

JUSTICE

Justice is the first virtue of social institutions, as truth is of systems of thought.

John Rawls, *A Theory of Justice*

Consider a typical situation for the practicing evaluator. He is hired by a government to evaluate an educational program for disadvantaged students. How does he proceed? What does he look for? The possibilities are greater than might be suspected. He may administer a standardized achievement test and compare the scores to those of students who do not have such a program. He may create a special test, perhaps based on objectives, to assess certain areas of academic "deficiency." He may measure student attitudes, opinions, or self-concepts.

On the other hand, he may solicit teacher opinions, attitudes, and judgments about the program. He may observe how teachers and students behave in the classroom, in the halls, in the streets. He may record and analyze what teachers say about students, how they grade them, what standards they set. He may ask parents about the program, about the teachers, about the schools. He may examine parent participation in school activities or school efforts to involve parents.

AUTHOR'S NOTE: This chapter is an extensive modification of the ideas in "Justice in Evaluation," in G. V. Glass (eds.) *Evaluation Studies Review Annual*, Vol. 1, Beverly Hills: Sage Publications, 1976, and "The Role of Theories of Justice in Evaluation," *Educational Theory*, Vol. 29, No. 4, 1980. The latter was a response to Strike's (1979) critique.

He may solicit the opinions of employers about the content or results of the program. He may match program content against job skills or trace student job opportunities, or pay, or performance after the program. He may examine grades in later school years. He might ask subject matter experts or political leaders to judge the program. He may render his own judgments. He may hold a mock trial and have members of the public judge the program. There are many other things he might do.

What does the evaluator do? There are a number of influencing factors—his experience, his training, what the government wants, what the people in the program want, etc. One of these influencing factors is the conception of justice that the evaluator and the sponsor of the evaluation hold, usually quite tacitly. What the evaluator believes is right and the prevailing conception of justice significantly affect the evaluation.

Admittedly, the influence is complex and far from direct. The evaluation is not a deduction from the conception of justice. The connection is more vague, more implicit—but nonetheless influential for all that. The dominant conception of justice limits the approaches one takes, what activities one finds legitimate, what arguments count as significant. Conceptions of justice act more as broad frameworks of consideration rather than as internally consistent machines for deducing conclusions. They distribute the burdens of argument in particular ways.

Evaluation techniques and practices do not per se belong to a particular theory of justice, like utilitarianism or Rawlsian justice-as-fairness. For example, multiple regression is not per se utilitarian—but an evaluator who held that there were a few quantitative indicators of success would find himself using multiple regression over and over. Someone who held that the interests of various publics should be directly represented in an evaluation would find himself using interviews far more than would a utilitarian. It is not the technique itself that is utilitarian but the reasoning that leads to particular choices of what the evaluator should do.

The evaluator proceeds partly by intuition, and philosophy can sharpen the evaluator's intuitions and sensitivities. The evaluator's intuitions are not only of justice but of self-interest, and philosophy can sharpen the better intuitions by reflections on theories of justice. Not only can people justify evaluations by reference to conceptions and

theories of justice, they do so often, albeit in the indirect manner suggested here.

The evaluator must choose. If he holds that increased standardized test scores will mean more schooling, a better job, more money, and presumably more satisfaction for the students—in short, that test gains are in their interest—he may well use that as the indicator of success for the program. If he believes that parental participation is central, or that the interests of the students are best represented by their parents' opinions, he will conduct a very different evaluation. He may still employ standardized tests. How the interests of the student are represented in the evaluation, and whose interests are registered will result in significantly different evaluations. That different interests be represented is a matter of democratic theory, but there are many conceptions of justice consistent with a democratic viewpoint. For example, a neo-Marxist evaluator may well search for inequalities in the school setting, whereas a political pluralist may strive to represent the opinions and values of many diverse groups.

Evaluation is by its nature a political activity. It serves decision-makers, results in reallocations of resources, and legitimizes who gets what. It is intimately implicated in the distribution of basic goods in society. It is more than a statement of ideas; it is a social mechanism for distribution, one which aspires to institutional status. Evaluation should not only be true; it should also be just. Current evaluation schemes, independently of their truth value, reflect justice in quite varying degrees. And justice provides an important standard by which evaluation should be judged.

In this chapter I outline three conceptions of justice—the utilitarian, the intuitionist/pluralist, and Rawls's justice-as-fairness—and relate these to current approaches to evaluation. I try to assess some of the weaknesses and strengths of the conceptions and arrive at an overall judgment as to which is preferable. As Strike (1979) has pointed out, there is no single approach to evaluation determined by one's theory of justice. However, I contend that one's theory does limit possible approaches to evaluation and is used to justify these approaches (House, 1979).

Utilitarianism

Utilitarian ethics, according to Rawls (1971), stipulates that a society is just when its institutions are arranged so as to achieve the greatest

net balance of satisfaction as summed over all individuals. The principle of utility is to maximize that net balance of satisfaction. Utilitarianism requires that there be a common measure or index of satisfaction in order that quantitative calculations of utility can be made. In education, this common measure is almost always construed to be standardized test scores. It is the surrogate index of satisfaction.

The basic structure of an ethical theory is given by how it relates two basic concepts: the good and the right. In teleological theories—which utilitarianism is—the good is defined *separately* and *prior* to the right. The right then becomes that which maximized the good—be it satisfaction, excellence, or whatever. Rationality itself comes to mean maximizing the good, a concept with great intuitive appeal. Not maximizing the good is looked upon as at least inefficient, perhaps fuzzy-minded, and even irrational. It is in utilitarian ethics *immoral*.

In the systems analysis approach in education, for example, test scores are established as the index of the good. That which maximizes them is both the right and proper thing to do. The best educational programs are those which, of the available alternatives, produce the greatest gains in test scores. Experimental design, instrumentation, and statistics are means for efficaciously determining the “best” alternatives by logically and quantitatively relating the right to the good. Social worth becomes a matter of maximizing test scores. Most evaluation discussion revolves around the efficiency of the evaluation as a means for determining the certainty and magnitude of gain.

The effect of such a theory is to separate moral judgments into two distinct classes. The good, the “value judgments,” are a separate class of judgments and can be prespecified without reference to what is right. Traditionally, the “right” includes the problem of the distribution of basic goods. In modern economics, heavily dependent on utilitarian thought, there is a very great emphasis on the production of basic goods but less attention as to how those goods are distributed. Attention is focused on increasing total production rather than on distribution of the goods. In utilitarian evaluation the emphasis is on increasing test scores, not on how those scores are distributed.

Utilitarian justice means that total net satisfaction is maximized. If someone has greater desires and expectations than everyone else, those must be considered, desire for desire, equally with someone with fewer desires. In a sense, if one has greater expectations, one is entitled to more. It is even possible to take away satisfactions from those with less in order to appease someone who “needs” more. This trade-off is one

cause of the failure of modern economics in developing countries. The gross national product of a country may be substantially increased by new industries. However, the money or the industry itself is often tuned to meeting the rising expectations of the new middle-class desires while the expectations of those lower down remain unimproved. Increasing total wealth does not necessarily improve the lot of everyone. The trade-off of lower-class satisfactions for higher-class satisfactions is consistent with utilitarian justice. Total satisfaction is maximized.

Likewise in education and other social areas—the desires (needs, objectives, etc.) are often taken pretty much as given. We mount “needs assessments” to determine these needs as objectively and quantitatively as we can. In our measurement theory, based on individual differences, the needs are intrinsic and lie there waiting to be measured, prioritized, and balanced against other needs. If upper-class people have very great expectations for their children, all their desires enter in the needs analysis and demand satisfaction. Their fulfillment will be reflected in the common measure of achievement and all the basic goods and values dependent in society on those scores. Distribution across people is not nearly so important as the total sum. Maximize everywhere is the dictum.

Since it is only the final net score that counts, one person’s loss may be balanced by another person’s gain. It is just for upper-class students to maximize achievement scores and advance their social position—even though they put lower-class children to disadvantage. Everyone must be free to advance his own good; and since what is right is that which maximizes the good, upper-class actions are both right and good.

Since all desires are taken at face value and since expectations and desires are much higher in the upper social and economic classes, utilitarianism has an upper-class bias built into it as a system of justice. One man’s desire for a Rolls Royce is just as valid a claim as another man’s desire for a Ford. The greater demands of upper- and middle-class students are honored at the expense of the lower classes.

In utilitarianism it is essential that everything be conflated into one system of desires so that justice can be served. Everything must be reduced to a unitary measure so that comparisons are possible. The basic method of comparison is the classic utilitarian “impartial spectator.” By being both impartial and sympathetic to all parties, the spectator (needs assessor) can organize all desires (needs) into one coherent system of desires (set of objectives). He weighs and balances

the intensity of desires and gives them appropriate weight in the overall system (evaluates). An administrator or legislator (the ubiquitous decision-maker) then adjusts the limited means in order to maximize the satisfaction of those desires. The result is efficient administration with the highest possible maximum satisfaction.

The utilitarian evaluator's resemblance to the impartial spectator is strong. In modern economics and modern evaluation the impartial spectator relies very heavily on quantitative measures for his impartiality. Modern econometricians push for "hard data" to measure the effectiveness of social programs (Rivlin, 1971). These can be converted into the most efficient design for society. The ideal evaluator also sympathetically records desires through needs assessments and definitions of objectives. He assigns weights and impartially measures results.

All educational needs are conflated by the evaluator into one system by familiar research techniques. Rawls notes: "Utilitarianism does not take seriously the distinction between persons, mistaking impersonality for impartiality." But this reductionism is essential for utilitarian justice to determine what is good and right to do.

As I indicated in Chapter 3, the systems analysis, goal-based, decision-maker, and goal-free models of evaluation seem to be based on a utilitarian conception of justice. They all try to arrive at the greatest utility by maximizing a small number of variables, a goal set, a decision-maker's preferences, or weighing outcomes. They try to arrive at a judgment of overall social utility, which, in turn, leads to maximum happiness in society. As in economics, cost is a prime consideration since wasted money could be used elsewhere to increase happiness; that is, utility could be improved.

Utilitarianism assumes that the essence of rational behavior is the maximizing of individual satisfactions or individual utilities (Macpherson, 1966). Human essence is rational action which maximizes utilities. Since people's desires for all kinds of satisfactions are unlimited, and the means of satisfying them always scarce, the problem is to find the system that would employ the scarce means to produce maximum satisfactions. Maximizing utilities is the ultimate good.

Utilitarianism is based on the idea that everyone is trying to get the most one can. Maximizing utilities means that one finds an arrangement by which people get satisfactions with the least effort. This leads to problems of how to identify and add together the satisfactions or utilities that people obtain from different things, of how to compare them on a single measuring scale. This is necessary in order to say that

one assortment of utilities is the maximum one, to say that one set of utilities adds up to a larger total utility (Macpherson, 1966).

Among the utilitarian approaches to evaluation, systems analysis tries to identify a critical set of variables like social indicators; the goal-based approach takes the program goals as given; the decision-maker approach identifies the key decisions of the decision-makers as critical; and the goal-free evaluator identifies the actual program effects. Each model takes a slightly different approach to maximizing utility.

Utilitarianism has been a brilliant and productive mode of thought. But by taking demands as they exist and by insisting on a common measure of welfare, it often favors the higher social classes at the expense of the lower. When the common measure turns out to be indicators like standardized achievement tests, the problem is compounded, for the tests themselves may be socially and racially biased; are based on a theory of individual differences; have been used historically for the explicit purpose of selecting people into and keeping them out of social groups (Karier, 1973); and may be used to legitimize inequalities. Utilitarian justice no longer coincides with many moral sensibilities.

Pluralism/Intuitionism

The second theory of justice is the pluralist/intuitionist theory. This theory asserts that there is an irreducible family of first principles which must be weighed against each other by asking which balance of the principles is most just (Rawls, 1971). There is no higher principle, such as the principle of utility in utilitarianism, which can be used for determining the weights of the ultimate principles. Hence, there is a plurality of first principles by which judgments of justice are made, and no explicit method or set of priority rules for weighing these principles against one another.

Most "common sense" notions of justice are of this type, as are most formal philosophical doctrines. Usually there are distinct groups of specific precepts which apply to particular problems of justice, like taxation or criminal behavior. These specific precepts are applied more or less intuitively within defined areas, some people balancing the fundamental principles one way, and some another. The relative weighting of specific principles can lead to the most intense debates, such as in freedom of speech, even when people agree on the funda-

mental principles. Custom and the interests of the parties involved lead to different weightings.

Intuitive balancing of principles provides no way of judging the justice of customary practices themselves, however. The intuitionist hopes that people can agree on the principles once they are identified or at least agree to some procedure for arriving at weights. At the same time the pluralist/intuitionist claims there is no supreme ethical conception which underlies the weights. The complexity of moral facts is too great to be encompassed by such a principle.

The origin and content of the principles of justice are different for the schools of evaluation that I have labeled pluralist/intuitionist. In the art criticism model of evaluation, the educational critic applies principles which he has arrived at through his own experience and training in educational programs. Presumably, each critic would have his own set of principles which he applies to a given situation in an intuitive fashion. Like the art critic, the educational critic would arrive at a composite judgment of the program or policy.

The professional model uses a set of principles which are often agreed upon in advance. For example, *Evaluative Criteria* lists many criteria suitable for judging English programs, another set for judging math programs, etc. English experts or math experts apply these stated criteria—and perhaps others—in an intuitive fashion. Each expert balances the principles according to his own disposition.

Similarly, the Council on Program Evaluation (COPE) has a long list of criteria which it used to judge university departments. These criteria could, in fact, be subsumed under a few principles like quality of research and teaching. The Council then arrived at a set of judgments by intuitively balancing these principles. Usually, research was given a very heavy weighting in these judgments. In both the art criticism and professional models, the criteria, ultimate guiding principles, and weightings are derived from educational professionals.

By contrast, the quasi-legal and case study models of evaluation generally derive their principles and weightings from participation of the people involved in the program or in making decisions about the program. In the quasi-legal model employing a mock trial procedure, the jury and judge determine what criteria and principles will be used in reaching a decision. Exactly what principles are employed depends on how the jury, judge, or hearing panel are selected. They may be members of the public at large, as in a real jury, or decision-makers who must take responsibility for the program at issue.

In many cases, such as in a “blue-ribbon” panel hearing, the panel may consist of prestigious members of the public or prestigious members of a professional subgroup generally thought to possess authority in that area, such as a group of tax experts or evaluators or decision-makers. For example, the joint Dissemination Review Panel in HEW reviews evaluations of federal programs to determine whether these programs should be “validated” for dissemination to the rest of the country.

Finally, among the pluralist/intuitionists are those who value participation of the people involved in the program. Generally, this participation consists of collecting the viewpoints and opinions of various people about the program or policy at issue. The evaluator faithfully records and portrays their viewpoints. In this manner, the principles, criteria, and weightings of the people involved are used to judge the program. Of course, in selecting and emphasizing some aspects and viewpoints at the expense of others, an inevitable occurrence, some of the evaluator’s own principles come into play.

In portraying the judgments of various groups associated with the program, this latter approach is pluralist not only in the philosophical sense of judging on the basis of several principles but also pluralist in the political sense of representing different political interests. Of course, the judgments of the participants and the evaluator are almost always intuitively balanced.

The pluralist/intuitionist theory of justice is the most commonsensical of all, but it leaves the evaluation subject to whatever principles and weightings the judges, whoever they are, happen to employ. At one extreme, it threatens total command of the judging process by professional principles and at the other extreme, a complete relativism in which everyone’s opinion is presumed to be as good as everyone else’s. Neither extreme would seem to coincide with what most people identify as justice.

Justice-as-Fairness

The most recent of the modern theories of justice is that of Rawls (1971). Rawls has suggested two principles of justice by which social institutions and arrangements can be judged. Rawls calls his conception “justice-as-fairness.” Unlike utilitarianism, “justice-as-fairness” assumes that there is always a plurality of ends and a distinctness of persons so

that one cannot conflate all desires into one system. An agreement by which the good can be distributed and by which disputes can be settled is arrived at first. Then, once the principles of distribution are agreed upon, individuals are free to determine their own good and to pursue it—but always in accord with the agreed-upon principles which determine what is right.

In defining how these principles are arrived at, Rawls uses the hypothetical construct of “the original position.” Assuming that all people are morally equal beings and are accorded fair treatment (and do not know what social disadvantages and advantages they will have in the actual society), what principles would they willingly agree to in order to define justice?

Rawls’s general conception of justice is this: “All social values—liberty and opportunity, income and wealth, and the bases of self-respect—are to be distributed equally unless an unequal distribution of any, or all, of these values is to everyone’s advantage” (Rawls, 1971: 62).

The specific principles of justice for institutions are these:

First Principle

Each person is to have an equal right to the most extensive system of equal basic liberties compatible with a similar system of liberty for all.

Second Principle

Social and economic inequalities are to be arranged so that they are both:

- (a) to the benefit of the least advantaged, consistent with the just savings principle, and
- (b) attached to offices and positions open to all under conditions of fair equality of opportunity [302-303].

It is clear from these principles that justice-as-fairness limits other ends. It puts boundaries around the things one may do. One does not take desires and aspirations as given. Rather, desires are restricted by the basic principles.

Justice-as-fairness specifically precludes imposing disadvantages on the few for the advantages of many. The priority rules for the two principles of justice specify that the first principle always has priority over the second. Basic liberties are to be maximized without regard to

social and economic benefits. Only then may social and economic inequalities be allowed. According to the second principle, these inequalities are allowable *only* if they benefit the *least* advantaged in the society. Inequalities are not allowable if the least advantaged are not benefited. Nor can there be any trade-offs of basic liberties for social and economic advantages.

The second principle, the “difference principle,” singles out the position of the least advantaged in society to judge whether inequalities are permissible. For example, if high salaries are necessary to attract people to positions that would benefit the least advantaged, then the inequalities may be permissible. (This judgment is in contrast to the “efficiency principle” of utilitarianism. The efficiency principle allows redistributions of primary goods only to the extent that giving to one social group does not take away from another group. This puts heavy emphasis on the existing order of things. A structure is maximally or optimally “efficient” when it is impossible to make some better off without making others worse off. In Rawls’s scheme, justice-as-fairness is *prior* to efficiency as a principle in this conception. One maximizes the long-range expectations of the least-favored position subject to the restrictions imposed by the first principle.)

Assuming for the moment that one accepts Rawls’s theory of justice, of what import is it to evaluating, if any? Clearly, the two principles are so abstract so as not to determine evaluation. On the other hand, as an important distributive mechanism, it seems reasonable that evaluation might attend to justice-as-fairness. The basic liberties protected by the first principle are political liberties, freedom of speech and assembly, liberty of conscience and freedom of thought, freedom of person and property, and so on.

In addition, the most important primary good, according to Rawls (p. 440), is self-respect. Without self-respect a person will not see that his own plan of life, whatever it is, is worth pursuing. He is cut off from a basic meaning in life. A major argument against utilitarianism is that some may be forced to give up their expectations in order to benefit the general utility. Having to involuntarily lower their expectations for the sake of others reduces their own basis of self-respect. This is not permissible in justice-as-fairness.

The first principle might be applied to evaluation in two ways: in what the evaluator looks for and in how the evaluation is conducted. The basic liberties are guaranteed by the first principle, including the right to self-esteem. For example, a major complaint of those who

oppose testing in particular is that test scores lower the self-esteem of many children in school. Not only do critics claim that tests are racially and class biased, but radical economists (Bowles and Gintis, 1972-1973) contend that tests are used precisely for the purpose of legitimizing the hierarchical structure of the social-class system so that even the lower members of society feel they deserve to be where they are. A lowering of expectations is built in.

Whatever the general truth of these charges, in justice-as-fairness these are serious concerns and the proper concern of the evaluator. Often it is easier to determine the worth of a *particular* practice, and that is closer to the evaluator's actual job. For example, sending a child home with his I.Q. score pinned to his coat, now a practice in some large cities, may be repugnant to even die-hard testers. Other practices, such as beating children to raise test scores, are usually ruled out—even if they raise test scores dramatically. There are certain things that cannot be done in the name of the common good.

Applying the two principles presumably reduces, but by no means eliminates, intuitive judgments. After one has decided that self-respect is a right that takes precedence over other considerations, whatever the program or sponsors want, one is still faced with applying the principle. This entails a series of intuitive judgments, and the evaluator's job in applying criteria to specific circumstances remains a difficult one.

Nonetheless, choosing primary criteria to apply over most other criteria (unnatural though it is to those of us raised in an empirical tradition) provides a moral base that is more helpful than might first appear. Sometimes the insights are informative. For example, if everyone has a right to self-esteem, it also follows in a just society that those least advantaged cannot violate the self-esteem of those better off. Self-denigrating activities are not allowable. This assumes, of course, that the self-esteem of those better off is not based on the lowered esteem of those worse off.

If the information is injurious to the self-respect of the person involved, it should not be included in the evaluation report—no matter how it affects the good of the project. The rights of the individual are to be protected over the good of the project. Only in the case of trading off this right for another equally important right could this principle be violated. Of course, at issue in a project would be the *total system* of basic equal rights of all concerned. A despotic director infringing on the self-esteem of others would not have to be treated so gingerly. Trade-offs of basic rights are permissible.

After the first principle is applied fully, one turns to the second principle. The essence of the second principle is that social and economic inequities are just only when they are arranged so as to benefit the least advantaged in society. A "representative man" from the least-advantaged sector must prefer his prospects with inequalities to his prospects without them. The social structure is judged from a particular social position—that of the least advantaged. The disadvantages of those lower in the social structure cannot be justified by the advantages that may accrue to the rest of the society, as would be permitted in utilitarianism. Of course, this judging must be done consistently with the first principle. One cannot trade off basic liberties for social and economic benefits.

As with the first principle, the second principle might be applied in two ways: as a criterion for the program and as a criterion for the evaluation itself. In the Illinois Gifted Evaluation, we kept in close touch with the program staff and were quite fair on the basis of the first principle. In retrospect, the major weakness was that we did not investigate possible deleterious effects from grouping talented children together. We did investigate the effects on the gifted children themselves, but we did not investigate the effects on nongifted students nor consider the broader social impact on the class system. Admittedly, these are not easy questions to resolve, partly because they have not been asked often enough. But they should have been addressed more than they were.

In justice-as-fairness, inequalities in natural talents are recognized but are not regarded as necessarily entitling one to greater material expectations in society unless they are used to benefit those less well off. For example, one might allow greater recompense for physicians in compensation for the training they endure and in expectation that attracting more talented individuals to medicine would benefit those least advantaged as well as everyone else. But one is not entitled to great expectations simply because one has greater natural talents. In a sense, the contract arising from the "original position" is an agreement to share fortune and misfortune as in a family. Members of a family do not maximize their good fortune at the expense of the least-advantaged family member.

What difference might the second principle of justice make in an evaluation? Consider Coleman's equality-of-educational-opportunity study (Coleman, 1966), essentially a liberal attack on a social problem. Although not strictly an evaluation, Coleman's study was an attempt to

assess the lack of equality of educational opportunity, a kind of needs assessment on which government policy was constructed. Coleman was presented with the problem without any clear conception of what equal opportunity meant. As Ennis (1975) notes, one's conception depends on what one means by "education" and what one means by "having an opportunity." Coleman chose education to mean academic education and among five conceptions of opportunity chose "equality of results" as measured by a vocabulary test.

The study presumed a single measure of utility. The mode of analysis was multiple regression analysis with background and school variables regressed against the test scores. From this Coleman reached the conclusion that schools were not redressing inequalities as reflected in test scores. Presumably, basic skills as reflected in test scores were what blacks needed for equality. Socioeconomic and racial isolation was deemed to be the primary problem and desegregation was seen to be the answer. Busing was the tactic.

As analyzed by Ennis (1975), much of this was implicit in the letter of transmittal which identified racial isolation as the major deterrent to educational opportunity and in the conception of equality held by Coleman and the government authorities. I would add that the conclusion also was implicit in the methodology of the study and the theory of justice within which it was formulated. Ennis (1973) noted that a causal chain of reasoning is only partly empirical; the other part is value. Values enter in determining where one might interfere to make a difference. Liberals might see busing as an answer but reject family interference while conservatives would approve neither. Ennis noted, "What he apparently did was to select factors for study, changes in which he thought (1) would not violate some set of values (at least mostly his), and (2) might make a difference. He then did an empirical study of these factors in an attempt to see which ones did make a difference."

By this approach the researchers determined the measure of utility and those factors they found acceptable for interference that might affect it. The empirical part of the study was in determining which of the predetermined factors would affect the predetermined ends. Even the analysis was subject to value judgments in the determination of which variables should be entered first in the stepwise multiple regression. By entering background factors first, the covariance attributable to background-school interaction was not attributable to the school. In addition, the residual error can be attributed to almost anything else,

e.g., "luck" in Jencks analysis (Levin, 1972). I do not believe the Coleman study was poorly done; but even though highly quantitative, it was *necessarily* full of hidden value assumptions.

The Coleman approach closed off several other available policy options. It neglected entirely the attitudes, if not the interests, of the groups most intimately involved in favor of a general measure of utility. In this case, desegregation led to busing black children to schools of white ethnic groups who were only slightly above the blacks in social and economic status and to busing white children to ghetto schools. This threatened the ethnic whites' own precarious position. Coleman has now questioned busing (1975) and has said about his original study: "It is clear that for this purpose to be achieved, there should have been far greater attention to the reactions of whites with the economic means to move." The viewpoint of these people was not considered in the study.

There was even a greater omission, however: Neither was the viewpoint of blacks considered—and these were the people who the desegregation policies were supposed to help. Black viewpoints—or white ethnics' viewpoints—could conceivably have suggested alternative policies for social justice. Not only were the governmental policies removed from the influence of those most affected, the needs assessment study did not reflect their viewpoints and interests. However noble the intents of the policy-makers, the approach was heavy-handed in its paternalism.

In this case, those who paid the highest price for the social reform were the blacks and the white ethnics—the two least-advantaged groups in this setting. In a sense, the gain for the blacks was attempted at the expense of those who were most naturally antagonistic to them and who had the most to lose. The social groups who formulated and implemented the policy were substantially removed from being affected by it.

There was one additional injustice. Within the educational system, the costs of desegregation were borne most heavily by the least advantaged—the children, and particularly the black children who were subjected to the ravings of the racists. Even within its own terms, the original study proved inadequate because where desegregation did occur, test scores were not necessarily increased (Coleman, 1975). Apparently, important factors were excluded from the causal analysis. Finally, above everything, the approach kept control outside the hands of those at the bottom.

A pluralist/intuitionist approach to the problem might have collected the viewpoints of everyone involved. It would not have established predetermined ends nor eliminated alternative causal possibilities without inspection. In addition, a justice-as-fairness approach would make certain that the viewpoints of the least advantaged were given priority over the others. After the problem was examined from the viewpoint of the least advantaged, the next priority would be given to the second least advantaged—the white ethnic groups in this case.

Using Theories of Justice

None of these dominant theories of justice is entirely satisfactory as a basis for evaluation. Utilitarianism comes closest to being the official government philosophy. Where possible, government officials try to reduce everything to common measures of money. Cost-benefit analysis is the ultimate mark of this approach, and a concern about money is often a sign that the evaluator harbors this theory of justice. There is much to be said in favor of utilitarianism. It does provide a strong criterion by which public decisions can be made and justified. Its major weakness is that it tends to favor the upper classes over the lower and leads one to judgments that do not always square with one's moral sensibilities. Employing utilitarian measures often leads to oversimplification in results, as in the Follow Through evaluation.

The pluralist/intuitionist theory of justice is most consistent with every-day, common sense notions of justice. Most people employ multiple principles in arriving at judgments and balance these principles intuitively. The principles employed vary from person to person and from situation to situation. The principles used to judge the justice of taxation policies are not necessarily the ones employed to judge the justice of salary schedules. This squares with commonsense and every-day usage, but is it sufficient for evaluation of public programs?

Would one be satisfied in knowing that one evaluator would employ one set of principles but another evaluator a different set in evaluating a particular program, especially if the program were one's own? The very cornerstone of justice is the notion of consistency, the notion that similar cases be treated similarly. Consistency seems to be threatened by such an approach. At one extreme is the threat of relativism—that everyone's principles are as good as anyone else's. Hence, any judgment is as good as any other. This seems to make something of a shambles of

evaluation. Politically pluralist approaches that use the judgments of involved groups as the ultimate criteria for the evaluation are threatened by this possibility.

At the other extreme of the pluralist/intuitionist approaches, principles derived entirely from professionals are not likely to gain full acceptance by the public. A teacher's notion of treating her students justly may not square with the parents' notion. Which does the evaluator employ? Few people would be willing to let those who run the railroads have sole control over how that should be done. Similarly, professional approaches to evaluation in which professionals evaluate other professionals, whether they are teachers, surgeons, or engineers, have lost credibility over the last decade. The public has lost confidence in the professionals regulating themselves. It would seem that people are not willing to leave judgments of justice, which are essentially decisions about distribution, entirely to others.

In the most modern theory of justice, Rawls has provided a theory which is more egalitarian than utilitarianism and more absolute and determinate than intuitionism/pluralism. Rawls thinks it squares better with common sense. But justice-as-fairness has its own difficulties as a theory. Many criticisms have been made of it. One of the most common is that the absolute priority of the first principle over the second means practically that the difference principle would never be applied. Others claim that even the smallest advantage to the disadvantaged would justify taking everything away from the advantaged.

Nor is justice-as-fairness always consistent with one's intuitive thoughts about justice. Justice-as-fairness provides only for consideration of the disadvantaged; it does not make special provision for their participation in deciding their own fate, a direction I find highly desirable. The most telling logical criticism for the evaluation of public programs is that the collective good cannot be derived from individual wants in the way that Rawls does. The individual good is not equivalent to the collective good, an idea I shall develop in the next section of the book.

Nonetheless, my own judgment is that justice-as-fairness is superior to utilitarianism as a theory of justice, which is not to say that utilitarianism is worthless. Justice-as-fairness is deeper in moral sensibility. I would prefer a pluralist/intuitionist conception of justice over either, but only one in which certain values are specified. I will argue in the next section that an appropriate theory of justice would take into

consideration the values of moral equality, moral autonomy, impartiality, and reciprocity, as well as the aggregative principle of utility.

This judgment is not a universal denunciation of all utilitarian thought in evaluation. There are many instances in which utilitarian thinking is appropriate and does not lead us astray. For example, an evaluation of a plan to maximize use of hospital beds may revolve around the notion of efficiency to the exclusion of most other considerations. It is when there are other pressing considerations that utilitarian thinking goes awry.

Generally, as I have suggested at the beginning of this chapter, one's theory of justice does not determine the type of evaluation one does. There are too many other factors influencing an evaluation than that. One's theory operates implicitly and subtly. If one is a utilitarian, one will be led repeatedly back to considerations of utility and money. If one is a political pluralist, one will be led to portray the opinions of different groups about the program. If one is a Rawlsian, one will be led to considerations of the rights of individuals and the interests of the disadvantaged. If one is a neo-Marxist one will be led to uncover social and economic inequalities in the situation. One's implicit conception of justice is in the long run significant.

I suspect that each person's conception of justice is far more varied, idiosyncratic, situation-dependent, and incoherent than the great theories I have discussed here. Only a few people have the ability to articulate a coherent theory of justice; their names grace the backs of books. It is not necessary for the evaluator to subscribe to one of these grand theories (each of which is deficient) or to articulate his own. What is professionally responsible is for the evaluator to use these grand theories as signposts telling us where we are and where we want to go. They serve as guides for our direction.

PRINCIPLES OF EVALUATION

Thus to respect another as a moral person is to try to understand his aims and interests from his standpoint and to present him with considerations that enable him to accept the constraints on his conduct.

John Rawls,
A Theory of Justice, 1971, p. 338