivability of these different *factors* from the union of (1) WAIF and (2) the identification of utility with well-being.

Deriving welfarism, obviously, poses no challenge to this task. Nor, in fact, does welfarist consequentialism in general. If nothing other than well-being can count as a morally fundamental fact and if well-being is just utility, then it will be necessary to evaluate actions and other choice variables in terms of utility consequences only.

The factor that remains unmarshalled even in this demanding framework is sum ranking, *viz.*, the view that the utilities are best valued aggregatively by summing them and by ignoring every other aspect of utility distributions, including the extent of inequality. Nothing in WAIF *or* in the identification of well-being with utility can make sum ranking follow automatically without invoking some consideration that makes us unconcerned with equality of well-being. It is, therefore, a matter of considerable interest that arguments have been presented by some utilitarians that make sum ranking allegedly inescapable simply on grounds of logic, mathematics, or rationality.¹⁷ How is this unusual feat performed?

Consider the Vickrey-Harsanyi model¹⁸ of "ethical preferences"

¹⁷ See, for example, John Harsanyi, "Nonlinear Social Welfare Functions: Or Do Welfare Economists Have a Special Exemption from Bayesian Rationality?" *Theory and Decision*, vi, 3 (August 1975): 311-332; M. McDermott, "Utility and Distribution," *Mind*, 91, 364 (October 1982): 572-578.

¹⁸ W. S. Vickrey, "Measuring Marginal Utility by Reactions to Risk," *Econometrica*, xiii (1945): 691-707; Harsanyi, "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economy*, LXIII (1955): 309-321.

based on choosing between different states of affairs in a state of primordial equiprobability in which the chooser has an equal chance of being anyone.¹⁹ If the *as if* behavior of this chooser satisfies certain "rationality" conditions (such as the von Neumann-Morgenstern axioms²⁰), then the position of being any person i in any state x can be given a value $W_i(x)$ such that the chooser's behavior can be represented as maximizing the mathematical expectation of these individual W_i values, i.e., maximizing $\sum_i p_i W_i$, when p_i is the probability of being person i . With the assumption of equiprobability of being anyone, this is no different from maximizing $\sum_i W_i$. If W_i is seen as person i 's utility, then it looks as if sum ranking has somehow been deduced simply from rational behavior. Rationality must, it would seem, preclude any concern for equality of well-being in aggregating different persons' well-being!²¹

But what are these W_i values? They are simply accounting values for predicting choice under uncertainty. They need not coincide with any concept of utility that has independent meaning, such as happiness, or satisfaction, or desire fulfillment. Not only may the values of W_i differ from the utilities U_i as perceived by person i himself (his happiness, his desire fulfillment, etc.), but they may also differ from what the *chooser* expects to enjoy in the counterfactual position of becoming person i . The fact that W is the sum

¹⁹ I abstain in the present context from questioning the acceptability of this general approach to moral valuation, since I am concerned here only with the specific question of how it deals with the issue of inequality. For discussions of different types of problems raised by this moral approach, see P. A. Diamond, "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility: A Comment," *Journal of Political Economy*, LXXV (1967): 765/6; my CCSW, ch. 9; Rawls, *A Theory of Justice* (1971); P. K. Pattanaik, *Voting and Collective Choice* (New York: Cambridge, 1971); Levi, "Four Types of Ignorance," *Social Research*, XLIV (1977): 745-756; E. F. McClennen, "The Minimax Theory and Expected Utility Reasoning," in C. A. Hooker, J. J. Leach, and McClennen, eds., *Foundations and Applications of Decision Theory*, vol. I (Boston: Reidel, 1978); C. Blackorby, D. Donaldson, and J. A. Weymark, "On John Harsanyi's Defences of Utilitarianism," Discussion Paper, University of British Columbia, 1980; M. Machina, "'Expected Utility' Analysis without the Independence Axiom," *Econometrica*, 1. (1982): 277-324; J. Broome, "Uncertainty and Fairness," *Economic Journal*, xciv (1984): 624-632. Harsanyi's own assessment is to be found in his *Essays on Ethics, Social Behavior, and Scientific Explanation* (Boston: Reidel, 1976).

²⁰ J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior* (Princeton, N.J.: University Press, 1944). For an excellent discussion of the contents, interpretation and implications of the axioms, see R. D. Luce and H. Raiffa, *Games and Decision* (New York: Wiley, 1957).

²¹ See also Harsanyi, "Non-linear Social Welfare Functions: A Rejoinder to Professor Sen," and my "A Reply to Professor Harsanyi," in Robert E. Butts and Jaakko Hintikka, eds., *Foundational Problems in the Special Sciences* (Boston: Reidel, 1977).

of the W_i values does not tell us anything at all about what the relationship is between W and the independently characterized utilities U_i of the respective persons. In terms of the language of the von Neumann-Morgenstern model of choice, we might employ the term 'utility' for W_i , but W_i has no necessary role other than yielding accounting values in terms of which the chooser's behavior can be predicted.

In fact, the numbering system of W_i works the way it does precisely by *incorporating* attitudes toward risk in the W_i values themselves. And, in this framework, these attitudes toward risk correspond neatly to attitudes toward inequality.²² If the chooser's attitude toward inequality changes (but "consistency" of choice is retained in the form of continuing to satisfy the von Neumann-Morgenstern axioms), then the W_i functions will change correspondingly, even though the utilities U_i of the persons in the form of happiness or desire fulfillment may remain completely unaltered. The fact that W is the *sum* of the W_i values does not imply sum ranking of the utility values U_i , and it presents an additive accounting procedure only through the *devised* values of W_i . No *substantive* sum ranking result is entailed by the accounting procedure.

It is, in fact, not surprising that the relevance of equality in moral valuation cannot be settled without assessing equality itself. The informational foundation of well-being yielded by WAIF leaves it open how the well-being of different people are to be combined in assessing states and actions. And this issue is not obliterated even when "rational" behavior under uncertainty with primordial equiprobability is taken to be the basis of moral valuation. Rationality—even in the limited sense of satisfying von Neumann-Morgenstern axioms—is no enemy of valuing equality of well-being.

While the issue of equality has to be faced, the fact that there are many different notions of equality implies that conflicts among these different egalitarian views have to be faced as well. In an important study called *Equalities*, Douglas Rae has argued that the "one idea that is more powerful than order or efficiency or freedom in resisting equality" is "equality itself."²³ Peter Westen has gone further and has suggested that equality is "an empty form having no substantive content of its own."²⁴ It may be tempting to dismiss

²² See A. B. Atkinson, "On the Measurement of Inequality," *Journal of Economic Theory*, II (1970): 244-263; P. Dasgupta, A. Sen, and D. Starrett, "Notes on the Measurement of Inequality," *Journal of Economic Theory*, VI (1973): 180-187; M. Rothschild and J. E. Stiglitz, "Some Further Results in the Measurement of Inequality," *Journal of Economic Theory*, VI (1973): 188-204.

²³ Cambridge, Mass.: Harvard, 1981; p. 150.

²⁴ "The Empty Idea of Equality," *Harvard Law Review*, VC (1982): 537-596, p. 596.

concern with equality in moral evaluation on grounds of this alleged "emptiness."

However, the multiplicity of notions of equality arises from various sources, and often enough these sources do not appear to be peculiar to equality at all. A person's advantage can be judged in many different ways, and corresponding to these different ways, there are various characterizations of equality. Utility, income, opulence, need-fulfillment, etc., provide different ways of seeing a person's success, and each of them would lead to a corresponding approach to equality. But this plurality is not special to equality. Any derivative notion for which individual advantage provides the informational base, would have this plurality. For example, the notion of "efficiency" in terms of achievements (that the situation is such that no one can be made more advantaged without reducing someone else's advantage) will have exactly the same variability. If advantage is seen in absolute terms *and* identified with utility, then the notion of efficiency immediately becomes the concept—much used in economics—of "Pareto optimality." But related to the other notions of advantage—income, opulence, need fulfillment, etc.—there will be the corresponding notions of efficiency.

In rejecting the claims of utility to represent well-being, we have outlined the need for an alternative approach. The issue of equality will again arise in whatever alternative concept of well-being we accept. I return now to the characterization of well-being.

V. FUNCTIONINGS AND WELL-BEING

It is useful to contrast two closely related ideas, which are quite clearly different. One is the idea of being "well off" and the other of being "well," or having "well-being." The former is really a concept of opulence—how rich is she? what goods and services can she buy? what offices are open to her? and so on. This refers to a person's command over things outside—including what Rawls calls "primary goods." Having "well-being," on the other hand, is not something outside her that she commands, but something in her that she achieves. What kind of a life is she leading? What does she succeed in doing and in being? Being "well off" may help, other things given, to have "well-being," but there is a distinctly personal quality in the latter absent in the former.

In looking for an appropriate approach to well-being, there are two different dangers to avoid, coming from two different directions. One problem lies in taking a basically subjectivist view in terms of one of the mental-state metrics of utility. The difficulties of this approach have already been discussed (sections II and III). Being happy may be a valuable thing to be, and not having one's

desires frustrated may also have value, but there are clearly other things that are also valuable to do or be. We have to value the different types of doings and beings, but “desiring” is not a valuational activity in itself, though it may be the consequence of one. The linkage may give desiring an “evidential” value, which was discussed, but the limitations of which, especially in interpersonal comparison, also emerged clearly enough. We are, thus, pulled in the objectivist (and, in one sense, “impersonal”) direction, looking for a criterion that does not get messed up by circumstantial contingencies (as the “evidential” value of desire information clearly can).

What form would an “objective” criterion of well-being take? “By an *objective criterion* I mean a criterion that provides a basis for appraisal of a person’s well-being which is independent of the person’s tastes and interests,” says Scanlon.²⁵ But why should an objective criterion not take note of the objective basis of differences of different people’s “tastes and interests”? As discussed in the first lecture (in the context of the quite different problem of position-relativity of moral judgments of actions and states, 183/4), the personal features that make a *relevant* objective difference (and which others too can see) can be built *parametrically* into an evaluation function without losing objectivity.

One reason why “being well off” is quite different from “being well,” despite the fact that the former is causally related to the latter, is the variability of personal characteristics which makes the causal relations *person-specific*. The Charybdis of overrigidity threatens us as much as the Scylla of subjectivist variability, and we must not lose sight of the important personal parameters in developing an approach to well-being.

Consider a disabled person, who, with the same level of real income and opulence, cannot do many things that a normal person can. His tastes in commodities will be different (and typically more expensive), and his interests will make very different demands on others. In what way does this disabled person differ in terms of this issue from the person with “expensive tastes,” who has received such a lot of attention recently (and, to indulge in an aside, whose taste for having expensive and exquisite attention must by now be well satisfied by what he has got from Kenneth Arrow, Ronald Dworkin, John Rawls, and Tim Scanlon, among others)?²⁶

²⁵ “Preference and Urgency” (1975), p. 658. Scanlon qualifies the statement in several different ways (pp. 658/9). Also, his concern is primarily with what is morally significant and not with developing a notion of well-being relevant in other contexts as well.

²⁶ Arrow, “Some Ordinalist-Utilitarian Notes on Rawls’s Theory of Justice,” this

One important difference between the two cases relates to the likely fact that the expensive-taste person may be able to *do* most things a normal person can and be handicapped only in the context of achieving contentment or satisfaction, whereas the disabled person may be handicapped in terms of a much wider range of activities.²⁷ As Scanlon describes the refined being, he could not be raised to a “normal level of satisfaction” without very high expenditure (659). The Arrow-Dworkin creature hooked on “plover’s eggs and pre-phylloxera claret” is unhappy unless he gets these extraordinary provisions, even though there is nothing to suggest that he cannot do the usual things that other people can, except that of being happy. In contrast, the deprivation of the handicapped person relates also to his inability to do many entirely normal things (unless, of course, we are considering some exotic type of handicap). The disadvantages that he suffers in his “being” do not have to be seen in terms of utilities at all.

In fact, handicapped people are often cheerful enough, and do not lead a life of frustrated desires, since their desires adapt to the reality of their handicap (a point that relates to the issue of circumstantial contingency of desires, which was discussed earlier, 190/1). The primary disability certainly would be in matters of leading a normal life—being able to get around, being able to hear or see, or something like that. That it is reflected in a lowered utility level is secondary, if true, and it may not even be true. There is no real comparison with the “plover’s egg” fan.

The primary feature of well-being can be seen in terms of how a person can “function,” taking that term in a very broad sense. I shall refer to various doings and beings that come into this assessment as *functionings*. These could be activities (like eating or reading or seeing), or states of existence or being, e.g., being well nourished, being free from malaria, not being ashamed by the poverty of

JOURNAL, LXX (1973): 245–288; R. Dworkin, “What Is Equality? Part I: Equality and Welfare,” *Philosophy and Public Affairs*, x, 3 (Summer 1981): 185–246; Rawls, “Social Unity and Primary Goods,” in Sen and Williams; Scanlon, “Preference and Urgency” (1975).

²⁷ Another difference emphasized particularly by Dworkin, Rawls, and Scanlon relates to the “responsibility” of the person with expensive taste in the genesis of his disadvantage. However, although that consideration is of obvious interest in judging a person’s *moral claim* for support, it may not be of direct relevance in judging a person’s *actual well-being*. In terms of the categorization used in these lectures, the issue of responsibility relates immediately to the agency aspect of a person, rather than to the well-being aspect. There is a further problem arising from the fact that the responsibility may have to be traced over a long period of time, and this is made more complicated by issues of personal identity; see Derek Parfit, *Reasons and Persons* (New York: Oxford, 1984), and Michael Slote, *Goods and Virtues* (New York: Oxford, 1983).

one's clothing or shoes (to go back to a question that Adam Smith discussed in his *Wealth of Nations*²⁸). I shall refer to the set of functionings a person actually achieves as the *functioning vector* (though it would have been more accurate to describe it as an *n-tuple*, since some of the functionings may not be numerically representable).

The primary feature of a person's well-being is the functioning vector that he or she achieves. The functioning vectors can be ranked and partially ordered in line with some common valuations, or in the light of what Scanlon calls "urgency," or some other acceptable criterion. This, then, is the secondary representation, in the form of the place of a functioning vector in a partial order of such vectors. If the ordering is complete, we should be able to get even a real-numbered "index" of well-being (in the absence of some technical peculiarities in that ordering²⁹). In fact, even a partial order will typically have numerical representation with some "transparency" properties.³⁰ The "natural" form of well-being ranking is indeed that of a partial, incomplete order. It would be just as extraordinary if every possible pair of functioning vectors could be compared in terms of over-all well-being, as it would be if none of them could be. There will, of course, be many agreed valuations, and many decisive judgments, and the clarity of these cases is not compromised by the muddiness of others.

VI. OPULENCE, FUNCTIONINGS, AND UTILITY

The contrast between opulence and functioning achievements is simple enough, but its significance in the context of well-being may be worth spelling out. The example of the crippled (or otherwise handicapped person) is, in some respects, particularly misleading. It may suggest that personal variations are uncommon and unusual things. In fact, however, interpersonal variations in "transforming" goods into functionings is extremely common. Take, for example, the consumption of *food*, on the one hand, and the functioning of *being well nourished*, on the other. The relationship between them

²⁸ *An Inquiry into the Nature and Causes of the Wealth of Nations* (1776; Everyman's Library, London: Dent, 1954), vol. II, p. 352.

²⁹ An example of a non-numerically-representable complete order is a "lexicographic" order over the unit square: see Gerard Debreu, *Theory of Value* (New York: Wiley, 1959), ch. 4. Since lexicographic orderings have some interest in moral philosophy, as with the "lexicographic maximin" version of Rawls's difference principle [see my *Collective Choice and Social Welfare* (1970), p. 138; Rawls, *A Theory of Justice* (1971), pp. 42-45], it is perhaps just worth recording the elementary fact that a lexicographic order—or indeed any other order—over any *finite* set (no matter how large) will never have such a representation problem.

³⁰ See Mukul Majumdar and Amartya Sen, "On Representing Partial Orderings," *Review of Economic Studies*, XLIII (1976): 543-545. The real problem arises with guaranteeing the continuity of such a numerical function.

varies with (1) metabolic rates, (2) body size, (3) age, (4) sex (and if a woman, whether pregnant or lactating), (5) activity levels, (6) climatic conditions, (7) presence of parasitic diseases, (8) access to medical services, (9) nutritional knowledge, and other influences.

Interpersonal variability of the relation between goods and functionings turns out to be quite central to many important policy issues. For example, the comparison between the positions of men and women in terms of nutritional achievements in the poorer countries is often distorted by the use of food intake figures, which are not only unreliable, but also difficult to relate to well-being. The comparisons are usually made with some arbitrarily assumed sex-specific nutritional requirements, and there is some evidence that the most widely used requirements norms (the so-called "FAO-WHO standards") not only are inappropriate in a great majority of cases, but also involve a systematic sex bias, helping to conceal the extent of the relative deprivation of women in many of these countries. In contrast, direct information on the actual *functionings* achieved by men and women (in this case, avoiding undernourishment, escaping avoidable morbidity, etc.) are both easier to use and more sensibly interpreted in terms of well-being.³¹

Interpersonal variations are not, of course, confined to such elementary functionings as being well nourished, or escaping avoidable morbidity. Personal and social influences on the conversion of opulence into functioning achievements can be powerful in many other fields as well. The resource requirements of such functionings as "appearing in public without shame" (emphasized by Adam Smith), taking part in the life of the community, being able to visit and entertain friends, being acceptably well informed, and so on, vary a good deal with the nature of the community in which one lives and with other contingent circumstances.³²

Opulence does, of course, *contribute* to functioning, and the command that it gives can indeed be important for a person's well-being and advantage. Rawlsian analysis of "primary goods" has forcefully demonstrated the usefulness of that perspective compared with reliance on the mental metric of utility.³³ But an index of opu-

³¹ Some of these issues are discussed, along with presenting some empirical results, in Jocelyn Kynch and Amartya Sen, "Indian Women: Well-being and Survival," *Cambridge Journal of Economics*, vii (1983): 363-380, and in Essays 13, 15, and 16 in my *Resources, Values and Development* (Oxford: Blackwell, and Cambridge, Mass.: Harvard, 1984), hereafter, RVD.

³² Some of these issues are discussed in my "Poor, Relatively Speaking," *Oxford Economic Papers*, xxxv (1983): 153-169, reprinted in RVD; and "The Living Standard," *Oxford Economic Papers*, xxxvi (1984): 74-90.

³³ Rawls, *A Theory of Justice* (1971), and "Social Unity and Primary Goods" in Sen and Williams. Rawls is concerned with agency as well, not just well-being, in his analysis of primary goods.

lence, or of primary goods, cannot be seen as an index of well-being as such. Nor can the "advantage" of different persons (or groups) be ranked on the basis of the primary goods index, because of interpersonal variations in what people can *do* with primary goods, depending on biological, social, and other parameters.

Nor, of course, is it a matter of utility, as we have already seen. Happiness may be a valuable part of well-being, and being happy is certainly a momentous functioning, but there are other important functionings that are not congruent with happiness and for which happiness cannot serve as a surrogate, especially in interpersonal comparisons. Desire intensities may also be relevant to the valuation of functionings for evidential reasons, but—as was discussed earlier—the problem of valuation of functionings cannot be avoided by concentrating instead on the observation of desires and their intensities. The central feature of well-being is the ability to achieve valuable functionings.³⁴ The need for identification and valuation of the important functionings cannot be avoided by looking at something else, such as happiness, desire fulfillment, opulence, or command over primary goods.

It is, of course, clear that the functioning approach is intrinsically information-pluralist.³⁵ Although some consensus of valuations and the usability of various objective criteria may take us a great distance in establishing an extensive partial ordering in the comparison of well-being, it would be quite wrong to expect anything like the kind of complete ordering that utilitarians have made us prone to demand. As was discussed in the first lecture, such a completeness requirement is neither methodologically sensible nor substantively plausible (nor in fact actually met by the utilitarian calculus itself, except under the artificial assumption of complete and unproblematic interpersonal comparability). Incompleteness is *not* an embarrassment.

VII. CAPABILITY AND FREEDOM

The information pluralism of the functioning approach to well-being has to be further extended if we shift attention from the person's actual functionings to his or her *capability* to function. A

³⁴ The problem of evaluation of functioning vectors has been discussed in my *Commodities and Capabilities* (Amsterdam: North-Holland, 1985).

³⁵ The feature of plurality can, however, be reduced by the use of "invariance restrictions" in the analysis of well-being, e.g., "being well nourished" defining an isoinformation set (ignoring differences between being well-nourished through eating bread or rice, etc.), and "being entertained" as another such category (ignoring differences between entertainment from the cinema or the theatre, etc.). There will, of course, be gradations in being well-nourished, being entertained, etc., with corresponding layers of isoinformation sets.

person's *capability set* can be defined as the set of functioning vectors within his or her reach. In examining the well-being aspect of a person, attention can legitimately be paid to the capability set of the person and not just to the chosen functioning vector. This has the effect of taking note of the positive freedoms in a general sense (the freedom "to do this," or "to be that") that a person has.

Importance may well be attached to checking whether one person did have the opportunity of achieving the functioning vector that another actually achieved. This involves comparison of actual opportunities that different persons have. If one person could have achieved all the relevant functioning vectors that the other could, then in some important sense the first person had at least as much freedom to live well. The general idea of the freedom to achieve well-being can be called *well-being freedom*.

To illustrate, consider two persons with identical actual functioning vectors, including—as it happens—both of them starving. Person *A* is starving because she is very poor and lacks the means to command food. Person *B* is starving out of choice, because of his religious beliefs, which have made him decide to starve and undergo the consequent suffering. In terms of the misery caused by the starvation, we learn that there is no difference between *A*'s experience and *B*'s. Even if it were plausible to say that *A* and *B* both have the same level of well-being, in terms of being undernourished, miserable, etc. (that complex issue will be examined presently), there would nevertheless remain an important difference between the two cases, viz., *B could have* in a straightforward sense, chosen an alternative life style which *A* could not have chosen. This is a difference that is relevant to assessing the well-being aspect of the two persons respectively, though it is not necessarily a difference between the *actual* well-being levels achieved by the two. In judging a person's advantage (and the kind of "deal" that he or she has got), the importance of well-being freedom must be recognized.³⁶

A second reason for being interested in the capability set arises from a different type of consideration altogether. The existence of genuine choice may actually affect the nature and significance of the functionings achieved. The first—and primitive—characterization of functionings in terms of doing this or being that may have to be modified and made more sophisticated by taking note of more

³⁶The problem of ranking capability sets involves many difficulties; on this see my *Commodities and Capabilities* (1985). Cases of dominance (either in the form of one set being a proper subset of another, or in the sense of one set being *element-by-element* dominated by the other in terms of the partial ordering of functioning vectors) would of course induce a *minimal* partial ordering on the space of capability sets.

complexly characterized doings, e.g., choosing to do x rather than y . *Fasting* involves an abstention from eating that goes beyond just starving, and it can be argued that in the example being considered, rich and religious B could choose to fast, whereas poverty-stricken A could not. The possibility of starving was open to both, but *fasting* by choice was open only to B .³⁷ Examining capability sets of primitive functionings may permit us to move to a richer analysis of functionings, taking note of the choices exercised.

It can be argued that the well-being that a person actually enjoys is often more closely related to such *refined* functioning achievements. This view relates to the idea that the good life is *inter alia* also a life of freedom.³⁸ It is a view that has found rather eloquent expression in the writings of many social and political philosophers.³⁹

On this line of reasoning, the extent of the capability set is relevant to the significance and value of the respective functionings. This might look like introducing a circularity in the relationship between functionings and capabilities, and between well-being and well-being freedom. But what it, in fact, does is to force us to see these concepts as mutually dependent, taking note of the simultaneity of the relationships involved. The refined characterization of functionings has to take note of the capability set of the "primitive" functionings, but the refined functionings, in their turn, yield a (corresponding) refined capability set.⁴⁰

VIII. CONCLUDING REMARKS

In this lecture I have argued (1) *against* accepting the adequacy of WAIF (well-being as informational foundation), (2) *against* the

³⁷ I am grateful to Ian White for helpful discussions on these issues.

³⁸ On related matters, see Jon Elster, "Sour Grapes," in Sen and Williams, pp. 227/8, and *Sour Grapes* (New York: Cambridge, pp. 128/9).

³⁹ One of them is, of course, Karl Marx, whose political philosophy gives a special role to bringing "the conditions for the free development and activity of individuals under their own control" and who outlines a vision of a liberated future society which "makes it possible for me to do one thing to-day and another tomorrow, to hunt in the morning, fish in the afternoon, rear cattle in the evening, criticize after dinner, just as I have in mind, without ever becoming hunter, fisherman, shepherd or critic" [K. Marx and F. Engels, *The German Ideology* (1845/46; republished New York: International Publishers, 1947), p. 22]. This perspective of freedom has a strong bearing also on other aspects of the Marxian tradition. Examples of different types of connections can be found in G. A. Cohen, *Karl Marx's Theory of History: A Defence* (New York: Oxford, 1978); J. E. Roemer, *A General Theory of Exploitation and Class* (Cambridge, Mass.: Harvard, 1982); G. G. Brenkert, *Marx's Ethics of Freedom* (London: Routledge & Kegan Paul, 1983); J. Elster, *Making Sense of Marx* (Cambridge: Cambridge, 1985); S. Lukes, *Marxism and Morality* (Oxford: Oxford, 1985).

⁴⁰ It is, of course, possible to do more than one round of refining, though the advantages of such extended refinement might be rather limited.

identification of well-being with utility (in any of its forms: happiness, desire fulfillment, choice), and (3) *against* treating sum ranking as trivial and equality as having no independent force. More positively, I have argued *for* seeing well-being in terms of functioning vectors and the capability to achieve them. The well-being aspect of a person leads to a particular concept of freedom, which has been called *well-being freedom*.

FREEDOM AND AGENCY

IN the last lecture it was argued that the well-being aspect of a person calls for dual accounting—in terms, respectively, of freedom and achievement. The significance of the distinction relates particularly to the agency aspect of a person. Values other than the pursuit of well-being may figure prominently in a person's assessment of choices. Not making maximal use of the freedom to pursue well-being may not reflect any failure of evaluation or action. In judging what kind of a deal a person has in terms of the opportunity to pursue his or her own well-being, the concept of *well-being freedom* is, therefore, necessary, and its role cannot be subsumed by the observation of actual *achievements* of well-being. The autonomy of the agency aspect of a person, which has the effect of making WAIF an inadequate basis for substantive moral reasoning, also has the result that the well-being aspect of a person must be seen in terms of freedom as well as actual achievement.

I. AGENCY FREEDOM

Well-being freedom is freedom of a rather particular type. It concentrates on a person's capability to have various functioning vectors and to enjoy the corresponding well-being achievements. This concept of freedom, based on the well-being aspect of a person, has to be clearly distinguished from a broader concept of freedom, related to the agency aspect of a person. A person's "agency freedom" refers to what the person is free to do and achieve in pursuit of whatever goals or values he or she regards as important. A person's agency aspect cannot be understood without taking note of his or her aims, objectives, allegiances, obligations, and—in a broad sense—the person's conception of the good. Whereas well-being freedom is freedom to achieve something in particular, viz., well-being, the idea of agency freedom is more general, since it is not tied to any one type of aim. Agency freedom is freedom to achieve whatever the

person, as a responsible agent, decides he or she should achieve. That *open conditionality* makes the nature of agency freedom quite different from that of well-being freedom, which concentrates on a particular type of objective and judges opportunities correspondingly.¹

This open conditionality does not imply that the person's view of his agency has no need for discipline, and that anything that appeals to him must, for that reason, come into the accounting of his agency freedom. The need for careful assessment of aims, objectives, allegiances, etc., and of the conception of the good, may be important and exacting. But despite this need for discipline, the use of one's agency is, in an important sense, a matter for oneself to judge. When the requirement of careful assessment cannot be fulfilled (e.g., in the case of young children, or with persons mentally ill in ways that rule out such assessment), the agency aspect will be, obviously, less important. This does not indicate that the demands of the agency aspect for such a person would have to be assessed by *others* on his or her behalf, but only that no such demands would then be made. This would not, of course, in any way compromise the importance of their *well-being aspect*. Indeed in the absence of the relevance of their agency aspect, it is their *well-being achievement* that would uniquely command attention.

The importance of the agency aspect, in general, relates to the view of persons as responsible agents. Persons must enter the moral accounting by others not only as people whose well-being demands concern, but also as people whose responsible agency must be recognized. Although well-being achievement is geared to a specific objective, and well-being freedom is seen in terms of the freedom to achieve that objective through the choice of functioning vectors, agency freedom cannot be examined in terms of any prespecified objective. The open conditionality referred to above is central to that concept.

II. PLURALITY, WELL-BEING, AND AGENCY

In terms of the analysis presented so far, we have a plural informational base for substantive moral analysis, which has to take note of several distinct considerations involving well-being and agency. These different considerations may quite possibly pull in different directions.

A person's actual well-being has intrinsic importance. This tends to give WAIF some apparent plausibility, and is reflected more

¹ Agency freedom has claims to being taken as freedom *tout court*, which well-being freedom clearly does not have.

generally in the hold of “well-being” on moral thinking (including utilitarian moral thinking, though—as was argued in the last lecture—utility provides rather limited ways of seeing well-being). For responsible adults, the concept of “well-being freedom” has a clear relevance in judging the opportunities a person has for pursuing his or her own advantage. The supplementation of well-being by well-being freedom, in the case of responsible adults, involves a refinement in the assessment of the well-being aspect of a person. But well-being freedom is only a specific type of freedom, and it cannot reflect the person’s over-all freedom as an agent; we have to turn to the notion of agency freedom in that context. It is hard to see how any part of this plurality (involving both the well-being aspect and the agency aspect of persons) can fail to have some intrinsic importance.

This plurality would have been a matter of considerable embarrassment if for methodological reasons we were obliged to reject any information-pluralist framework, with “monism” as the only acceptable approach. But, as was discussed in the first lecture, there is no good reason for that monist presumption, and none for imposing monism as an a priori requirement.

Furthermore, as was also discussed in the first lecture, it is not a matter of embarrassment that there might be scope for variation in the assessment of the relative importance of the different elements in a pluralist structure. Although “dominance” reasoning will, as was discussed, be able to provide a minimal partial ordering in a pluralist structure, and variable valuations within given *ranges* of weights will extend that partial ordering further, nevertheless even such enhanced partial orderings may well remain incomplete. The need to consider both the agency aspect and the well-being aspect of persons is not compromised by the possibility that the combined structure may leave a large subset of unordered pairs. Those who see substantive moral reasoning as worthless unless it completely orders the space of all actions and states will, of course, find much to complain about in this incompleteness. But, as was seen in the first lecture, there is no good reason why we should have to yield to this external pressure for completeness.

It can, however, be argued that while plurality and incompleteness are not themselves objectionable, nevertheless in this particular case there happen to be redundant elements in the proposed plural structure. I have already discussed, in the second lecture, why well-being is inadequate as an informational basis of moral judgments, given the intrinsic importance of agency. But I have not yet explicitly provided any reason why agency information may not be ade-

quate for moral judgment (thereby making it redundant to incorporate well-being information—in the form of achievement *or* freedom—in the basic structure). Agency information includes each person's pursuits and choices. The hypothesis of the adequacy of agency information might appear to have considerable plausibility,² especially since agency freedom can be seen as incorporating well-being freedom, *inter alia*. "Open conditionality" indicates that if a person were to pursue well-being, then assessment of the agency freedom of that person would have to include his or her well-being freedom as well. And if the person were unconcerned about pursuing his or her own well-being, then why should it matter whether the person had well-being freedom or not?

In responding to this issue, the important thing to recognize is that the well-being aspect and the agency aspect of persons have dissimilar roles in moral accounting. They invite attention in *disparate* ways. At the risk of oversimplification it can be said that the well-being aspect of a person is important in assessing a person's *advantage*, whereas the agency aspect is important in assessing what a person can do in line with his or her conception of *the good*. The ability to do more good need not be to the person's advantage.

Perhaps an illustration may help to bring out the distinction a little more clearly. You are enjoying eating a sandwich on a nice spring day, sitting on the bank of the river Avon. Meanwhile, far away, a man who cannot swim is drowning. You cannot do anything to save him—he is hundreds of miles away. Consider now a *counterfactual* situation in which that man, instead of drowning at that lonely, distant spot, is about to drown right in front of you. How are *you* affected by this counterfactual change? Your agency freedom is, in an important way, increased. You can now very likely save the man. You will, let us assume, value this opportunity highly and decide to save him, and, further, you will even succeed. You will chuck your sandwich, jump into the cold river, and haul the man out. Your agency freedom would have been well used, and the state of affairs would have been better, as you would yourself judge it.

But your *well-being* is, quite possibly, reduced. You may, of course, enjoy saving a life, and a part of your well-being may certainly be positively influenced by the opportunity to do good. But it is possible that, everything considered, your own well-being is no

² I am grateful to Bruce Ackerman for providing good arguments in favor of this possibility and for insisting that *either* agency freedom be regarded as adequate *or* its inadequacy be explicitly established.

higher, and let us assume (to pursue our contrasts) that it is actually lower—with the quiet lunch being replaced by a ducking in the chilling river, risking your own life in undertaking the life-saving operation. An expansion of agency freedom can go with a reduction in actual well-being, through your own choice, even though you are not, by any means, indifferent to your own well-being.

What about your “well-being freedom?” That was certainly not enhanced by the counterfactual change. The additional opportunity of saving the drowning person did not give you a better (or even as good a) way of pursuing your own well-being. In fact, in an important way, the counterfactual change actually *reduced* your well-being freedom. You were no longer free to eat your sandwich *without anxiety*, and, in terms of the “refined” functionings discussed in the last lecture (pp. 201/2), there was a genuine loss of opportunity to pursue your own well-being.

Since agency freedom includes *inter alia* the freedom to pursue your own well-being, it can, of course, also be said that *in that respect* your agency freedom too would have been adversely affected. But if it is the case (as it is here assumed to be) that, given the choice, you would rather have the opportunity of saving the drowning person than eat your sandwich without anxiety, and that your valuations as an agent confirm the correctness of this choice, then there is clearly a *net* gain in terms of your *agency freedom*. As an agent you value the new opportunity (saving the person) more than the lost opportunity (eating the sandwich without anxiety). But in terms of your *well-being freedom*, there is a net loss, since the lost opportunity is more valuable from the narrow perspective of your own well-being than the new opportunity. It is, in this sense, arguable that even though agency freedom does “include” well-being freedom, the latter can quite possibly go down when the former goes up (and, of course, vice versa). The ranking of alternative opportunities from the point of view of agency need not be the same as the ranking in terms of well-being, and thus the judgments of agency freedom and well-being freedom can move in contrary directions.³ So, even though agency freedom is “broader” than well-being freedom, the former *cannot* subsume the latter.

Insofar as each person’s advantage commands attention and respect in moral accounting, the well-being aspect of the person has to be directly considered. This role cannot be taken over by agency

³The relationship between the ranking of *elements* and the ranking of *sets* of elements is discussed in my *Commodities and Capabilities* (1985), ch. 7, in the specific context of judging “capability sets” and the corresponding “freedoms.”

information. Although the agency aspect and the well-being aspect both are important, they are important for quite different reasons. In one perspective, a person is seen as a doer and a judge, whereas in the other the same person is seen as a beneficiary whose interests and advantages have to be considered.⁴ There is no way of reducing this plural-information base into a monist one without losing something of importance.⁵

The well-being aspect may be particularly important in some specific contexts, e.g., in making public provisions for social security, or in planning for the fulfillment of basic needs.⁶ In judging what a person may expect from social arrangements, the demands of well-being (and, in the case of responsible adults, also of well-being freedom) may loom rather large. On the other hand, in many issues of personal morality, the agency aspect, and one's responsibility to others, may be central. The well-being aspect and the agency aspect both demand attention, but they do so in different ways, and with varying relevance to different problems.

III. FREEDOM: POWER OR CONTROL?

Within the idea of freedom itself, there tends to be an internal plurality. There are, in fact, two distinct elements in the idea of freedom (they may be called, respectively, "power" and "control"), and the assessment of freedom has to deal with this internal plurality as well.⁷ The distinction is of special relevance to particular issues of freedom, such as liberty and autonomy.

A person's freedom may well be assessed in terms of the power to achieve chosen results: whether the person is free to achieve one outcome or another; whether his or her choices will be respected and the corresponding things will happen. This element of freedom, which I shall call *effective power*, or *power* (for short), is not really concerned with the mechanism and procedures of control. It does not matter for effective power precisely *how* the choices are

⁴ This dichotomy has a close connection with the contrast between "interest aggregation" and "judgment aggregation" in social choice theory, discussed in my "Social Choice Theory: A Reexamination," *Econometrica*, XLV (1977): 53-89; reprinted in my CWM (1982).

⁵ However, while the agency and well-being aspects are quite distinct from each other, they are *not* independent of each other. The dependence arises from at least two obvious facts: that a person's agency may affect his or her well-being, and that being well may help a person's ability to act in pursuit of other objectives as well.

⁶ This applies both to formal modeling of plan priorities [see, for example, G. Chichilnisky, "Basic Needs and Global Models: Resources, Trade and Distribution," *Alternatives*, vi (1980)] and to broader planning "strategies" [see, for example, P. Streeten, *et al.*, *First Things First* (New York: Oxford, 1981)].

⁷ The contrast is discussed in my "Liberty as Control: An Appraisal," *Midwest Studies in Philosophy*, vii (1982): 207-221.

“executed.” Indeed the choices may not even be directly addressed. Effective power can take note of *counterfactual choice*: things might be done because of knowledge of what the person would choose if he actually had control over the outcome.

In contrast, a person’s freedom may also be assessed in terms of whether the person is himself exercising *control* over the process of choice. Is he actively doing the choosing in the procedure of decision and execution? This element of freedom may be called *procedural control*, or *control* (for short). It does not matter for freedom as control whether the person succeeds at all in achieving what he would choose.⁸

It is fair to say that, in the literature on liberty and freedom in political philosophy, it is the control element that has tended to receive the lion’s share of attention. There is, of course, the related but distinct question of whether freedom should be seen in positive or negative terms, and whether one’s obligations should take the form of not interfering with other people’s control over spheres of their lives rather than the form of duties to help positively. I shall take up these questions later, but the distinction between effective power and procedural control is, in a sense, more elementary. While the control view is often combined with a constraint-based approach (to wit: *do not interfere* with other people’s control over some specific matters), there is a general issue of whether the control view is *itself* an adequate approach to freedom, even when its use is not restricted to the negative perspective. I would argue that this concentration on control tends to produce an unacceptably narrow conceptualization of liberty and freedom.

To illustrate the contrast, I begin with a shortened version of an example I have used elsewhere.⁹ Your friend is injured in an accident and is unconscious. The doctor says that either treatment *A* or treatment *B* can be used and one would be just as effective as the other, but that your friend would suffer less from *A* because of its

⁸ In commenting on an earlier version of this text (and on the paper cited in fn 7), G. A. Cohen has argued that the terms ‘power’ and ‘control’ may be misleading. When a person has “effective power” in the sense specified, even if he does not have control over the process of selection, it *could be said* (Cohen points out) that he *had*, in fact, exercised control. This point might well be correct, but, if the basic distinction is accepted, it may not matter much whether we see it as a contrast between “procedural control” and “control” in some suitably broad sense, or between “procedural control” and “effective power”, as it is presented here. The important issue relates to the fact that in modern society many powers are indirectly exercised, and, in that context, concentration on the levers of control provides an inadequate approach to freedom.

⁹ “Liberty and Social Choice,” (1983), pp. 18–20.

smaller side effects. However, you happen to know that your friend would have chosen treatment *B*, since treatment *A* is associated with some experiments on live animals of which your friend disapproves totally. He would, in fact, agree that treatment *A* would have been better for his well-being, but as a free agent he would have nevertheless chosen treatment *B*, if he were given the choice. If you now decide to ask the doctor to give your friend treatment *B*, you are giving your friend effective power even though he is not exercising the control himself. And your ground for choosing treatment *B* for him is surely not his well-being (he too agrees that *A* would be better for that), but his freedom (in the general form of *agency freedom*). It is also worth noting that your decision regarding what treatment your friend gets is *dependent* on what your friend would himself have chosen. It is not only that he happens to be getting what he would have chosen; he is getting it *because* he would have chosen it.

The contrast between effective power and procedural control is important in practice. It is often not possible to organize society in such a way that people can directly exercise the levers that control all the important aspects of their personal lives. To try to see freedom exclusively in terms of control is to miss out on the demands of freedom when control cannot feasibly be exercised by the persons themselves.

If, for example, we want the freedom not to be mugged in the streets and the liberty to move about without being so treated, we are looking for effective power. Precisely who exercises the control may be less important than the ability to achieve what we would have chosen. If the streets are cleared of the muggers *because* we would choose not to be mugged,¹⁰ our freedom is being well served, even though we have not been given control over the choice of whether to be mugged or not.

To avoid possible misunderstandings, it may be useful to clarify what is being asserted here, and what is not. First, it is, of course, typically the case that not being mugged also serves our well-being. But in asserting that our freedom—even *agency freedom*—is well served by the streets being cleared of the muggers, something more is being asserted. It is that we would have *chosen* not to be mugged, and the streets are being cleared precisely because that is what we would have chosen.

Second, with the streets cleared of the muggers, we have, of

¹⁰ While the dependence of the choice of antimugging policy on what we would choose is important for the “power” view of freedom, it is not necessary for that approach that this be achieved though some formal *procedure* of referendum-based execution.

course, the opportunity to walk as we please. This reflects the positive freedom, in most cases, to walk or not. But in that part of the story there is *no* contrast between control and power. We have the power to walk and also control over *that* decision. The contrast arises in considering the alternatives of being mugged and not being mugged. We do not control *this* choice ourselves, and we do not have the option of "being mugged" once the streets are cleared. But though that control is not ours, it was exercised in line with what we would have chosen and because of it.¹¹ It is in this sense that we have effective power, though no procedural control.

Third, in the example of mugging, the power is enjoyed not by each individual separately, but by the individuals taken together. There are other examples in which the individual on his own has effective power without control, e.g., in the example of your friend getting the treatment he would have chosen. Given the interdependencies of social living, many liberties are not separately exercisable,¹² and effective power may have to be seen in terms of what all, or nearly all, members of the group would have chosen, e.g., in this case *not* to be mugged. What Isaiah Berlin calls "a man's, or a people's, liberty to choose to live as they desire"¹³ has implications for effective power of individuals taken together *as well as* separately.

Concentrating on the control view of freedom and liberty has a long history. But in recent years it has tended to be particularly emphasized in the context of criticizing the practice in social choice theory of seeing freedom and liberty in terms of effective power only.¹⁴ Robert Nozick, Peter Bernholz, Peter Gärdenfors, Robert

¹¹ There can, however, be *directional asymmetries* in effective power. For example, although the choice of "not being mugged" may be respected and executed, that of "being mugged" may not be similarly guaranteed. The real way to check whether we are getting *x* because we would have chosen *x*, is not so much to examine whether we would invariably have got "not *x*" if that is what we would have chosen, but to ascertain whether *x* is guaranteed in the first case but not in the second. Many standard liberties do, of course, have directional asymmetries; e.g., Miss *A* has the liberty *not to marry* Mr. *B* no matter what he wants, but not the liberty *to marry* Mr. *B* irrespective of his decisions.

¹² The issue of interdependence is well discussed by Christian Seidl, "On Liberal Values," *Zeitschrift für Nationalökonomie*, xxxv (1975): 257-292, and Jonathan Barnes, "Freedom, Rationality and Paradox," *Canadian Journal of Philosophy*, x (1980): 545-565.

¹³ Isaiah Berlin, *Four Essays on Liberty* (New York: Oxford, 1969), p. 170.

¹⁴ See my CCSW, ch. 6; and Allan Gibbard, "A Pareto-consistent Libertarian Claim," *Journal of Economic Theory*, vii (1974): 388-410. The "social choice" literature on this subject is now quite vast. For references, see the bibliographies in W. Gaertner and L. Krüger, "Self-supporting Preferences and Individual Rights: The Possibility of Paretian Libertarianism," *Economica*, xlvii (1981): 241-252; Sen, "Liberty and Social Choice" (1983); K. Suzumura, *Rational Choice, Collective Decisions and Social Welfare* (New York: Cambridge, 1983); and J. Wrigglesworth, *Libertarian Conflicts in Social Choice* (New York: Cambridge, forthcoming).

Sugden, Brian Barry, and others have pursued the control view, and though their formulations have differed from one another, there seems to have been solid agreement on the folly of seeing liberty in terms of power rather than procedural control.¹⁵ It is not my intention to deny that the control element of freedom is indeed important in many contexts. But it is not the only important element in freedom. The power element cannot be neglected in any adequate formulation of freedom or liberty.

One may, of course, attach importance to having control over certain mechanisms of decision making and procedures of choice. In matters of what to eat, what to drink, what to wear, etc., it may not be adequate just to get what one would have chosen; one must actually do the choosing oneself. However, it is, in fact, also possible to argue that if achievements are characterized in a more articulate and apposite way than they are in the conventional formulations, then the motivation underlying any intrinsic concern with the control element of freedom can be *subsumed* within the notion of the power to achieve results. This question relates to the issue of the contents of states of affairs, which was discussed in the first lecture (181–183). Since states of affairs include actions and since agency information is part of the information regarding the states, the power to achieve particular states of affairs may have to include the power to use specific controls as well. Thus, the power view can be broadened to make room for attaching importance to control in the assessment of outcomes. The evaluation of the power to achieve different states of affairs may be control-sensitive, and the power view of freedom can, *inter alia*, accommodate whatever is important in the control view.

IV. AGENCY AND CONSEQUENTIALISM

Taking note of actions, agency, and control in the valuation of states of affairs raises an important question regarding the status of

¹⁵ Nozick, *Anarchy, State and Utopia* (1974); Peter Bernholz, "Is a Paretian Liberal Really Impossible?" *Public Choice*, xix (1974): 99–107; Peter Gärdenfors, "Rights, Games and Social Choice," *Nous*, xv (1981): 341–356; Robert Sugden, *The Political Economy of Public Choice* (Oxford: Martin Robertson, 1981); Brian Barry, "Lady Chatterley's Lover and Doctor Fischer's Bomb Party: Liberalism, Pareto Optimality and the Problem of Objectionable Preferences," mimeographed, 1982, to be published in J. Elster and A. Hylland, eds., *op. cit.*, forthcoming. A somewhat similar view is taken by Allan Gibbard, "Rights and the Theory of Social Choice," in *Logic, Methodology and Philosophy of Science*, vi (Amsterdam: North-Holland, 1982), in which volume contrast the analysis of Peter Hammond, "Liberalism, Independent Rights and the Pareto Principle." See also Isaac Levi, "Liberty and Welfare," and Peter Hammond, "Utilitarianism, Uncertainty and Information," in Sen and Williams, and Peter Bernholz, "A General Social Dilemma: Profitable Exchange and Intransitive Group Preferences," *Zeitschrift für Nationalökonomie*, xi (1980): 1–23.

that influence. If the moral approach in question incorporates “welfarism” (as the standard versions of utilitarianism do), then the value that can be attached to actions, agency, and such, must be entirely derivative. They must be based on their influence on utilities, since welfarism insists that the valuation of states must be a function of the utility information regarding the respective states. In the first lecture, this derivative view was disputed (sections VI and VII). In particular, arguments were presented against the “invariance restriction” which insists on *positional invariance*, i.e., the requirement that within a given moral approach the valuation of a state must be independent of the evaluator’s position vis-à-vis the state, including his agency.

Positional invariance raises some difficult issues. It may look superficially like a condition of elementary consistency. How can two persons, sharing exactly the same moral approach, value the same state differently? The issue of consistency can be dealt with in formal terms easily enough. The valuation of the state may take note of such positional information as the agency of persons involved in bringing about that state, and the influence of that agency can vary *parametrically* depending on whether the evaluator is one of the agents involved (and if so, what has been his specific role). The real issue is not formal consistency, but cogency.¹⁶ *Why* should the evaluator’s own agency make a difference to the valuation made by him?

The issue of cogency relates to the recognition (discussed in the first lecture) that action and agency are integral parts of the state, and furthermore, that agency can be an important aspect of the person (a question also discussed in the second lecture). It would be extraordinary to require that a person himself must take no more note of his own agency in evaluating a state of affairs than others should and that a morally relevant assessment of the state can be made in this oddly detached way. When the evaluator happens to be also the agent whose action helped to bring about the state in question and when the action for which the person is peculiarly responsible is an integral part of that state, it is not at all clear why the evaluator must ignore this important connection in assessing the state, for moral reasoning. Lady Macbeth may well have got a little carried away in the state of affairs resulting from the murder

¹⁶ I have argued elsewhere that “consistency conditions” used in the literature of rational choice and decision theory often make it difficult to see the real issues involved; and that there are no “natural” conditions—universal in scope—of “internal consistency of choice” (“Consistency,” Presidential Address to the Econometric Society, 1984; to be published in *Econometrica*).

she committed when she asked, "Will these hands ne'er be clean?", but on the other hand we might mislead her a little if we were to inform her that, in the post-murder state, the perfumes of Arabia *would* sweeten her "little hand" as easily as any other hand. The responsibility of agency surely extends to evaluating states including the evaluator's own role, and that responsibility cannot suddenly disappear in the morally relevant evaluation of the state.

One reason why agency invariance may appeal, is a tendency to confuse this type of invariance with "authorship invariance," which was discussed in the first lecture. Authorship invariance does have some claims to being a reasonable requirement of moral evaluation (and, as was argued in the first lecture, can also be seen as a requirement of objectivity of moral beliefs). This is because the authorship of a statement is a very contingent fact. *A* can easily consider what he should say if he were in fact in *B*'s position rather than in his own. But the *positions* themselves are not without significance in evaluating the states, and this position relativity can be combined with authorship invariance and indeed objectivity of the corresponding type. The general thesis was presented in a broad way in the first lecture, but the gravity of the agency aspect of a person (further pursued later) has the implication of underlining the substantive importance of the position of being an agent, without in any way affecting the case for authorship invariance. The agency position has to be parameterized in the evaluation of states in a way that authorship need not be.

There is an interesting question whether these considerations require that the much-discussed condition of "consequentialism" should also be rejected. I don't believe that the outlined considerations entail either a rejection or an endorsement of consequentialism. The introduction of position relativity does have the effect of relaxing the overtight informational constraints imposed by the traditional combination of welfarism and consequentialism, and in this sense, it does make room for less blind judgments under consequentialist evaluation. But it does not require either the acceptance or the rejection of consequentialism as such.

Take Bernard Williams's famous example of Jim having to consider killing one person himself to save nine lives; the one to be killed by him would be killed anyway.¹⁷ Whereas others have a straightforward reason to rejoice if Jim goes ahead, Jim has no option but to take serious note of his own responsibility in that state

¹⁷ "A Critique of Utilitarianism," in J. J. C. Smart and B. Williams, *Utilitarianism: For and Against* (New York: Cambridge, 1973).

and his agency in killing someone himself. Position relativity will certainly be adequate to bring in a deep asymmetry between Jim's position and those of others even in consequential evaluation.

Nevertheless, even with this modification, Jim's decision may be ultimately straightforward in a consequentialist framework. In the state that will result from his killing one person, nine people are saved, and the one who is killed would have been killed anyway. Jim would have to attach an extraordinary degree of importance to the position of not being a killer himself to come to the conclusion that the value difference between one person being killed by Jim rather than by someone else is enough to outweigh the gain of nine people being saved from sure death. Even with position relativity, the decision is fairly straightforward. If, in fact, Jim's decision problem is *not* straightforward, as Williams argues it is not, then the difference must rest somewhere else.

It is possible to argue that there is a difference between an *action* being right and *the state resulting from that action* being the best feasible state, after taking note of the value or disvalue of the action itself (including agency). The first judgment concerns an *action*, and the second a *state* (including *inter alia* that action). There is no *necessity* that the two must correspond to each other in the rigid way that consequentialism would demand.¹⁸ Whether they should, in fact, be linked in this way or not is a matter for further moral analysis, and the need to examine this link remains even after position relativity has been incorporated in the evaluation of states of affairs. Jim could still ask: "Can it be right for me to kill this man, just because the state of affairs resulting from that would be better (even after taking note of the disvalue of my action, judged as a *part* of the state of affairs)?"

The question of consequentialism will have to be settled by other arguments, since this one leaves that issue open. It is certainly true (as can easily be shown¹⁹) that agent relativity of every kind that has been proposed in the literature (by such authors as Bernard Williams, Thomas Nagel, and Derek Parfit²⁰) can be accommodated *within* a consequentialist structure, once position relativity in the evaluation of states is admitted and used. But the *force* of the deontological considerations may not be fully captured in this way,

¹⁸ See my "Evaluator Relativity and Consequential Evaluation," *Philosophy and Public Affairs*, xii (1983): 113-132.

¹⁹ Discussed in my "Rights and Agency," *Philosophy and Public Affairs*, xi (1982): 3-39.

²⁰ Williams, "A Critique of Utilitarianism" (1973), and *Moral Luck* (1981); Nagel, "Limits of Objectivity" (1980); Parfit, *Reasons and Persons* (1984).

since part of that argument rests on distinguishing between the nature of action judgments and that of state judgments, no matter how inclusively (and position-relatively) the states are judged. The approach to moral reasoning utilized in these lectures does not force us to take a position one way or the other on this question. The welfarist part of utilitarianism (scrutinized in this lecture and the last) and the requirement of sum ranking (examined in the last lecture) stand rejected in a way that consequentialism is not; but neither is it positively endorsed.

However, the exacting demands of consequentialism must be distinguished from the requirement of "consequence sensitivity," and a firmer view can be taken about the latter. The analysis presented in these lectures would certainly be hard to reconcile with moral approaches that try to judge actions in a consequence-*independent* way. Given the intrinsic importance of well-being, and indeed of agency, it is not credible that a person can morally evaluate his or her actions without taking note of their effects on the well-being and agency aspects of others (including their well-being freedom and agency freedom).

Even the attempt to get a limited sphere of consequence-independent evaluation of actions by declaring some particular types of action to be invariably rejectable, no matter how good their consequences are, leads to problems. This "openness" is relevant even in assessing acts that are terrible judged on their own, e.g., killing. In fact, even in Williams's example involving Jim's problem, the dilemma would amount to nothing at all if some unexceptionable deontological constraint were definitively to rule out Jim's killing anyone. Whereas a nonconsequentialist Jim is *not* obliged to go ahead and shoot, simply because the resulting state of affairs (including his having killed someone) would be better, the dilemma with which Williams is concerned arises also because Jim cannot decide, in an equally straightforward way, that he *must not* shoot under any circumstances. The decision is difficult because there *is* a case for doing *either* alternative, and no unexceptionable constraint. That indeed is precisely the problem. Consequences *are* relevant even if consequence judgments do not enjoy a hegemony over action judgments.

V. RIGHTS AND NEGATIVE FREEDOM

This issue becomes particularly relevant in examining the possibility of having a substantive moral system that gives absolute priority to certain constraints related to rights, as, for example, in the system explored by Robert Nozick in *Anarchy, State and Utopia*. In such a system rights impose unrelaxable constraints, which have the effect

of excluding certain alternatives. People must obey these constraints, no matter what else they may or may not do.

I have discussed elsewhere²¹ why we may have good grounds for violating someone's rights if that prevents even worse consequences, e.g., a worse violation of the rights of others, arguing that this type of consideration is inescapable given the problems of interdependence that moral analysis has to face. For example, the violation of the property rights of some can prevent a developing famine and decimation.²² Or, to take a different type of example, the violation of the rights of privacy of someone may yield the necessary information to prevent someone else from being battered, kidnapped, raped, or murdered. It is not at all clear why violating property rights, privacy, etc., must still remain rejectable, everything considered, in the cases specified. This necessity of considering, as it were, "everything" makes the proposals of consequence-independent absolute rights hard to defend. And the considerations of well-being and agency discussed in these lectures are impossible to reconcile with any proposal of consequence independence.

The nature of the rights considered by Nozick raises two different types of question. One concerns their *force*, in particular, their irresistibility; and the other relates to their *form*, namely their focus on "control" and more specifically their concentration on the "negative" perspective—imposing constraints on what others are permitted to do. The question of force, which has already been briefly addressed, must be distinguished from the issue of the appropriate form.

Considerations of well-being freedom and agency freedom need not, of course, have this negative concentration. Concern for well-being, well-being freedom, and agency freedom can all figure in a system in which conflicting positive contributions have to be weighed against each other when we are faced with a choice. Some of the relevant freedoms can also yield straightforward notions of rights. For example, minimal demands of well-being (in the form of basic functionings, e.g., not to be hungry), and of well-being freedom (in the form of minimal capabilities, e.g., having the means of avoiding hunger), can well be seen as rights that command attention and call for support. This perspective, which is close to the traditional format of the pursuit of positive freedoms, emerges naturally from the analysis that I have tried to present in these lectures.

²¹ "Rights and Agency" (1982).

²² Discussed, and empirically illustrated, in my *Poverty and Famines: An Essay on Entitlement and Deprivation* (New York: Oxford, 1981).

Some demands of *agency* freedom can also be seen as rights. The importance of self-agency has easily recognized urgency, and the force of such considerations as autonomy and liberty relate to the capability to take decisions in certain personal spheres in one's own life and to choose for oneself certain basic features of one's personal existence.²³ Although, as has been argued earlier in this lecture, the formulation of these freedoms requires us to look beyond control at effective power, this reformulation does not change the reasons for which such rights are taken to be important. One's ability to decide on certain matters in one's own personal life is given general support, and this support is not confined only to valuing procedural control or merely to wanting others not to violate these controls.

It is, in this general context, legitimate to ask: Where does *negative* freedom fit into this picture? Or does it not at all? The identification of the *source* of negative freedom has to be distinguished from that of its relative *strength* and its comparative force in a pluralist approach. In particular, the possibility that the source may be entirely *instrumental* is not ruled out by the general recognition that considerations of negative freedom may limit other people's actions very forcefully. Giving people powerful rights of, say, immunity from interference by others may help in furthering various intrinsic goals, including that of enhancing well-being freedom and agency freedom. Focusing on instrumentality (*vis-à-vis* utility) is, of course, precisely the way Jeremy Bentham reconciled attaching importance, on the one hand, to negative freedom (and to institutionalized rights) and, on the other hand, also regarding "natural rights" as "simple nonsense" and "bawling upon paper," and "natural and imprescriptible rights" as "rhetorical nonsense, nonsense upon stilts".²⁴

It is, in fact, quite possible that much of the actual force of negative freedom and the rationale of institutional rights come from instrumental considerations of one kind or another. But without trying to rival the "instrumental" approaches through an alternative ("intrinsic") view that would yield these rights with *comparable force*, it is possible to argue that there are other—more intrinsic—sources of value in negative freedom. A violation of negative

²³ I have discussed some of these issues in "Liberty and Social Choice" (1983).

²⁴ For an interesting and useful analysis of Bentham's treatment of moral and natural rights, see Ross Harrison, *Bentham* (London: Routledge & Kegan Paul, 1983), ch. IV. The instrumental approach need not, of course, be utilitarian. On this see Scanlon, "Rights, Goals and Fairness," in S. Hampshire, ed., *Public and Private Morality* (New York: Cambridge, 1978), pp. 104/5.

freedom involves a direct failure on the part of the violator as a moral agent. And this failure of agency can be seen as having some significance of its own.

Consider a case in which person *A* pushes person *B*, who does not swim, into the river. *B* in fact drowns, and *A* cannot justify his action, we assume, in terms of some greater good that would follow from this drowning. It is possible to distinguish between two bad things that have happened. First, *B* is drowned and his positive freedom to live on and not to die prematurely has been violated. Second, *A* has failed rather badly as a responsible moral agent. So something worse than *B*'s drowning has occurred, especially as it should be viewed from the position of *A*.

This way of seeing the problem can appear to be erroneous, involving double-counting. Everything follows, it can be argued, from the bad thing happening to *B*. But is such an objection really sustainable? *A*'s failure is, of course, not unrelated to the bad thing happening to *B*, but there is a source of badness in *A* as well, in terms of his agency aspect, since he has failed so badly as an agent. Indeed, *A*'s failure as a responsible agent could have taken place even without the same bad thing happening to bedraggled *B*. For example, *A* pushes *B* with all his might, but unexpectedly *B* somehow does manage to stay on land; or *B* does fall, but is saved by the accidental presence nearby of a good swimmer. *B*'s positive freedom is, thus, very little affected, but *A*'s agency failure remains serious, if the saving happened to be unexpected and accidental.

In evaluating the states and actions, agent *A* cannot but take note of his own agency in the attempted drowning. His moral assessment must take note of both *B*'s loss, if any, and his own failure as a moral agent (if he were to take the step of pushing *B* into the river). Once the agency aspect of a person is given a central place in moral accounting, there is no escape from including this type of consideration in evaluating states and actions, and what looks like double-counting is, in fact, the union of two different, though not independent, bad things. The state in which *B* drowns accidentally—not pushed by anyone—cannot be as bad, especially as it should be seen by agent *A*, as the state in which *B* drowns after being deliberately pushed by *A* himself into the river. Thus, negative freedom does have some intrinsic importance of its own, *in addition to its instrumental role in the pursuit of positive freedom.*

There are many ambiguities in analyzing responsibility, and it is not my intention to suggest that this way of seeing negative freedom is unproblematic. But the coupling of agency considerations with well-being considerations does provide a much wider basis for

analyzing the moral value of various traditional concepts—negative freedom among them.

To prevent a possible misunderstanding, I should also say that this type of agency-based consideration is not specifically confined to negative freedom only. The agency failure may well take a different form. For a variant of this case, consider one in which *B* falls accidentally into the river, with swimming champion *A* standing close to the spot, doing absolutely nothing to save *B*. Though *A* has not violated *B*'s negative freedom, it can be argued that this is also a case of agency failure on the part of *A* and that the state of affairs is, in a morally relevant way, worse for agent *A* than it would have been had *B* drowned without *A* standing nearby, or had *A* tried to save *B* unsuccessfully (with, of course, exactly the same result, so far as *B*'s life is concerned). Agency failure is a more general consideration than failure to respect negative freedom, which does not stand, by any means, alone in supplementing considerations of positive freedoms. But the relevance of agency failure provides some reason for intrinsic interest, *inter alia*, in negative freedom, in addition to whatever instrumental relevance negative freedom may have.

VI. CONCLUDING REMARKS

In these lectures I have tried to explore some of the implications of seeing persons from two different perspectives, *viz.*, well-being and agency, neither of which can subsume the other. The substantive moral discussions drew on the examination of methodological issues in the first lecture, using an informational approach to ethical questions, and dealing with such matters as pluralism, incompleteness, positionality, agent relativity, and objectivity.

The adequacy of the foundational structure that Scanlon calls "philosophical utilitarianism" (more descriptively renamed here as "well-being as informational foundation"—WAIF for short) was examined and rejected. But that informational foundation provided one of the two major bases of moral reasoning used in the second and the third lectures. While the attractiveness of well-being as a foundation accounts for a good deal of the appeal of utilitarian moral approaches, it was argued in the second lecture that the identification of well-being with utility fails for each of the three main interpretations of utility (in the form of choice, happiness, and desire fulfillment). It was further argued that even if such an identification were, for the sake of argument, to be accepted, utilitarianism still could not be derived substantially from WAIF. Contrary to some claims, the issue of sum ranking is not trivial nor unproblematic.

More positively, it was argued that a more plausible view of well-being is provided by the accounting of functioning vectors, refined by considerations of capabilities to function. This approach yields partial orderings which are typically incomplete but eminently usable.

The third lecture has been largely concerned with examining the relevance of agency and the nature of freedom. Each of the two basic aspects of persons, viz., well-being and agency, leads to a particular concept of freedom. Using this dual base, I have tried to comment on a number of related issues concerning control, power, rights, consequences, agency, and positive and negative freedoms. Both well-being freedom and agency freedom have important—though disparate—relevance to the assessment of states of affairs and actions. And the informational base of the moral approach explored in these lectures cannot, in general, dispense with either agency information or well-being information, each of which has intrinsic interest—in the form of freedom as well as that of achievement. These lectures have been much concerned with examining these plural informational elements and their respective roles in moral valuation and judgment.

AMARTYA SEN

All Souls College, Oxford

ERRATUM NOTICE

An item was omitted in the Index to Volume LXXXI, December 1984. On page 806 of the December issue, the following item should have appeared:

TYE, MICHAEL: The Debate about Mental Imagery. 678 (11).

Professor Tye's article was indexed under 'Debate' on page 800 and under 'Mental Imagery' on page 803. The editors regret the error.

NOTES AND NEWS

The editors regret to report that Paul Seligman, Professor Emeritus of the University of Waterloo, died on March 2, 1985, after a long illness. Professor Seligman was born in Bingen, Germany, in 1903. In the late twenties, he was active in the cultural life of Frankfurt as an author and film maker, left Germany in 1936, and studied philosophy at the University of London in the forties. He taught at the University of Waterloo from 1963 to 1975, when he retired as Professor. In 1975, he received a distinguished teaching award from the Ontario Faculty Associations; he tutored and supervised students, wrote, and published until shortly before his death.