OUTCOME MEASUREMENT IN NONPROFIT AGENCIES: DOING WELL AT DOING GOOD

BY ROBERT E. O'HARA

Presented to the Public Administration Faculty
at the University of Michigan-Flint
in partial fulfillment of the requirements for the
Master of Public Administration

March 9, 1999

First Reader _____

Second Reader _____

Introduction

Nonprofit, non-governmental organizations are big business in the United States. Nearly 181,000 nonprofit agencies which provide some type of service to Americans, without deriving a profit from these services, filed tax returns in 1995. Two decades earlier, the number of agencies completing this step was barely 82,000 (Chronicle of Philanthropy, January 28 1999). In 1996, more than $150 billion was given to charitable organizations in this country. Eighty percent of this largess was contributed by individuals (Auchincloss, 1997).

Despite the huge amount provided to charity, the amount given as a percentage of the gross domestic product has remained nearly the same for the last thirty years (Auchincloss, 1997). Nonprofit organizations, meanwhile, are beset by enormous challenges: Exponential growth in the number of charities, cuts in government social programs, and a growing demand for their services. It is both more important than ever, and more difficult, for nonprofits to solicit financial support from individuals (Bendapudi, et al., 1996).

People give to charity through various organizations. The largest recipient group of charity dollars is universities (Blum, et. al., 1998). Churches are also recipients of large amounts of donor dollars, as are hospitals and other medical institutions, United Ways and international aid organizations. The list of tax-deductible charity recipients, as defined by the IRS, extends down to community-based agencies to benefit the poor, and includes a wide variety of agency missions.

Why do people give to charity?

The reasons for giving gifts also vary widely. Donors may perceive a need for their support (Bendapudi, et al., 1996), or their motives may be to improve their own welfare (Martin, 1994; Del Prete, 1996). They may be

responding to interest group pressure, e.g., a corporation's gifts may be driven by the opinions of their customers and employees (Auchincloss, 1997; Stricherz, 1998). They may be exhibiting their own sense of wealth and power (Glazer, 1996; Harbaugh, 1998). College gifts are given to support the educational purposes of the institution, but also to support sports programs (Baade, 1996).

On a deeper level, Loseke (1997) argues that the idea of private charity gains its power from the moral effect and connotation it has developed, in relation to religion, politics, economics and human compassion.

Whatever the reason people have for giving, more people who give apparently desire to see concrete results from their philanthropy, and often give to specific projects, not for the general purposes of the recipient agencies (Auchincloss, 1997).

In giving to these specific projects, donors are often looking for specific *outcomes*. Outcomes are "benefits or changes for individuals or populations during or after participating in program activities" (Hatry, et al., 1996). Outcomes, then, are not simply the programs or services offered to clients, but the changes in clients' lives due to the programs. Hatry continues, "Consumers of services and volunteers who provide services want to know that programs to which they devote their time really make a difference. That is, they want better *accountability* for the use of resources. One clear and compelling answer to the question of 'Why measure outcomes?' is: To see if programs really make a difference in the lives of people...there is an even more important reason: To help programs improve services." (*Italics added*)

A simple way of characterizing this drive for accountability is to say that donors want to know that their money is being spent for good purposes. At times, the donors themselves might specify these purposes. At other times, they might indicate general purposes for the funds, and leave it to the agency to

develop specific purposes. These agency-based outcomes are likely to reflect the strengths and limitations of the agency, as well as client benefits. The clients of the agency probably have a third set of outcomes they would prefer. If the government (on any level) is involved in the program, there are often prescribed outcomes from this source, as well.

Given that each of these interest groups advances slightly different outcomes (sometimes grossly different), it is necessary to ask, "Who drives the outcome process?" That is, who identifies what the targeted outcomes of a certain program or activity are to be? What are the relative weights (relative both to each other and to the desired outcomes from other agencies, programs or activities) to be assigned to these various outcomes? And, what are the target values for each of these outcomes; that is, how much of each outcome does the outcome-identifier wish to require of the program or activity?

Who drives the outcome process?

Who drives the outcome process? Is it donors who dictate outcomes through the marketplace of donations, or agencies that sell their most effective and outcome-oriented programs to impressed donors? A review of related literature seems to come to the conclusion that, while agencies have a great deal of influence in setting the parameters of the discussion on any given human service issue, they are not free to run programs without a defined value. In order to effectively solicit revenue, and to develop continuing revenue streams, agencies must produce results (Witt, et al., 1995).

There are a number of possible actions agencies can take in response to this demand for results. The first, obviously, is to achieve the desired results. These results must, of course, also be sufficiently impressive to instill in donors the desire to support the agency. Meaningless results earn no support.

The second possible course of action is to convince donors that results are being achieved, even if this is not the case. Do agencies present a fictionalized tale of success if they cannot produce concrete results? Do they hide behind financial scams, because it is easier than making a good-faith effort to fulfill their mission? While this strategy sounds implausible, it has happened. John G. Bennett pleaded no contest in 1997 after he was charged with 82 counts of fraud, money laundering, and tax-code violations, and began serving a twelve-year jail term. Bennett's Foundation for New Era Philanthropy collapsed in 1995, after taking in more than $350 million from more than 1,000 nonprofit agencies and individual donors (Williams, 1998). Bennett was found to have been running an elaborate Ponzi scheme. He promised donors that a wealthy, anonymous donor would match their contributions dollar-for-dollar. This would have the effect of doubling the contributions of all donors. Agencies who gave money to the New Era Foundation were promised that they would receive double their money back. In truth, Bennett was living well off part of the proceeds of the foundation. Contributors to the foundation got only a small portion of their money back, if any, when the scheme collapsed.

The third possible course of action is to identify achievable and fundable outcomes and work toward those, regardless of whether those outcomes reflect the mission of the agency. This can result in a successful agency, if not an effective one. This course of action, in turn, could have two possible results. One is that the agency's nominal mission remains the same, while its actual mission becomes that of raising money. The second is that the agency changes its mission to one which is more conducive to raising funds. This may or may not be a good thing. Loss of that mission from the community of nonprofit agencies may leave a hole in the safety net of social services which has grown up in most communities (Moore, 1998).

A fourth possible course of action is to ignore outcomes, and simply encourage potential donors to feel they must support the agency, even if it has no firm goal. Bendapudi, et al., (1996) identify the factors necessary to develop a perception of need in a donor's mind. It is conceivable that an agency could manipulate these factors, while simply failing to deal with outcomes. Like Bennett's New Era Foundation, this strategy is likely to work only in the short term. However, it can work remarkably well. Gorman (1998) cites Thomas Sowell's theory that advocates of ineffective policies and programs simply argue that problems would have been much worse without their programs. Gorman argues that this seems to have justified increasing allocations to drug abuse prevention programs during the last decade, despite their failure to produce results. This is a prime example of an issue universally perceived as presenting a problem for our society, but in which increased financial support has not resulted in verifiable, concrete benefits. A related danger is that practitioners and researchers might take the approach of trying to use research to advance their theories, rather than as a tool for expansion of knowledge (Gambrill, 1995). In such a case, a partisan cynically uses money available for research to support thinly disguised propaganda.

Do agencies change their mission to produce revenue-enhancing results? Nonprofits regularly wrestle with the temptation to change their mission to produce revenue-enhancing results (Eisenberg, 1999). This commonly happens either because the agency stands to earn a large contribution through the change, or because the agency is in financial crisis, and needs to take drastic measures to improve its financial condition. Generally, the reasons given for the change are couched in more modest terms. The agency might say it is just "adding a program focus" or its "clients' needs have changed". Despite the tendency to downplay the effects of financial considerations, the choices in

these situations are often quite stark. In many other cases, the divisions are gray, rather than black and white, and agencies need to focus tightly on their mission, to avoid changing that mission in any way in order to increase revenue. With the management of agencies in the hands of a relatively small group of people (the common model), the lines between ethical standards and organizational survival often blur.

What is an effective nonprofit human service agency?

In order for a nonprofit agency to continue to receive support from its community and stakeholders, it is necessary that it establish a reputation as being effective and responsible. That is, the outcomes of the agency's work are important (relative to other agencies' work and to the needs of the community), it accomplishes these outcomes satisfactorily, and the agency proves and reports its outcomes. When such agencies receive public funds, many see such proof as a civic responsibility (Cook, 1986).

Green and Attkisson (1984) report that interest in statistical or scientific data regarding government policy issues first appeared around 1800. Beginning with medical services and spreading into welfare, education and community services, it expanded to military personnel screening and the evaluation of psychotherapy during the 1940's. The widest introduction of program evaluation, however, came during the Great Society programs of the 1960's.

Changes in our society over the years have been reflected by changes in the human services field, and especially in nonprofit agencies. Agencies that had previously been assumed to do good as a matter of course, were forced to

prove their effectiveness during the 1960's and 1970's. Some of these evaluations generated more controversy than knowledge (Peterson, 1987).

This demand for greater accountability has spread throughout the human service field during the last two decades (Green and Attkisson, 1984; Consumer Reports, 1997).

While the Great Society proved that program evaluation could be widely useful, federal belt-tightening during the Reagan Revolution of the 1980's discouraged large evaluation projects as wasteful. This happened in the same environment in which spending for anti-drug programs grew exponentially, and while government leaders and the public called for more accountability (Green and Attkisson, 1984). This seeming contradiction has required nonprofit agencies, especially, to try to do more with less.

Agency Competition and Targeted Money

The great increase in the number of nonprofit agencies in the United States during the last two decades has caused a serious rise in the amount of effort agencies need to put forth just to stay in business. Competition with other agencies has made them work harder and expend more resources to find money to support their work.

During the same time period, political pressures to solve such social problems as substance abuse and gang violence have resulted in large expenditures targeted for programs to fight these ills (though not to provide ongoing financial support to the agencies which implement the programs). The end result has been a catfight between agencies for operational dollars, and a political battle on the part of agencies and interest groups to secure funds to run closely-targeted programs of questionable value. The success of these last

efforts has not helped our society, in many cases, and may have actually caused harm.

Gorman (1998) basing his work on Thomas Sowell's book The Vision of the Anointed, argues that broadening efforts and adding programs to fight social problems is often little more than a waste of both time and money. He cites the case of the school-based drug prevention efforts of the late 1980's and early 1990's. As detailed earlier, despite early and significant proof that programs such as D.A.R.E. were having no effect, funding for them continued to grow. Gorman maintains that empirical evidence is largely irrelevant to the federal funding process in this case. Our nation's experience with substance abuse prevention programs seems to prove this point. Programs were funded not on proven need and effectiveness, but on perceived need and assumed effectiveness.

Despite these examples of wasted or ineffectual efforts to assess the value of nonprofit-based human service programs, the demand for proof that programs work, continues unabated. This is probably an outgrowth of the sentiment that, in the 1980's, called for less evaluation and more accountability.

Measures of effectiveness

The question of measurement of a program's effectiveness is not simply a question of whether to measure or not. There is a further question of what to measure. Many researchers argue that we must measure the effects of the processes used to serve clients (Schorr, 1989). Others assert that outcome measures are the correct approach, but we must be careful to articulate all important objectives and watch for unanticipated outcomes (Herman, 1987). Schorr believes that we must also include values (such as altruism, community and justice) in measurement as well, not simply quantifiable outputs or

outcomes. Regardless of the measures used, evaluation must meet the needs of the practice it is evaluating (Slonim-Nevo, 1997).

The most widely accepted measures of a program's effectiveness fall into two broad categories: Process measures and outcome measures. Two determining factors of outcome measures are the beneficiary of the service and the time frame of the program. Following is a brief description of each of these factors.

Factors in the Evaluation of Human Service Programs		
Process Measures	Outcome Measures	
	Beneficiary	Time Frame
Efficiency	Client	Immediate
Inputs	Funder	Short-term
Activities	Agency	Intermediate
Outputs	Community	Long-term

Process measures

Efficiency. Efficiency is the relationship of input to output. A low rate of input and a high rate of output denotes an efficient organization (Scioli, 1986). Measurements of efficiency can also include such factors as the demand for services (Yeh, et al., 1997). An efficiency study cannot tell, however, whether a program should exist or not; that is, whether there is a valid need in the community for the service. An efficiency study can only speak to whether the program can be more efficient. This is the first problem with efficiency measurements. Says Scioli, "...the efficiency model does not need to question the impact of the program on the community or its citizens. It questions merely

intended output in relation to program input" (Scioli, 1986). Comparing Scioli's efficiency model with Hatry's model of outcomes, efficiency is (by definition) not an outcome measure, but a process measure.

The United Way, for example, has traditionally asked agencies to provide the number of units of service provided by the agency. Comparisons of the units of service between agencies naturally followed, with a presumption that all units of service are relatively equal, and that efficiency can be extrapolated through comparison of agency statistics. This issue highlights the second problem with the measurement of efficiency; that is, such measurements often cannot compare apples with apples.

While many argue that efficiency is an outcome; that is, that efficiency is a required result for any successful program, this writer disagrees. The definition of outcomes, as stated earlier specifies that outcomes are measured by the difference in clients' lives, not the quality of financial management of the agency.

Quite to the contrary, efficiency is a factor in assessing effectiveness. A program which produces significant benefits to program clients, but at a high relative cost, will be seen as relatively less effective than an efficient program. Effectiveness is the only true outcome, as seen from a funder's point of view. There are, however, gradations of effectiveness, and an inefficient program would have lower value.

Inputs. Inputs include the financial resources necessary to support a program, but also include the professional skill to develop a program in an appropriate manner. This may include collaborations between agencies to provide a comprehensive solution (Dryfoos, 1996) or recognition of the specific characteristics and risk factors of a client group (Yowell, 1996).

Activities. Some researchers have focused on the importance of the activities of a program as a determinant of its success. Such researchers promote community youth development, for example, as a preferable objective compared to a tight focus on depressing specific problem behaviors or promoting micro-skills (Allen, et al., 1997; Jarvis, et al., 1997).

One question that should always be asked during an evaluation is whether the program was actually implemented. This concern can relate to inputs, or to activities. If the program did not receive adequate resources (financial or otherwise), then a lack of appropriate inputs dooms implementation plans. On the other hand, failure to correctly and completely carry out the planned activities of the program can also lead to failure (Rindskopf and Saxe, 1998; Gottfredson, et al., 1998). Regardless of the specific reason for failure to carry out activities, the community at large could see the complete program as a failure and no longer support it in any case.

Outputs. Outputs are the products of a program's activities (Hatry, et al., 1996), often related in a quantified manner, and thus confused with outcomes. An output might be the number of art classes provided by a youth development agency, or the number of counseling sessions provided by a substance abuse prevention program.

Outcome measures

The other major category of effectiveness measures is outcome measures. Outcome measures can be behavioral (Hatry, et al., 1996) or opinion-based (Stuntzner-Gibson, et al., 1995), client-centered or family-centered (Loff, et. al., 1987). The United Way, a leader in the nonprofit human

services field, takes a more limited view, seeing outcomes as only those results which can be measured, quantified and compared (Hatry, et al., 1996).

In this paper, I relate two factors of outcome measures: the beneficiary and the time frame in which the program is implemented. Clients tend to want immediate results from an agency or program; that is, they wish to use agency activities to complete their personal agendas. They wish no further agency intervention into their lives. Funders (and, often, managers) wish to see immediate results from their support, leading to a tendency toward short-term outcomes. Their focus, however, is often not on the personal desires and goals of the clients, but on a benefit to society as a whole. For example, clients of a Salvation Army food kitchen probably have, as their desired outcome, a meal. The funders, however, (in this case, the Salvation Army themselves) desire to save the souls of their clients, and ameliorate the problems which prevent this long-term outcome. These short-term challenges may be drug abuse, alcoholism, or other social problems. The Salvation Army, in this case, has a desired outcome of decreasing drug abuse. For this reason, clients of Salvation Army food kitchens are required to attend worship services or listen to gospel readings, in return for their meal.

Agencies, because they have a desire for ongoing funding (this is more cost-effective and efficient for the agency), take an intermediate-term look at the situation. The community in which all of this interaction transpires, requires positive long-term effects, and is relatively incognizant of, and uninterested in, shorter-term effects.

Beneficiary Focus

Beneficiary focus is one factor that presents a fault line along which agencies break their programs and outcome measurements. Child day care provides a good example of such a break.

Clients. Social service professionals often take a parent-focused view of child day care, seeing it as a useful service for parents. Day care providers, on the other had, focus their programs and their outcomes on the children who are their direct clients (Roditti, 1995). This divergence of focus is a problem because a focus on the child *as a unit of the family* has the effect of limiting prevention activities to family-centered activities. The outcome dictated by family-based treatment may not always be the most effective outcome for the child involved. This presents an obvious problem for program designers (especially when the social service professionals also act as funders of the program), as to the focus of the outcomes of the program.

Agencies. Agency outcomes may develop from yet another perspective. Agency personnel tend to believe that their programs are having a positive effect on their clients, regardless of whether there is objective evidence to back up this belief. This effect may be based on staff perceptions of their program's impact on their clients' risk factors, rather than on measurements of their clients' positive or negative behaviors. Impressionistic evaluations tend to focus on individual problem cases, and see greater effect on behaviors than can be established through measurement (Gilham, et al., 1997). Staff perceptions, which color agency evaluations, differ from client perceptions, which may differ as well from client behaviors. Such varied results in possible measures of program outcomes makes it difficult to compare outcomes between programs, much less between agencies.

Funders. Corporations, as program funders, sometimes favor certain projects, which may be driven more by the image left with the corporations employees or customers, than by the benefit to the direct clients of the service (Stricherz, 1998).

As an example of the influence funders can have on the outcomes of a program, I refer to the federal government. When Congress passed the Omnibus Budget Reconciliation Act of 1993, it included language intended to create a family-focused national system of childcare, through the individual states (Roditti, 1995). The federal system of childcare supports is an important part of many child care providers' revenue streams. This action, then, changed the mission for many of these childcare providers from service to their primary clients (children) to service to their families, which are represented by the children's parents. The same effect has been observed in youth development agencies, vis-a-vis drug prevention programs. Agencies are coerced, albeit often willingly, into defining agency objectives in drug prevention terms as opposed to youth development terms (Peyrot, 1991).

Community. The community in which a client lives and interacts is too often ignored as a client of the service-providing program (Jessor, 1993). Especially when a program has been presented to a donor as providing value to the community, the community should be looked at as a beneficiary of the program, and outcomes should be identified and tested to measure this benefit (Herman, 1987).

Time Frame

The time frame over which a program is studied is a factor in any evaluation. Programs may demonstrate short-term effectiveness, but be unable to sustain those results over time. The funder, meanwhile, might be expecting long-term change in the program's clients or community. On the other hand, a funder may have a short-term reason for wanting quick results, but the agency being supported works in longer-term program cycles.

Short-term. Our instinctive societal response to social problems is action-oriented and immediate (Levine and Zimmerman, 1996). Because of the nature of the political process, discussions on outcomes in the political arena focus on short-term outcomes (Schorr, 1989). This has a profound effect on the field of human services, because a large percentage of human services initiatives are funded through governmental programs (Kautz, et al., 1997). Because most evaluations focus on short-term impact of prevention programs, it is unlikely that major changes in behavior (community measurements of change) could be proven (Gilham, et al., 1997).

Long-term. Herman (1987) warns that a researcher or program planner needs to consider both short-term and long-term outcomes. This can cause problems in research design and fund development, as some outcomes may be so long-range as to be practically impossible to evaluate within a useful timeframe.

Gilham, et al. (1997) warn, however, that short-range programs and measures may actually create smaller change than long-term programs and measures. They also point out that those measures most susceptible to change are also those most easily observed by third parties.

On the other hand, long-term support of programs can result in colossal, expensive failures. The D.A.R.E. program appears to be an example. As of 1997, D.A.R.E. was still the most widely utilized school-based drug use prevention program in the United States. Early reviews of its performance showed no significant differences in drug use behaviors or attitudes between treatment and control groups (Aniskiewicz and Wysong, 1990; Wysong, et al., 1994). A review of its long-term effectiveness revealed that the most favorable assessment possible could only argue for a possible sleeper effect on the use of harder drugs by teenage males, with no significant effect among teens for the use of "softer" drugs, such as alcohol, marijuana and tobacco (Dukes, et al., 1997).

Other Factors in Outcome Measurement and Interpretation

One problem, for agencies that wish to measure the outcomes of their work, is that of capacity. Capacity is the ability of an agency to deal with problems or to administer work. Outcome research must accurately reflect the desired goal, must be useful, and must be carefully designed and evaluated (Bloom, 1995). This term relates to the current problem in two ways. First, agencies must have the capacity to develop programs that can be evaluated critically. If agencies do not have the professional capacity to understand the science of evaluation, they are unlikely to develop programs that can be usefully evaluated. Further, even if they are lucky enough to develop a program that can be evaluated, they will be unable to develop and implement a worthwhile evaluation. While professional researchers have the necessary skills and tools to conduct outcome effectiveness studies, most nonprofit agencies have not. Sadly, the financial condition of most nonprofit agencies is such that they pay

low salaries (tending to result in less-talented and less-educated agency professionals) and have little or no money to pay for professional evaluation.

The second way in which capacity effects agencies' ability to utilize evaluations is in their capacity to implement the changes dictated by evaluation results. Agencies must have the financial and professional capacity to allow flexibility and change, which can keep the agency working for the benefit of clients. Without this capacity, agencies are likely to become slaves to process, never changing to meet the altered needs of their client communities.

Those agencies that can afford proper evaluations also need to worry about focusing too narrowly on specific outcomes. A broad combination of outcomes is likely to have more lasting and greater effect on the quality of life of a client than will a single targeted outcome (Schorr, 1989). However, this same combination of outcomes is more likely to cause problems in the evaluation stage than is a single outcome (Bloom, 1995).

Case Studies in Outcome Measurement

Case studies of several outcome assessments may be useful, both to clearly identify the relationship between the various factors affecting outcomes, and to delineate the utility of outcome measurement.

Three case studies follow, covering a spectrum of outcome measurement efforts and agencies.

D.A.R.E.

Drug Abuse Resistance Education (D.A.R.E.) is a collaboration between local police departments and local school districts. As originally developed by the Los Angeles Police Department and the Los Angeles Unified School District, its curriculum was designed to build self-esteem, decision-making skills, resistance to peer pressure, to increase institutional bonding, and to delay the onset of experimentation with alcohol, tobacco and other drugs (Dukes, et al., 1997).

The following table represents the structure of the D.A.R.E. program, as it was originally conceptualized.

Outcomes of "D.A.R.E."			
Inputs	Activities	Outputs	Outcomes
Trained police staff	Class Sessions	Number of classes	Decline in drug use
School support	Resistance education	Number of students	Delay in first use
Facilities	Resistance skills Practice		Increased knowledge
Educational materials			

Dukes, et al. (1997) in a six-year effects study of D.A.R.E., found a possible sleeper effect of D.A.R.E. on 12th-grade males, as related to hard drugs. No other effects, on the most popular drugs or on a broader population, could be established. Both in their own studies and a review of other evaluative research, the researchers came to the hard conclusion that no significant effect of the D.A.R.E. program could be found, in terms of drug use. Some effect was perceived in terms of increased knowledge.

How, then, is D.A.R.E. to be rated? In terms of effectiveness, only one of three desired effects was found in the long term, and then on a relatively weak level.

If we review the one factor of effectiveness not yet mentioned in this analysis, that of efficiency, D.A.R.E. loses out on its very face. Using uniformed police officers as instructors, in the work location and working in conjunction with a certified teacher, hardly seems efficient. Both groups are relatively highly paid. A trained employee of a nonprofit agency could certainly operate the program more cheaply.

Why, then, is the program so popular? Perhaps because the inputs, activities and outputs of the program are so appealing. The program seems to make such good sense. Its basis is to use a strong authority figure, in a non-threatening environment, to teach a message of drug avoidance and resistance, as well as passing on resistance skills.

How could the program not work? It doesn't work because the program designers postulate all of the program's inputs, activities and outputs, and do so incompletely. From the perspective of many of the clients, other inputs include daily exposure to illegal substances and substance abuse in their neighborhoods (if not in their homes), threats on their safety if they resist involvement in a drug culture, and a wave of media support for alcohol and tobacco use. Activities include constant peer pressure to abuse these substances, and exploration of new experiences with their peer group. Outputs include an escape from the painful reality of their lives, increased stature among a subset of their peer group, and the possibility (for some) of monetary gain. These factors push the clients, 10- and 11-year old kids, to different and less happy outcomes. These inputs, activities and outputs are not accounted for in D.A.R.E.'s design.

<u>Nutrition Enhancement</u>

Nicklas, et al. (1998) undertook to assess the outcomes associated with a program to increase fruit and vegetable consumption among high school youths. This study featured both short- and long-term outcomes, as exhibited in the following table.

Outcomes of "Gimme 5"			
<u>Inputs</u>	<u>Activities</u>	<u>Outputs</u>	<u>Outcomes</u>
Money	Marketing campaign	Number of messages	Increased awareness
Educational support	Classroom activities	Five workshops	Increased knowledge, positive attitudes, and skills
	Parental involvement	Number of parent activities	Increased awareness & support of program by parents
Supplies	School meal modification	Additional exposure to fresh fruits and vegetables	Increased use of fruit by students Decreased cancer rate among subjects

The researchers found that the program did have its hypothesized effects. Awareness of the benefits of eating more fruit, and the practice of eating more fruit, increased significantly in the intervention group, as compared to the control group.

How, then, is Gimme 5 to be rated? Quite highly, if one follows the logical course of action based on its success. This program, however, received no

broad public support (as did D.A.R.E.). Coincidentally, the last year of the study coincided with the introduction of healthier menus into schools by the USDA. This had a broad effect on all students, both control and intervention, because the changes were universal in the schools. It might have also had an effect on public support, a topic that I will touch on in the conclusion of this section.

Lake Tahoe Watershed Project

The Lake Tahoe Watershed Project was a summer program for female middle school students in math and science. Its purpose was to "address some of the issues which confront self-concept and influence the choices of middle school age females" (Rohrer and Welsch, 1998). It was designed to "provide a non-threatening, all-female environment in which the participants could see and learn from female scientists and science teachers who were models of women successful in a math or science career." Rohrer and Welsch list the planned outcomes as being "for the students to demonstrate enthusiasm and evidence of experiencing success in daily math and science activities; exhibit evidence of positive interactions with peers while engaging in activities relating to science and mathematics; participate in the presentation of reports and supportive materials demonstrating their successful completion of the tasks described above; and demonstrate a continued interest in science and mathematics by continuing to enroll in courses prerequisite to college and subsequent careers in those disciplines."

In truth, only the last of these is truly an outcome, as shown in the following table.

Outcomes of "Lake Tahoe Watershed Project"

Inputs	Activities	Outputs	Outcomes
Money	Classroom instruction	Number of expressions of enthusiasm for scientific pursuits	Increased enrollment in science and math courses
Trained staff	Field study and experimentation		
Structured curriculum	Team building	Number of successful experiments	Increase in the number of female scientists and mathematicians
Facilities	Recreation		
		Number of positive interactions with peers	
		Number of reports presented	

As this analysis shows, most of the outputs of this project were presented as outcomes, and an apparent desired outcome of the project (an increase in the number of female scientists and mathematicians) was never presented. I conjecture that a possible reason for this exclusion is that the outcome could not possibly be measured within the limited period of the project. This is, as pointed out earlier, a weakness of outcome measurement: That long-term measures, even if possible, can be extraordinarily difficult and expensive.

How is the Lake Tahoe project to be evaluated? Rohrer and Welsch rate it as an unqualified success, based on clients' opinions of the activities. This conclusion, however, could not stand a true outcome assessment. The failure of this project to correctly spell out its outcomes is a serious fault. If outcome measurement is to be successful as a comparison and management tool across

the nonprofit world, standard definitions must be applied to all of the terms used in the description of outcomes.

These three case studies are accurate examples of the challenges facing nonprofits, even though two of them were carried out in an educational environment. Each of the two related to a characteristic of the educational system not directly related to its mission. That is, the Lake Tahoe project related to gender inclusiveness, not education in a pure sense; and the nutrition enhancement project related to meals. Each of the three related to a desired behavioral change in its clientele, a central characteristic of a nonprofit agency.

These studies, then, illustrate some of the challenges facing nonprofits as they enter the universe of outcome measurements:

- Nonprofits must correctly identify the outcomes for which the will strive.

- Nonprofits must ensure that their outcomes enjoy the support of both their clients and supporters.

- Nonprofits must ensure that the outcomes are significant enough to engender long-term support.

- Nonprofits must ensure that their client base is broad enough to develop political support for their programs.

- A special challenge, referred to in the case study of the nutrition enhancement project, is that government support of a program, or government mandates for program standards, can sometimes discourage private support.

Conclusions

What works best? It seems probable that outcome-based evaluation is not the final answer for determining the effectiveness of human service programs. An evaluation which is effective and useful over the long-term will require a high-quality plan and sufficient resources (inputs), a well-run program to offer the service (activities), clear client-centered and agency-centered goals (outputs) and results of value to the community (outcomes) (Hatry, 1996; Ostiguy and Hopp, 1995). A clearly defined relationship and program flow between the structural, process and outcome components of a program are required to ensure its success (Donabedian, 1980).

Schorr (1989) worries that America has tired of trying to solve our social problems. She lists six challenges facing our efforts to improve the quality of life of the disadvantaged: Knowing what works, proving we can afford it, attracting and training personnel, avoiding replication through dilution, moderating the effects of bureaucracy, and devising a variety of replication strategies. "Many Americans have soured on 'throwing money' at human problems that only seem to get worse...Fear of actually doing harm while trying to do good, together with the threat of unmanageable costs, have paralyzed national policy making", says Schorr. The issue of outcomes and their measurement directly affect performance in Schorr's first two challenge areas, and indirectly affects the last three. If America is to avoid the collapse of public support for community-based social programs, each of these challenges must be successfully met. Strict dependence on outcome measurement, without a drive for inclusion of the other factors necessary for success, could precipitate collapse (if outcome measurement models are followed closely) or could lead to sham assessments (if evaluators do not adhere to outcome measurement models). Either of these extremes would have a negative effect on the human services network.

The United Way, a leader in the nonprofit human services field, has thrown its full weight behind outcome evaluation of programs' effectiveness, as opposed to process evaluations. This means that results must be quantifiable, and not impressionistic.

While there is a significant body of research on the effectiveness of specific programs or treatment methods in human service agencies, there is a dearth of comparative evidence. "Performance measurement", say Cook and Eisenstein (1986), "assumes some type of comparative framework." The answers to such questions as the comparative value of programs focusing on effectiveness measures as opposed to efficiency measures, or the effect of administrative cost percentages on the outcomes of agencies, are lacking.

I would suggest the following questions for further research:

Is there a correlation between "efficiency" and "effectiveness" as outcome measures? This question appears to be inherent in the ratings given to agencies by management experts such as Peter Drucker, and financial advisors such as Money magazine. Salzer, et al. (1997), seem to point up service (customer perception of quality) as an important factor in differentiating between agency outcomes, given that relative costs are equal.

While some attempts have been made to do cost-benefit and cost-effectiveness analyses in human services (e.g., Cartwright, 1998), I found no research on the correlation between efficiency and positive program outcomes.

What is the status of outcome measurement in nonprofit agencies? While most resource providers and management advisors seem to favor outcome measurement, the actual level of its use in agencies is unclear. Some researchers suggest that the problem may be not just in the use (or lack thereof) of outcome measures, but in the misuse and manipulation of measured results, to satisfy political pressures (Brown and Kreft, 1998).

In the end, the question of which agency to support with one's charitable gifts remains a private choice, with personal criteria being a determining factor in the final choices. However, clear definitions of the various methods and objectives of programs, and their inter-relationships, may make the choice easier. This is the purpose of outcome measurement. The next step in the development of this field, should be to develop a model which requires outcome measurement, but also develops a successful process development component.

REFERENCES

1. Allen, Joseph P.; Philliber, Susan; Herrling, Scott and Kuperminc, Gabriel P. "Preventing Teen Pregnancy and Academic Failure: Experimental Evaluation of a Developmentally Based Approach". *Child Development* 68(4), August 1997, pp. 729-742.

2. Aniskiewicz, Rick and Wysong, Earl. "Evaluating DARE: Drug Education and the Multiple Meanings of Success". *Policy Studies Review* 9(4), Summer 1990, pp. 727-747.

3. Auchincloss, Kenneth. "The Land of the Handout". *Newsweek*, September 29, 1997.

4. Baade, Robert A. and Sundberg, Jeffrey O. "Fourth Down and Gold To Go? Assessing the Link Between Athletics and Alumni Giving". *Social Science Quarterly* 77(4), December 1996, pp. 789-801.

5. Bendapudi, Neeli; Singh, Surendra N.; and Bendapudi, Venkat. "Enhancing Helping Behavior: An Integrative Framework for Promotion Planning". *Journal of Marketing* 60(3), July 1996, pp. 33-52.

6. Bloom, M., Fischer, J., & Orme, J.G. Evaluating Practice: Guidelines for the Accountable Professional. Boston: Allyn & Bacon, 1995.

7. Blum, Debra E.; Demko, Paul; Dickey, Marilyn; Hall, Holly; and Marchetti, Domenica. "A Banner Year for Big Charities". *The Chronicle of Philanthropy* 11(2), November 5, 1998, p. 1.

8. Brown, Joel H. and Kreft, Ita G. G. "Zero Effects of Drug Prevention Programs: Issues and Solutions". *Evaluation Review* 22(1), February 1998, pp. 3-10.

9. Cartwright, William S. "Cost-benefit and Cost-effectiveness Analysis of Drug Abuse Treatment Services". *Evaluation Review* 22(5), October 1998, pp. 609-636.

10. "A Snapshot of Nonprofit Finances". *Chronicle of Philanthropy* 11(8), January 28 1997, p. 24.

11. "Charitable Giving: How to Make Your Generosity Count". *Consumer Reports* 62(12), December 1997, pp. 54-55.

12. Cook, Thomas J. "Symposium: Performance Measurement in Public Agencies (Introduction)". *Policy Studies Review* 6(1), August 1986, pp. 61-70.

13. Cook, Thomas J. and Eisenstein, James. "Measuring the Performance of Criminal Courts". *Policy Studies Review* 6(1), August 1986, pp. 137-144.

14. Del Prete, Dom. "Good Deeds Can Equal Many Happy Returns". *Marketing News* 30(10), May 6 1996, p. 11.

15. Donabedian, A. Explorations in Quality Assessment and Monitoring: The Definition of Quality and Approaches to Its Assessment. Ann Arbor MI: Health Administration, 1980.

16. Dryfoos, Joy G. "Adolescents at Risk: Shaping Programs to Fit the Need". *Journal of Negro Education* 65(1), Winter 1996, pp. 5-18.

17. Dukes, Richard L.; Stein, Judith A.; and Ullman, Jodie B. "Long-Term Impact of Drug Abuse Resistance Education (D.A.R.E.): Results of a 6-Year Follow-Up". *Evaluation Review* 21(4), August 1997, pp. 483-500.

18. Eisenberg, Pablo. "The New Philanthropy Isn't New – or Better". *Chronicle of Philanthropy* 11(8), January 28, 1999.

19. Gambrill, Eileen. "Less Marketing and More Scholarship". *Social Work Research* 19(1), March 1995, pp. 38-45.

20. Gilham, Steven A.; Lucas, Wayne L.; and Sivewright, David. "The Impact of Drug Education and Prevention Programs: Disparity Between Impressionistic and Empirical Assessments". *Evaluation Review* 21(5), October 1997, pp. 589-613.

21. Glazer, Amihai and Konrad, Kai A. "A Signaling Explanation for Charity". *American Economic Review* 86(4), September 1996, pp. 1019-1028.

22. Gorman, D. M. "The Irrelevance of Evidence in the Development of School-Based Drug Prevention Policy, 1986-1996". *Evaluation Review* 22(1), February 1998, pp. 118-146.

23. Gottfredson, Denise C.; Gottfredson, Gary D.; and Skroban, Stacy. "Can Prevention Work Where It Is Needed Most?" *Evaluation Review* 22(3), June 1998, pp. 315-340.

24. Green, Rex S. and Attkisson, C. Clifford. "Quality Assurance and Program Evaluation". *American Behavioral Scientist* 27(5), May/June 1984, pp. 552-582.

25. Harbaugh, William T. "The Prestige Motive for Making Charitable Transfers". *American Economic Review* 88(2), May 1998, pp. 277-282.

26. Hatry, Harry; van Houten, Therese; Plantz, Margaret C.; and Taylor, Martha. Measuring Program Outcomes: A Practical Approach. Alexandria, VA: United Way of America, 1996.

27. Herman, J. (Ed.) Program Evaluation Kit (2nd Edition). Newbury Park, CA: Sage Publications, 1987.

28. Jarvis, Sara V.; Shear, Liz and Hughes, Della M. "Community Youth Development: Learning the New Story". *Child Welfare* 76(5), September-October 1997, pp. 719-741.

29. Jessor, Richard. "Successful Adolescent Development Among Youth in High-Risk Settings". *American Psychologist* 48(2), February 1993, pp. 117-126.

30. Kautz, James R. III; Netting, F. Ellen; Huber, Ruth; Borders, Kevin; and Davis, Tamara S. "The Government Performance and Results Act of 1993: Implications for Social Work Practice". *Social Work* 42(4), July 1997, pp. 364-372.

31. Levine, Irene S. and Zimmerman, Jane D. "Using Qualitative Data to Inform Public Policy: Evaluating "Choose to De-Fuse". *American Journal of Orthopsychiatry* 66(3), July 1996, pp. 363-377.

32. Loff, Christine D.; Trigg, Linda J.; and Oassels, Christina. "An Evaluation of Consumer Satisfaction in a Child Psychiatric Service: Viewpoints of Patients and Parents". *American Journal of Orthopsychiatry* 57(1), January 1987, pp. 132-134.

33. Loseke, Donileen R. " 'The Whole Spirit of Modern Philanthropy': The Construction of the Idea of Charity, 1912-1992". *Social Problems* 44(4), November 1997, pp. 425-444.

34. Martin, Mike W. Virtuous Giving: Philanthropy, Voluntary Service and Caring. Bloomington, IN: Indiana University Press, 1994. (As quoted in Bendapudi, et al. [op. cit.])

35. Moore, Jennifer. "A Corporate Challenge for Charities". *Chronicle of Philanthropy* 10(22), August 13, 1998.

36. Nicklas, Theresa A.; Johnson, Carolyn C. and Myers, Leann. "Outcomes of a High School Program to Increase Fruit and Vegetable Consumption: Gimme 5 – a Fresh Nutrition Concept for Students". *The Journal of School Health* 68(6), August 1998, pp. 248-253.

37. Ostiguy, Lisa and Hopp, Robert. "Partners in Inner-City Youth Programming". (Leisure Today: Youth Program Success Stories) *JOPERD--The Journal of Physical Education, Recreation & Dance* 66(4), April 1995, pp. 32-33.

38. Peterson, Paul. "Evaluating Employment and Training Programs: Some Thoughts on the Lessons Learned". *Policy Studies Review* 6(4), May 1987, pp. 771-776.

39. Peyrot, Mark. "Institutional and Organizational Dynamics in Community-Based Drug Abuse Treatment". *Social Problems* 38(1), January 1991, pp. 20-33.

40. Rindskopf, David and Saxe, Leonard. "Zero Effects in Substance Abuse Programs: Avoiding False Positives and Negatives in the Evaluation of Community-Based Programs". *Evaluation Review* 22(1), February 1998, pp. 78-94.

41. Rohrer, Jane C. and Welsch, Sue. "The Lake Tahoe Watershed Project: A Summer Math Program for Female Middle School Students in Math and Science". *Roeper Review* 20(4), May/June 1998, pp. 288-290.

42. Roditti, Martha G. "Child Day Care: a Key Building Block of Family Support and Family Preservation Programs." *Child Welfare* 74(6), November-December 1995, pp. 1043-1068.

43. Salzer, Mark S.; Nixon, Carol T.; Schut, L. James A.; Karver, Marc S.; and Bickman, Leonard. "Validating Quality Indicators: Quality as Relationship Between Structure, Process and Outcome". *Evaluation Review* 21(3), June 1997, pp. 292-309.

44. Schorr, Lisabeth B. Within Our Reach: Breaking the Cycle of Disadvantage. New York: Doubleday, 1988.

45. Scioli, Frank P., Jr. "Problems of Controlling the Efficiency of Bureaucratic Behavior". *Policy Studies Review* 6(1), August 1986, pp. 71-89.

46. Slonim-Nevo, Vered. "Evaluating Practice: the Dual Roles of Clinician and Evaluator". *The Journal of Contemporary Human Services* 78(3), May-June 1997, pp. 228-239.

47. Stricherz, Mark. "Any Volunteers? Why Charity Is Not Enough". *The New Republic* 218(1), January 5 1998, pp. 12-13.

48. Stuntzner-Gibson, Denise; Koren, Paul E. and DeChillo, Neal. "The Youth Satisfaction Questionnaire: What Kids Think of Services". *Families in Society: The Journal of Contemporary Human Services* 76(10), December 1995, pp. 616-624.

49. Williams, Grant. "New Era Founder Took Comfort – and Cash – From Former Treasury Secretary". *Chronicle of Philanthropy* 10(18), June 18 1998.

50. Witt, Peter A.; Crompton, John and Baker, Duane. "Evaluating Youth Recreation Programs". *JOPERD -- The Journal of Physical Education, Recreation & Dance* 66(4), April 1995, pp. 27-30.

51. Wysong, Earl; Aniskiewicz, Richard; and Wright, David. "Truth *and* DARE: Tracking Drug Education to Graduation and as Symbolic Politics". *Social Problems* 41(3), August 1994, pp. 448-452.

52. Yeh, Joy; White, Kenneth R. and Ozcan, Yasar A. "Efficiency evaluation of community-based youth services in Virginia." *Community Mental Health Journal* 33(6), December 1997, pp. 487-499.

53. Yowell, Constance M. and Gordon, Edmund W. "Youth Empowerment and Human Service Institutions". *Journal of Negro Education* 65(1), Winter 1996, pp. 19-29.