

Individual and Societal Utility

DR. DAVID S. ALBERTS

*Coordinator, Computer Science Program
Assistant Professor of Operational Research
New York University*

DR. RICHARD C. CLELLAND

*Chairman, Department of Statistics
and Operational Research
University of Pennsylvania*

ABSTRACT

Inherent in all social actions is a possible conflict between the goals of society at large and the individuals affected. This paper explores the extent of this conflict in the case of a welfare program. A methodology for constructing two measures of effectiveness is presented—one based upon the goals of the individual and the other upon those of a democratic society. The two measures are computed in approximate form within the context of two actual welfare programs and the results compared.

Introduction

One of the important philosophical problems underlying all social welfare programs involves the possible conflict between societal goals, as expressed by the agencies carrying out the programs, and individual goals, as expressed by the persons receiving the programs' benefits. One can imagine an open society in which specific welfare programs would acquire specific target groups and

take their direction entirely from the goals of the members of these groups. In fact, some recent experiments^{1,2} in our society have been of this sort. On the other hand, in oligarchic societies, the aims of welfare programs have always been set by the ruling class. And in democratic societies today the goals of most welfare programs are determined by the funding agency without prior consultation with the beneficiaries of the program.

This question may be approached at two levels. The first is a matter of fact: Is there conflict between the individual goals of the participants in social welfare programs and the societal goals which the programs are attempting to serve? The second is a point for debate: Should there be such a conflict in a democratic society? This paper reports on an initial and very limited attempt to investigate the question of fact; we defer consideration of the second issue.

In particular, this paper outlines a methodology for constructing two measures of effectiveness for a social welfare program—one based upon the goals of the individual and the other upon those of a democratic society. These two measures are then computed in approximate form within the context of two welfare programs and the results compared. The comparison suggests that, in the case of these particular programs, there was disagreement between the two methods of evaluation. The application is viewed as illustrative; it cannot be generalized to other programs, and the details of each measure's construction and approximation are under continuing review. However, the problem is an important one, and the systematic approach given here is much needed in the evaluation of social programs generally. The results obtained for the specific programs evaluated are also of some interest. Other methods for evaluating welfare programs have been given by Levinson,³ Levine,⁴ and Jahn and Blenkner.⁵

Formulation of An Individualistic Function

Let us speak of a social welfare program having r participants. The program may affect a participant in two different ways—it may either change his environment or his responses to his environment. The environment consists of entities which we shall call *objects*; these may be tangible or intangible, and it is convenient to consider them as occurring in discrete units when one attempts to evaluate an actual program. We shall use the term *need* only in reference to a participant's need for one of these objects.

Let N_{ik} ; $i = 1, \dots, m$; $k = 1, \dots, r$ be the need of the k^{th} participant for the i^{th} object either as perceived by himself or as externally identified. Let $n_{ik}(t)$ be a measure of the degree to which each of these needs is satisfied at time t —the level of satisfaction of the k^{th} participant's i^{th} need. We define $S_k(t)$, the state that the k^{th} participant is in at time t , to be the set of all his levels of satisfaction.

$$S_k(t) = \{n_{ik}(t)\}; i = 1, \dots, m; k = 1, \dots, r.$$

The needs of an individual for various objects may differ. Let $a_{ik}(t)$ be the "relative value" of the i^{th} object to the k^{th} participant. A large a_{ik} indicates a strong need for the object in question while a small a_{ik} shows the opposite. We use the relative value of the several objects to the participant as weights in constructing a utility measure for his state at a given time, that is, we define the utility of that state, $U[S_k(t)]$, as

$$U[S_k(t)] = \sum_{i=1}^m a_{ik}(t)n_{ik}(t); k = 1, \dots, r.$$

The first way that a program can affect a participant is by changing his environment. His state then changes in response to the resulting shift in his level of satisfaction with respect to at least one object. The second way is by changing his responses to his environment that occur when the relative values of certain objects change, thus altering the utility of the participant's state. It should also be recognized that the relative value of a unit of a particular object varies with the number of units he possesses and, hence, with his state.†

Once the participant's utility function for his state has been constructed, it is a simple matter to form a statistic representing the effect of a particular welfare program. Let us speak of j programs, $j = 1, \dots, q$. Let $U_k(P_j)$ be the utility of the j^{th} program to the k^{th} participant. Let t_0 and t_* be times representing the beginning and the end of the program. Let $S_k(t_0)$ and $S_k(t_*)$ be, respectively, the states that the participant is in at times t_0 and t_* . Thus, we define $U_k(P_j)$ to be:

$$\begin{aligned} U_k(P_j) &= U[S_k(t_*)] - U[S_k(t_0)] \\ &= \sum_{i=1}^m a_{ik}(t_*)n_{ik}(t_*) - \sum_{i=1}^m a_{ik}(t_0)n_{ik}(t_0); k = 1, \dots, r. \end{aligned}$$

that is, $U_k(P_j)$ is the gain in utility ascribed to the terminal state over that ascribed to the initial state.

As a measure of a program's effectiveness we need some aggregative

† Suppose a participant has a craving for snuff. Assume that one ounce of snuff constitutes a unit and that the participant "needs" four ounces per week. At present he is getting one ounce per week. The utility of his state may be altered either by seeing that he receives two additional ounces of snuff per week (changing n_{ik}) or by educating him as to the deleterious effects of snuff so that the a_{ik} weight for snuff is reduced. In the first instance the relative value of the fourth ounce of snuff is likely to be much less than that of the second ounce.

measure of its utility for its participants. We shall use the arithmetic mean, but other averages are possible. $U(P_j)$, the utility of the j^{th} program, is defined to be:

$$U(P_j) = \frac{\sum_{k=1}^r U_k(P_j)}{r} \quad ; j = 1, \dots, q.$$

A comparison of the effectiveness of several welfare programs from the individual point of view can be based upon the quantities $U(P_j)$.

Formulation of a Societal Function

All societies attempt to provide some of their members with services designed to facilitate the attainment of the goals of these members. Most of the services society performs can be grouped under five headings: inspiring, manufacturing, distributing, providing, and protecting. Modern society operates primarily through governmental bodies although it also exerts strong influence through quasi-establishment organs such as the private enterprise system, organized religions, communications media, and educational institutions. The major distinction between a democratic society and other types is that, in theory, and to some extent in practice, the democratic form attempts to provide the above types of services for all its members not just for some of them.

A societal welfare function for a program P_j should measure the average change in level at which the five types of service are performed for the participants as a result of the program. In a democracy, it is possible to be more specific about these services and to consider the following statements as societal goals:

1. To inspire each member to seek only nondestructive[†] future states;
2. To create an adequate amount of each object so that it is possible for each member to have the objects he needs as part of his environment;
3. To distribute objects and information so that existing objects are accessible to all members who need them;
4. To provide each member with sufficient resources so that he can have the objects he needs as part of his environment;
5. To protect each member from denial of objects, denial of access to them, and denial of resources for possessing them as long as they constitute elements of his nondestructive future states.

[†] A future state for an individual is said to be nondestructive if it represents a gain in utility to him over his present state and if its realization does not decrease any other individual's utility for his state.

Since we choose to state the goals of a democratic society in these terms, we can now use them as a necessary background for evaluating a welfare program from the point of view of a democratic society. Let $v_{hk}(t)$; $h = 1, \dots, 5$; $k = 1, \dots, r$; be the degree to which the above five goals are realized for the k^{th} participant in program P_j . We define, as in the case of the individualistic function, the participant's state at time t :

$$W_k(t) = \{v_{hk}(t)\}; h = 1, \dots, 5; k = 1, \dots, r,$$

and a set of relative value weights $b_{hk}(t)$. The difference is that both $v_{hk}(t)$ and $b_{hk}(t)$ are measured and expressed by the administration of the program rather than by the participant. Then, the social utility of his state is given by:

$$V[W_k(t)] = \sum_{h=1}^5 b_{hk}(t)v_{hk}(t); k = 1, \dots, r;$$

and the societal utility of the program to him by:

$$V_k(P_j) = V[W_k(t_*)] - V[W_k(t_0)].$$

To determine the societal utility of the program one could again use the arithmetic mean of the $V_k(P_j)$, but these are good arguments against this procedure. Programs can have a large variation in their effect upon individual participants. Certain persons may be helped greatly while others may be harmed. We feel that in a democracy a program with a large variation in effect upon participants is less desirable than one which makes more equal contributions to their welfare. Hence, we shall use a weighted mean to define the societal utility of the program as a whole:

$$V(P_j) = \sum_{k=1}^r c_k V_k(P_j); j = 1, \dots, q.$$

where the weights c_k are chosen so as to penalize extremes. If we set all $c_k = 1/r$, we would, of course, be using the arithmetic mean again. If we chose c_k to be small for both large and small $V_k(P_j)$, we would be considering a program that greatly helps a small number of people as desirable as one that helps the same number of people very little. To avoid this we shall optimally choose c_k only to penalize for extreme values below the mean. Thus, programs which are extremely good for a few people are preferred to those which are extremely bad for few people, given the same mean effect. We shall refer to the set of all c_k as a *fairness function*. The general shape of an optimal fairness function is given in Fig. 1. Of course, a comparison of the effectiveness of several welfare programs from the point of view of society can be based upon the quantities $V(P_j)$.

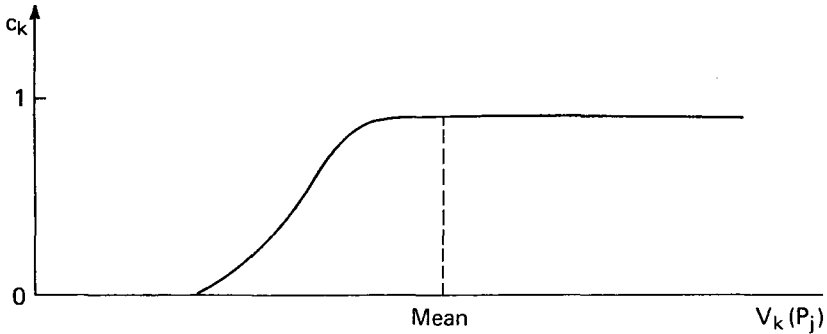


Figure 1. Optimal Fairness Function

Approximation to Individualistic and Societal Welfare Functions

The functions presented in the two previous sections are theoretical constructs which in the case of an actual welfare program would be either impossible or extremely expensive to implement. The measurement problems are obvious. However, they serve the purpose of forcing us to keep in mind what it is that we *should* be measuring. It is necessary to use rough approximations to these functions in most actual cases.

For the individualistic function it is necessary to limit the objects under consideration to some quite small set of particularly important items. It is also often desirable to combine objects when they are substitutable (i.e., soap-detergent; lamb-beef; coat-sweater) or complementary (i.e., coffee-milk-sugar; automobile-gasoline). Instead of experimentally determining each participant's level of satisfaction for each object, a standard level (one overcoat; one lavatory; 3,000 calories per day) can be set externally. This can also be done for the relative value of each object. And the whole evaluation may be based upon a sample of participants, particularly when the program affects a large number of persons or when data on some are unavailable.

In the case of the societal function, it is necessary to select measurable quantities—called indicants—which reflect the (unknown) degree to which each of the five societal goals previously specified is being realized. Levels of satisfaction are externally determined from these indicants. Since the appropriate $b_{hk}(t)$ are a matter of judgment and differ from program to program, one possible procedure is to set all $b_{hk}(t) = 1$, thus weighting the five services equally. Again a sampling procedure can be employed.

When it is impossible to obtain initial and terminal observations on the same set of r persons (as occurs when there is a high rate of turnover among participants or when an initial set of data is available from one program and a terminal set from a different but comparable program), one is forced to

proceed using a difference of averages instead of an average of differences. For the individualistic function, define:

$$\bar{U}[S_k(t_*)] = \sum_{k=1}^{r_*} U[S_k(t_*)]/r_*$$

$$\bar{U}[S_k(t_0)] = \sum_{k=1}^{r_0} U[S_k(t_0)]/r_0$$

where there are r_0 participants at t_0 and r_* at t_* .† Then we shall define the effectiveness of the program as:

$$U'(P_j) = \bar{U}[S_k(t_*)] - \bar{U}[S_k(t_0)].$$

For the societal function a comparable procedure can be used.

$$\bar{V}[W_k(t_*)] = \sum_{k=1}^{r_*} V[W_k(t_*)]/r_*$$

$$\bar{V}[W_k(t_0)] = \sum_{k=1}^{r_0} V[W_k(t_0)]/r_0$$

Notice that since in such a case we have no notion of the magnitude of any individual's gain in utility, we must take simple averages—that is, use a fairness function for which $c_k = 1/r$. Then the effectiveness of the program is defined as:

$$V'(P_j) = \bar{V}[W_k(t_*)] - \bar{V}[W_k(t_0)].$$

Other ways of approximating can easily be found. We shall be very specific about our approximations when we analyze real data.

The Data Base

We now turn our attention to the evaluation of a particular program. It was decided to apply these methods to a set of data gathered by the Research Center of the University of Pennsylvania's School of Social Work‡ during the period 1965-1966. The population under consideration consisted of all individuals living in the city of Philadelphia. These were divided into seven strata and four substrata, as shown on the following page:

† In practice, t_0 and t_* are also approximate. In the case discussed later, the timing of the questionnaire is such that the evaluation for one group occurs shortly after the program's start (t_0), while that for the other group occurs several years after the termination of the program (t_*).

‡ We are grateful to Dr. Julius A. Jahn, director of the study that collected these data, for making them available, and for many helpful suggestions.

- A 1: Former Clients of Intensive Service.
- A 2: Former Clients of Standard Service.
- B 1: Current Clients of Standard Service.
- B 2: Rejected Referrals of Standard Service.
- B 3: Current Clients of Other Agencies.
- B 4: Neighbors.
- B 5: Non-Clients, Non-Neighbors.

Group A 1, *Former Clients of Intensive Service*, are members of families included in the Intensive Service Project (ISP) of the Philadelphia Society to Protect Children (PSPC). They received this service from PSPC during the years 1959-1962.

Group A 2, *Former Clients of Standard Service*, are members of families who received service from PSPC during the period of time that ISP was provided, but did not participate in the ISP.

Group B 1, *Current Clients of Standard Service*, are members of families who "currently" (i.e., during a time interval that was defined to be about the time of the field experiment, 1965-1966) had been referred to PSPC for service and accepted.

Group B 2, *Rejected Referrals of Standard Service*, are members of families referred to PSPC, but not accepted for service in the same period of time as defined in B 1.

Group B 3, *Current Clients of Other Agencies*, are members of families or individuals accepted for service by other agencies (which agencies are sources of referrals for families in Groups B 1 and B 2 but not included in B 1 and B 2). In this particular case only one other agency was considered: Youth Service, Inc. (YS).

Group B 4, *Neighbors*, are members of families whose place of residence is in the same block as families in Groups B 1, B 2, and B 3, but who are not included in these groups.

Group B 5, *Non-Clients, Non-Neighbors*, are all other individuals in the defined population.

The substrata consist of the following age groups:

1. Adults who are parents of children under 18.
2. Older children (12 to 18).
3. Younger children (under 12).
4. Other individuals.

Note that groups A 1 and A 2 are mutually exclusive as are B 1, B 2, B 3, B 4, B 5, but that a member of one of the A groups might be included in one of the B groups because of the differences in the time periods involved.

Selection of the Sample

Within each of the agency strata (A 1, A 2, B 1, B 2, and B 3) all persons who were in these groups on the date of sampling were listed, and simple random samples were drawn. The sampling ratios were determined so as to obtain about 30 families in each of the seven strata. Cost considerations necessitated the reduction of the sample selected from stratum B 1 by one-half.

For the non-agency groups, B 4 and B 5, the sampling method involved, first, a random selection of city blocks listed in the 1960 United States Census, and next a listing of dwelling units within the specified blocks. A random selection of dwelling units from this list was made, followed by a random selection of households. One adult member from the selected households was predesignated for interviewing. The female parent was preferred; the second choice was the male of the household. For the children in the household, one from each of the age groups, 12-18, under 12, was randomly selected. The selection of the blocks differed slightly between these two groups. For B 4, the blocks were selected by first accumulating lists of all families registered† in the local Social Service Exchange. Next, a 1% sample was selected, and non-Philadelphia residents were excluded. A subsample of nine families was randomly selected; the blocks in which they resided were used. For B 5, the blocks were selected directly from the census listing adjusted for vacant blocks. The actual sample sizes are entered in the second column of Table 1.

Interviewers were selected from persons with some experience in social work with clients living in the areas covered by the survey. Two weeks of training were given to all interviewers. Interviews were carried out during the period between February and May, 1966. The interviewers were instructed to attempt to contact the sample families or persons on at least three separate occasions before classifying the questionnaire as incomplete. Table 1 also shows the number completed in each stratum; Table 2 shows the distribution of reasons for the noncompleted interviews.

In any statistical survey, observations sometimes cannot be made; this may introduce bias into the analyses. In this particular case, subjects have the right to refuse to be observed. Refusals accounted for less than 10% of noncompleted questionnaires for those individuals or families taking part in some welfare program; they accounted for 100% of the noncompleted questionnaires for those not participating in a program. Most of the questionnaires for participating individuals that are listed as incomplete reflect the fact that the people in question could not be located. It would

† All persons accepted for service at member organizations are registered. Virtually all important sources of aid are members of the Social Service Exchange.

Table 1. Percentage and Number of Research Schedules Completed by Stratum

<i>Stratum</i>		<i>Number Assigned</i>	<i>Number Completed</i>	<i>Per Cent Completed</i>
Intensive-Former	A 1	28	11	39
Standard-Former	A 2	33	12	36
Standard-Current	B 1	14	7	50
Rejected	B 2	25	5	20
Other Agency	B 3	36	34	95
Neighbors	B 4	23	9	39
Block Sample	B 5	26	16	62

have been desirable to resample refusals, but realistically this was not possible. For participants, refusals really are not a problem because of their small number and because 80% of them came from rejected applicants who would be disinclined to participate in the study. The individuals who could not be located are more likely to be a cause of bias. They may be unlocatable because their situation has either improved or worsened, resulting in their either moving or being forced to move. It is interesting to note that almost the same percentage of participants from each of the former programs was not located. Actually there appears no reason why noncompleted questionnaires from any of the groups should be much different from any other noncompleted ones. If any bias exists, it would seem fairly consistent over groups and to have no major effect upon the comparisons.

The number of schedules completed refers to families, thus, one completed schedule may also contain information collected from the head of the household and one or two children. In addition, some of the items that

Table 2. Distribution of Reasons for Noncompleted Schedules

<i>Stratum</i>	<i>Not Located</i>	<i>Refusal</i>	<i>Other</i>
A 1	17	—	—
A 2	20	—	1
B 1	6	1	—
B 2	12	4	4
B 3	2	—	—
B 4	—	14	—
B 5	—	10	—

we selected for our analysis were not filled out in some of the “completed” interviews. Thus, the numbers of schedules shown as completed in Table 2 do not correspond exactly in all cases to sample sizes presented later in this paper. In the analysis a distinction will be made between children and adults as the effects of a given program may differ considerably between these two groups.

A Basic Assumption

Why was this program selected for evaluation? Both the individualistic and societal functions previously developed are based upon two sets of data, one taken at the beginning of the program and one at its conclusion. Each measure of effectiveness is then a type of average change in utility over the duration of the program. But we are not aware of *any* social welfare program for which both initial and terminal data have been collected in anything like the degree of detail needed to approximate the individual’s state at times t_0 and t_* . Consequently, we used a program for which (1) a single extensive† set of relevant data had been collected, and (2) certain features of the study’s design allowed estimates of both initial and terminal states.

More specifically, in the Jahn study it is reasonable to assume that all clients accepted by an agency (strata A 1, A 2, B 1, and B 3) are drawn from the same population.‡ We shall concentrate upon them in an attempt to evaluate the programs for strata A 1 and A 2. The 1965 interview is regarded as initial for B 1 and B 3 and as terminal for A 1 and A 2. Then, the initial results for A 1 and A 2 are assumed to be the same as those obtained for B 1 and B 3. We emphasize the possible sources of error inherent in this procedure. Different persons make up the various strata. Therefore, we must measure effectiveness in terms of differences of averages $U'(P_j)$ and $V'(P_j)$ rather than the preferable averages of differences $U(P_j)$ and $V(P_j)$. Also the “terminal” evaluation takes place several years after the completion of the first set of programs.

Analysis

The questionnaire filled out by the interviewer was divided into 10 sections entitled: General Information, Housing, Occupation, Financial Situation, Physical Health, Adjustment, Interpersonal Relations, Activities and Interests, Respondent’s Comments, and Interviewer’s Analysis. For the construction of the individualistic function, four sets of “objects” were selected and their “units” defined. The sets were: Household Possessions,

† The questionnaire filled out by the interviewers in this study contains 68 pages.

‡ In the case of strata B 2, B 4, and B 5 this is either untrue or unclear.

Health, Self-Image, and Interpersonal Relations. The relative values attributed to each object were obtained from the interviewer's judgments shown in the last section of the questionnaire and expressed on a scale running from zero to one.

For example, the "Housing" need consisted of 13 objects identified during the interview. Each object, for instance, a refrigerator, could be either possessed or not. Certain objects, for example, a washing machine, were considered to be possessed if the individual had access to them. If the individual gave answers that signified possession of 11 of these objects, his level of satisfaction with respect to "Housing" would be calculated to be .85. Similar calculations were performed for each need. This same individual happened to receive the following estimates for the other three needs: Self-Image .95; Health .48; Interpersonal Relations .75. To obtain an overall measure of individual utility, the relative importance of each of these needs was required to serve as a weight when calculating a single measure of individual utility.

Ratings of each individual's current status with respect to these needs were obtained on a scale of 1 to 11 (poor to excellent) as perceived by the interviewer. The relative importance of each need was taken to be inversely proportional to the rating received. For example, if the individual received ratings of 2 for Housing, 5 for Self-Image, and 6 for both Health and Interpersonal Relations the calculations of the respective weights would be as follows. First, the complement from 11 would be taken for each need (9, 6, 5, 5). Then, the sum of these would be calculated (25); finally, each complement would be expressed as a fraction of the total (.36, .24, .20, .20). In this case the individual's utility would be calculated as follows:

$$\begin{array}{cccc} \text{Housing} & \text{Self-Image} & \text{Health} & \text{IPR} \\ (.36 \times .85) + (.24 \times .95) + (.20 \times .48) + (.20 \times .75) = .78 \end{array}$$

These calculations were performed for all individuals; the results by stratum and by the additional child-adult classification are presented in Table 3.

Since our main purpose in analyzing the data from the Jahn study is to compare the individualistic and societal measures of effectiveness for strata A 1 and A 2, we will now (Table 4) focus upon those strata. Entries follow from Table 3.

In Table 4 both A 1 and A 2 are representative of a program being measured at time t_* (after completion), while the combined results of B 1 and B 3 were used as estimates to provide information on t_0 (start of program). Participants in all four programs are considered to come from the same population. Both programs show an overall average gain. While the average gain is slightly higher for the Standard program, the Intensive program seems to improve both the children's and adult's situations while the

Table 3. Summary of Results Obtained by
the Application of the Individualistic Function
in Terms of the Values $\bar{U}[S_k]$

Stratum	Children		Adults	
	<i>r</i>	$\bar{U}[S_k]$	<i>r</i>	$\bar{U}[S_k]$
A 1	7	.71	11	.72
A 2	7	.75	11	.66
B 1	3	.78	6	.67
B 2	—	—	4	.66
B 3	28	.68	6	.69
B 4	1	.70	9	.73
B 5	4	.76	15	.69

Standard program seems to do best in improving the lot of children. In fact, adults on the average seem to be slightly worse off after completing the Standard program.

For the construction of the societal function it was first necessary to choose a set of indicants. The best available from the data at hand seemed to be the following five: Net Contribution to Gross National Product, Number of Crimes Committed, Level of Education, Level of Religious and Community Participation, and Level of Information. These represent a compromise between the available data and the five stated goals given previously. Nevertheless, it is true the Gross National Product (GNP) is a reflection of both the number of objects which members of society can possess and their collective ability to possess them. In this way GNP can be considered an indicant of the second and fourth goals—those concerned with production

Table 4. Comparison of Strata A 1 and A 2
Using the Individualistic Function

Program	Sample Size		$\bar{U}[S_k(t_0)]$	$\bar{U}[S_k(t_*)]$	$U'(P_j)$
	r_0	r_*			
A 1 Adults	12	11	.68	.72	.04
A 1 Children	31	7	.69	.71	.02
A 1 Combined	43	18	.68	.71	.03
A 2 Adults	12	11	.68	.66	-.02
A 2 Children	31	7	.69	.75	.06
A 2 Combined	43	18	.68	.72	.04

and provision. The number of crimes committed related to the fifth goal—that of protecting individuals from unlawful denial of objects and access to them. The level of education relates to both the first and third goals of society—those dealing with inspiration and distribution. The level of religious and community participation relates to the first goal—that of inspiring individuals to seek nondestructive ends. Finally, the level of information provides a further indication of the degree to which the third goal of society is being realized with respect to informational distribution. We will next consider the way in which the five indicants were computed.

First, the individual's contribution to GNP was estimated by breaking down his total income by source. Sources such as Income from Employment; Social Security; Pensions or Annuities; and Interest, Dividends, and Rentals were considered to provide positive contributions to GNP. Sources such as Public Assistance, Unemployment Insurance, Gifts or Contributions, and Support for Dependents from OASDI are considered as negative contributions to GNP. The crime rate or the number of crimes was based upon reported convictions.

The level of education was based upon years of formal education completed. The level of religious and community participation was calculated by using information about the number of organizations and committees in which the individual participated and the extent of his association in terms of the positions held and the time involved. The level of information was based upon a list of 27 activities including both those in which information can reasonably be expected to be transmitted and those in which information transmission is minimal. The number of an individual's activities in which information could be transmitted determined his level of information.†

To obtain a single-valued measure for the societal function, a weighted sum of the contributions of the five indicators was taken. Crime rate was transformed to "crimeless" rate so that all components indicated improvement by increasing. The needed weights were obtained by using the number of questions devoted to each subject as an indicant of the relative importance of each component. While many other weighting schemes are possible, we felt this one reflected the great experience of the authors of the questionnaire.

Table 5 summarizes the results obtained from the societal function, and Table 6 points out the results for strata A 1 and A 2 in a way comparable to Table 4.

Since the participants in program A 2 can reasonably be considered to be from the same population as those in A 1, these results indicate that A 1 is a more effective program from a societal standpoint.

† For example, knitting alone was considered to be devoid of informational transmittal, while writing or reading were considered to be activities in which information was bound to be transmitted.

Table 5. Summary of Results Obtained by the Application of the Societal Function in Terms of the Values $V[W_k]$

Stratum	Children		Adults	
	r	$\bar{V}[W_k]$	r	$\bar{V}[W_k]$
A 1	7	12.5	11	11.3
A 2	7	9.5	11	7.6
B 1	3	21.9	6	11.7
B 2	—	—	4	9.0
B 3	28	11.2	6	10.1
B 4	1	—	9	9.3
B 5	4	13.2	15	13.1

We note that while the individualistic function shows a greater increase for program A 2 (Standard) than for program A 1 (Intensive), the societal function shows a decrease for A 2 and no increase for A 1. Thus, if one were evaluating the two programs, the two functions yield conflicting results. Whether this is generally true, or whether it results from small samples and the major assumption that B 1 and B 3 may be used to estimate $\bar{U}[S_k(t_0)]$ and $\bar{V}[S_k(t_0)]$ is not clear at this time. Further discussion related to adults only is given in the conclusion to this paper.

Although the main purpose of this study was to compare the results obtained by the individualistic and societal measures of effectiveness for the various programs, it is of interest to examine differences which were obtained between pairs of strata within each measure. Those pairs of particular interest would represent differences in duration of contact with the programs. Table 7

Table 6. Comparison of Strata A 1 and A 2 Using the Societal Function

Program	Sample Size		$\bar{V}[W_k(t_0)]$	$\bar{V}[W_k(t_*)]$	$V'(P_j)$
	r_0	r_*			
A 1 Adults	12	11	10.9	11.3	.4
A 1 Children	31	7	12.2	12.5	.3
A 1 Combined	43	18	11.8	11.8	0
A 2 Adults	12	11	10.9	7.6	-3.3
A 2 Children	31	7	12.2	9.5	-2.7
A 2 Combined	43	18	11.8	8.3	-3.5

Table 7. Some Comparison Between Results
Obtained by the Individualistic Function
for Selected Pairs of Strata

<i>Higher Mean</i>	<i>Lower Mean</i>	<i>Sample Sizes</i>	<i>Difference in Means</i>	<i>Significance Level</i>
A 1 (adults)	A 2 (adults)	(11, 11)	.06	.125
B 1 (children)	B 3 (children)	(3, 28)	.10	.025
A 2 (children)	B 3 (children)	(7, 28)	.07	.040
A 1 (adults)	B 1 (adults)	(11, 6)	.05	.150
A 1 (adults)	B 2 (adults)	(11, 4)	.06	.100

presents results obtained using the Mann-Whitney test in comparing values of the individualistic function for pairs of strata. No other pairs yielded differences that were even close to significance at the $\alpha = .10$ level.

A given social action or welfare program cannot reasonably be expected to affect all or even most of these measures of social goals. In this particular case, the programs could not have been expected to affect the individual's contribution to GNP or their level of education. Of the other three, level of information, crime rate, and level of community participation the results

Table 8. Societal Measures of Community
and Religious Participation

<i>Program</i>	<i>Children</i>		<i>Adults</i>	
	<i>Sample Size</i>	<i>Mean</i>	<i>Sample Size</i>	<i>Mean</i>
A 1	7	15	11	18
A 2	7	11	11	7
B 1	3	25	6	11
B 2	—	—	4	5*
B 3	28	12	6	7
B 4	—	—	9	7
B 5	4	12	15	11

* Adjusted by deletion of extreme value equal to 101.

showed significant differences among strata only for the level of community participation. Table 8 presents these results for community participation; Table 9 presents the same pairs of stratum for comparison with significance levels computed by the Mann-Whitney test in Table 7.

Table 9. Comparison of Mean Values of Religious and Community Participation

<i>Higher Mean</i>	<i>Lower Mean</i>	<i>Sample Size</i>	<i>Difference in Mean</i>	<i>Significance Level</i>
A 1 (adults)	A 2 (adults)	(11, 11)	11	.04
B 1 (children)	B 3 (children)	(3, 28)	13	.025
B 3* (adults)	A 2 (adults)	(28, 7)	1	N.S.
A 1 (adults)	B 1 (adults)	(11, 6)	7	.15
A 1 (adults)	B 2 (adults)	(11, 4)	13	.10

* In reverse order from Table 7.

Final Remarks

From Tables 4, 6, 7, and 9 we note that, regardless of the measure used, the *adults* completing the Intensive program obtained higher values than those completing the Standard program. This is an indication of the effectiveness of the Intensive program. The weakness of the Standard program is shown by the fact that adults currently in the Standard program received higher values for both functions than did those who had completed it. However, those children completing the Standard program did receive higher values using the individualistic function than did those currently in programs of other agencies. This would seem to indicate either that the individuals who are clients of other agencies are worse off to begin with or that the effects of the Standard program are immediate or short-lived. Finally, those adults completing the Intensive program compared favorably, regardless of measure, with those who were accepted for treatment and are in the Standard program. Certain of these comparisons are statistically significant according to the Mann-Whitney test.

The most important substantive question posed in this paper is that of the agreement or disagreement of individualistic and societal measures of the

results of welfare programs. We have found in the case of the two programs studied (A 1 and A 2) that the two functions disagreed as to the order in which they valued the programs. Whether this result would hold if the measure were further refined is an open question. So is the whole matter of conditions under which the two functions could be expected to agree or disagree. We hope to investigate these questions further. We note without comment that if the study is arbitrarily limited to adults, both functions rated the Intensive program (A 1) higher than the Standard program (A 2).

We wish to repeat that the main purpose of this paper is to illustrate a method. If welfare programs are to be evaluated in a meaningful sense, this implies a formal procedure. We have outlined and applied such a procedure. The measures involved can be improved upon. The approximations used are very rough. Nevertheless, the idea that the contribution of a program is measured by the net change in a measure of utility is sound and basic. So is the idea of two points of view and two functions. We hope the illustration presented has served both to clarify certain problems related to the application of such measures and to indicate the formidable difficulties inherent in their use at this time.

REFERENCES

1. U.S. Department of Health, Education and Welfare, *Handbook on Programs*, 1963.
2. Office of Economic Opportunity, *Community Action Program Guide*, Vol. 1, February, 1965.
3. P. Levinson, "Evaluation of Social Welfare Programs—Two Research Models," *Welfare in Review*, March, 1966.
4. A. S. Levine, "Cost-Benefit Analysis of the Work Experience Program," *Welfare in Review*, September, 1966.
5. J. Jahn and M. Blenkner, *Serving the Aging Methodological Supplement, Part 2, Experimental Design and Statistical Tests of Hypothesis*, Community Service Society of New York, 1964.